THE AI TOOLCHAIN

NO. 075

Tail

THE DAILY RELEASE FIREHOSE

PUBLISHED OCTOBER 27, 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 | PREVIEW PANE

INFORMATION DENSITY

everything



BRIEFED | EVERYTHING

Every tool in the issue, in full.

all 482 changes

Do you prefer **scroll**? 👍 👎

▸ // CONTENTS SHOW 178 tools

▾ // THE BRIEF HIDE

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. Graphicle cut median `graphicle context` latency from ~3s to ~1.3s and peak memory from 3.5GB to B on a 30,315-file Kubernetes clone. Claude Code added a signed

vim ? help

model catalog for Bedrock/Vertex/Foundry. And Mistral shipped Mistral Large 4 and Voxtral 3, several with 1M-token context windows.

BUILD

Ask a code graph a question on a huge repo without a 3.5GB memory spike

Graphicle's per-episode query file lets a query decode only the slices it needs instead of loading the whole graph, which is what turns ~3s/3.5GB into ~1.3s/0.6GB on a 30,315-file clone. rails-ai-context's `get_schema` now covers PostgreSQL exclusion constraints, enum types, check constraints and generated columns, so an agent reasoning about a Rails schema stops guessing at the constraints the old dump silently omitted.

Graphicle · rails-ai-context

BUILD

Let a running agent survive its MCP servers changing underneath it

yoagent detects when a stdio MCP server process has exited and swaps tool sources mid-run when a server's tool list changes — previously a dead server surfaced as a mystery tool failure mid-task. pi-mcp-adapter adds async MCP Tasks and semantic tool search, so an agent facing hundreds of registered tools retrieves the relevant few rather than stuffing every schema into context.

yoagent · pi-mcp-adapter

GOVERN

Pin which models a deployment is allowed to call, and say why a tool call was refused

Claude Code's signed model catalog for Bedrock/Vertex/Foundry gives self-hosted deployments a verifiable list of permitted models instead of trusting whatever identifier a config file names. pi-mcp-adapter's structured deny reasons turn a silent tool refusal into something a caller can log, branch on, or show a user.

Claude Code · pi-mcp-adapter

BUILD

◆ AI Coding Agents

Anysphere Cursor

Feedback ⓘ snapshot-20261027 NOTES ↗

Cursor is a built to make you extraordinarily productive, agents turn ideas into code. Accelerate development by handing off tasks to Cursor.

Cursor launches the Agent SDK — a CLI and folder-based framework for defining, testing locally, and deploying code-driven bots to Cursor-managed infrastructure.

Introduces the `@cursor/agent-sdk` CLI: bootstrap a new agent project with `npx @cursor/agent-sdk init ./my-bot`, then iterate locally using `agent-sdk dev` to spin up a bot with a live playground.

Adds `--template` flag to `agent-sdk init` to scaffold from built-in starting points: security reviewer, owners-policy approver, PR autofixer, Linear triager, or Jira triager.

Adds `agent-sdk deploy` to publish an agent to Cursor-managed infrastructure, giving it a URL and access to dashboard-managed secrets and integrations.

Adds `agent-sdk validate` to lint an agent project and `agent-sdk info` to print the discovered project structure.

Adds `agent-sdk eval` for assertion-based testing of bot behavior — checks which tools ran, with what input, and the shape of the reply — runnable locally or in CI.

+2 more

NEW INTEGRATION Artemis Security integration

WAS THIS USEFUL? 👍 👎

yetone magpie

Feedback ⓘ v0.1.1710 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.

> Magpie's biggest shipments this window were an on-device PII-masking engine (with tunable whole-word matching and per-model confidence thresholds) and automatic quota failover across provider accounts, alongside several Codex- and DeepSeek-specific reliability fixes and a CoreML-accelerated PII encoder.

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

01 On-device prompt masking and tuning controls NEW 95

Adds on-device prompt masking (off by default), enabled via `Settings → Privacy`, that strips keys, passwords and emails from prompts before they leave the machine and reinserts them in the reply; values that must remain visible can be listed under `Never mask`. Adds a `WordsWhole bool` field to the `Options` struct in `internal/redact/redact.go` so masked words match only as whole words (`ABC` masks `'ABC'` and `ABC IS` but not `XABCD` or `ABCDEF`), exposed as `redactWordsWhole` in the settings API (`POST /api/settings`). Adds `redactModelScore` (a `map[string]float64` in settings, keyed by model ID) to set a per-model confidence threshold for local PII masking, with a `Confidence threshold` (`pmScore`) control in the privacy settings UI, visible once a local model is installed.

Enable on-device masking to prevent secrets from reaching vendor servers, then exempt a specific value that an agent needs to reason about.

> 📍 In the app, go to Settings → Privacy and toggle masking on. Then open the Masking page and add the token you need unmasked to Never mask.

Persist whole-word-only masking for a masked word via the settings API, so `ABC` is redacted in isolation but not when embedded inside `XABCD`.

```
$ curl -s -X POST http://localhost:<port>/api/settings \
  -H 'Content-Type: application/json' \
  -d '{"theme":"dark","redactWords":
["ABC"],"redactWordsWhole":true}'
```

— Names config keys, API endpoint, and matching rules with a runnable example v0.1.1699 v0.1.1689

02 CoreML acceleration for AstrLink Guard PII encoder IMPROVED

75

Accelerates AstrLink Guard's PII token encoder on macOS via CoreML, running on the Neural Engine or GPU — a 512-token window drops from ~400 ms (CPU) to ~12 ms on an M4 Pro; the compiled model is cached beside the ONNX weights and falls back to the CPU path automatically on failure or weight mismatch.

— Concrete before/after numbers and fallback behavior, but automatic v0.1.1694

---

03 **Automatic provider and gateway failover** NEW 70

Adds automatic quota failover so that when a provider account hits its quota limit, the next account in the group takes over and the agent never sees the error, and adds `WrongEndpoint` detection in the gateway so that when an OpenAI-compatible provider returns a 4xx error indicating a model is served on a different API (e.g. `/v1/responses` vs `/v1/chat/completions`), the gateway automatically retries the request against the next API the provider offers, using that API's key.

— Describes behavior only; no config or command named Names the exact endpoints and detection mechanism

v0.1.1710

---

04 **Per-provider vendor isolation for remote magpie** IMPROVED 65

Each provider behind a remote magpie instance now acts as its own vendor (via `remoteUpstream` and candidate.vendor()) — giving each provider its own conversation cache, reasoning seals, and compaction state, instead of sharing one account-level context across all remote providers.

— Mechanism and scope named but no user action required v0.1.1708

---

05 **Automatic Codex app-server cycling** IMPROVED 65

When magpie takes over a used-up ChatGPT account (or hands it back after reset), it ends the Codex app's own app-server process via `proc.End` so the app immediately starts a fresh one, re-reads `account/read`, and re-enables Send — on every platform, with no user action required.

— Names the exact call and endpoint but runs automatically v0.1.1704

---

06 **PWA install support for web UI** NEW 60

Serves `/manifest.json`, `/icons/apple-touch-icon.png`, and `/icons/icon-512.png` without requiring a session cookie, so mobile browsers can prompt Add to Home Screen and install the web UI as a PWA.

— Names the exact served paths and the install trigger v0.1.1690

THINNER COVERAGE BELOW

---

07 **Codex thread descriptions off setting** NEW 55

Adds `off` as a valid value for the `codexDescriptions` setting, answering thread-description requests with no description and sending nothing to any backend.

— Names the exact setting and value to apply v0.1.1698

08 **DeepSeek thinking-block signature preservation** IMPROVED 50

Preserves the `signature` field on DeepSeek thinking-block tool turns forwarded through the Anthropic gateway, so empty-signature thinking blocks no longer lose their `signature` key in transit.

— Names the field and path but is a background fix v0.1.1710

09 **Custom provider request attribution fix** IMPROVED 45

Requests made while a custom provider was being added (the Detect probe on the Add sheet) are now attributed to that provider, not to its bare hostname, so the Requests page no longer lists a phantom host-only entry alongside the real provider.

— Describes the fix and UI surface but no steps to use it v0.1.1697

NEW INTEGRATION SenseNova U1 image model support

NEW INTEGRATION PenguinHarness desktop app detection

v0.1.1707 CODE ↗ STABLE

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v0.1.1707 https://github.com/yetone/magpie.git
# already have the repo? check out this version:
$ git checkout v0.1.1707
```

v0.1.1696 CODE ↗ STABLE

Gateway keeps translated streams alive for up to 30 min of upstream silence; `pingEvery` throttles Anthropic pings to preserve Claude Code's watchdog budget.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v0.1.1696 https://github.com/yetone/magpie.git
# already have the repo? check out this version:
$ git checkout v0.1.1696
```

Adds `translatedQuietLongest` (default 30 minutes) to control how long a translated reply's upstream may be silent while magpie keeps the downstream client

alive via `keepThroughSilence`.

Adds `pingEvery` (default 1 minute) to rate-limit Anthropic `ping` events sent to a client, preventing Claude Code's 30-ping watchdog budget from being exhausted during long upstream silences.

WAS THIS USEFUL? 👍 👎

ogcode

Feedback ⓘ v0.53.0 NOTES ↗ RANK

An open-source, local-first AI teammate and agent operating environment for turning intent into completed work—with tools, memory, permissions, and human oversight.

> ogcode shipped persistent session-level memory, browser-integrated agent control through a new native-messaging bridge and ogweb Chrome extension, and a visual agents page with a relationship force graph, alongside one-shot scheduled reminders, multi-file read syntax, parallel tool calls, and adjusted tool-output limits.

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

01 Multi-file read syntax for change gate NEW

80

Adds `read files=[paths]` syntax to feed the change gate directly, letting you read multiple files at once and pipe them into the change gate for a focused diff review, e.g. `read files=[src/auth.go, src/session.go]`.

Read multiple files at once and pipe them into the change gate for a focused diff review.

```
$ read files=[src/auth.go, src/session.go]
```

— Exact syntax with runnable example command v0.52.0

02 One-shot scheduled agent reminders NEW

75

Adds one-shot agent reminders via a timer store and the `schedule` tool, polling every 10 seconds to trigger a follow-up task at a future time.

Schedule a one-shot agent reminder to trigger a follow-up task at a future time.

📍 schedule

— Names mechanism and poll interval with a usage example v0.52.0

03 Browser-integrated agent control via ogweb extension NEW

65

Adds a browser takeover bridge and native-messaging host plus the `ogweb` MV3 Chrome extension for browser-integrated agent control; the extension (`kacnfgnhmmnhdpokflkiemdehhjhkidp`) is allowlisted in boot-time native-messaging manifests so it can connect automatically at startup.

— Names extension ID and manifest mechanism, no setup steps shown v0.52.0 v0.53.0

04 **Session-level agent memory** NEW 60

Adds session-level memory — one per-conversation file stored under `.ogcode/memory/sessions`, shared with the whole agent tree.

— Names exact storage path but no usage example v0.53.0

THINNER COVERAGE BELOW

05 **Agents page with relationship force graph** NEW 55

Adds an agents page with a force graph visualizing loop wires, schedule gates, lazy browser toolset, intent, and teammate relationships.

— Describes visualization contents but no navigation detail v0.52.0

06 **Tool result size limits for read, file_map and grep** IMPROVED 50

Uncaps `read` and `file_map` tool results, which were previously subject to a cap, and adds a cap with a narrow-it notice to `grep` results when output is large.

— Names affected tools and before/after but no thresholds v0.52.0

07 **Parallel tool calls for provider endpoints** NEW 35

Adds `parallel_tool_calls` support to documented provider endpoints.

— Bare mention of a named flag, no mechanism or example v0.52.0

WAS THIS USEFUL?  

aerovato Operator Memory

Feedback ⓘ helper@v1.9.0

NOTES ↗ RANK

Operator Memory is a self-improving context engine for coding agents.

> Operator Memory's v1.9.0 release makes OpenCode V2 the default harness while freezing V1, and opens up adapter development documentation for the open-source community.

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

01 **OpenCode V2 promoted to default, V1 frozen** BREAKING

30

OpenCode V2 is now the default harness for Operator Memory, while OpenCode V1 is frozen (no further updates).

— Names the change but no migration steps or config details helper@v1.9.0

02 **Antigravity harness guides and guide standard** NEW

23

Adds new guides for the Antigravity harness along with a standardized guide format.

— Bare mention, no links or specifics on content helper@v1.9.0

03 **Adapter development docs opened to contributors** NEW

18

Opens up adapter development knowledge for open-source contributors, per the release notes.

— Very thin description with no concrete surface named helper@v1.9.0

WAS THIS USEFUL? 👍 👎

mvschwarz OpenRig

Feedback ⓘ v0.7.3 NOTES ↗ RANK

Build your own network of agents from Claude Code, Codex and Pi: persistent teams with roles, shared context and owned work.

> OpenRig expanded its gateway system with multi-person routing, Slack channels, and org/group policies, and gave seats live runtime switching between Codex and Claude Code, provider account switching, and refined per-seat permission controls — alongside new MCP handoff tools and tighter environment-variable scoping across releases.

└──▷ **WHAT SHIPPED · 17 FEATURES** most completely described first ? what's the number?

01 Multi-person gateway routing, Slack channels, and org policies

NEW 95

`rig gateway human add` registers more than one person in a gateway, and `rig gateway default <id>` designates who receives the last escalation step, notice digest, and workflow fallbacks — if a second person is registered, `rig gateway default <id>` must be run or these stall. `rig gateway channel add <id> --binding slack:…` registers a Slack shared channel as a gateway participant, with each message row naming the channel, sender, and policy, and `rig slack post-open-asks` posts still-open asks to Slack with reply hints that answer in the original thread. `rig gateway org set` and `rig gateway group set` attach a readable policy to organizations and groups, visible via `rig gateway participants`, and a new `whose <person|all|default>` Feed filter lets the TUI show whose ask queue is displayed.

— Names every gateway command but gives no usage counts or limits. v0.7.3

02 Per-seat permission management for Codex and Claude IMPROVED

95

`rig seat set-permissions <seat> --codex-permissions <file>` gives a Codex seat its own sandbox and approval settings passed on every managed launch, and `--codex-profile <name>` selects a Codex configuration profile for future launches. `rig seat set-permissions <seat> --rules <file.json> --reason "<why>"` applies per-seat Claude tool rules to an existing seat, and the rig spec gains `claude_permissions` with `allow` and `deny` lists so Claude members sharing a project folder can keep their own tool rules.

— Full flag set and config key named across both releases. v0.7.2 v0.7.3

03 Provider account switching and status checks NEW 90

`rig provider switch --seat <seat> --to-account <name>` (Codex, Linux) or `--token-file <path>` / `--stored-login` (Claude) moves a seat to another

subscription login mid-conversation, and `rig provider status --refresh --provider claude --claude-token-file <path> --credential-label <label>` reads a Claude token's usage on demand.

— Both subcommands and all flags named verbatim. v0.7.2

04 **Live seat runtime switching between Codex and Claude Code** NEW 88

`rig seat set-runtime <seat> --runtime claude-code|codex --reason "<why>"` moves a live seat between Codex and Claude Code runtimes, with a `--dry-run` preview mode and `--recover` options.

— Command and flags given; no detail on what state is preserved. v0.7.2

05 **Per-seat provider login via CLAUDE_CONFIG_DIR/CODEX_HOME** NEW 85

Seats can run on their own Claude or Codex login by naming an absolute `CLAUDE_CONFIG_DIR` or `CODEX_HOME` in the member's or rig's `env:`, and `rig seat status` / `rig provider status` report which home a seat uses. A one-off `rig up --env` no longer applies these two variables — they must be declared in the rig spec's `env:` instead.

— Exact env vars and commands named; migration note included. v0.7.2

06 **Environment variable scoping tightened across seats** BREAKING 78

Seats relying on the daemon's own environment variables (`ANTHROPIC_BASE_URL`, `OPENAI_BASE_URL`, `CLAUDE_CODE_USE_BEDROCK`) must now declare them in the rig's or member's `env:`, since the first seat on a new tmux server gets the same scoped environment as every other seat. On Linux, a seat's session-bus address must likewise be set in the rig's or member's `env:`, because seats now start with an address that cannot connect even when a shell startup file sets `XDG_RUNTIME_DIR`.

— Exact env vars named; no guidance on setting the bus address value. v0.7.2 v0.7.3

07 **Attention-clearing after declined permission prompts** IMPROVED 75

`rig seat clear-attention <seat> --reason "<what you saw>"` now clears the prompt named by the latest hook after a person declines a Claude permission prompt.

— Command and trigger condition named; no broader scope given. v0.7.2

08 **MCP handoff tools for held work** NEW 70

`rig_queue_handoff` and `rig_queue_unclaim` are added to OpenRig's MCP server, bringing it to 36 tools, so an assistant can pass held work on to another seat or return it in one step.

— Names the two tools and new total; no usage walkthrough. v0.7.2

09 **First-turn delivery for community harness presets** NEW

60

`first_input: true` in a community-harness preset config makes a harness receive its first startup turn at launch, instead of having it pasted into the pane.

— Config key named but no broader context given. v0.7.2

10 **Ask settlement restricted to the intended recipient** BREAKING

60

A registered person can no longer settle an ask directed to someone else — the attempt returns `answer_not_recorded`; from the TUI only the primary user can settle another person's ask.

— Error code named; no migration guidance given. v0.7.3

THINNER COVERAGE BELOW

11 **Wait-for-idle sends to mid-turn Codex seats** NEW 55

`rig send` to a mid-turn Codex seat now supports `--wait-for-idle` for separate work.

— Flag named but behavior described minimally. v0.7.2

12 **Readable-size seat capture** NEW 55

`rig capture --readable` reads a small seat at a readable size and returns its window afterward.

— Flag named; limited explanation of mechanism. v0.7.3

13 **Workspace path validation for mission creation** BREAKING

55

`rig scope --workspace <path> mission create` now refuses a path that does not exist.

— Command named; no detail on resulting error. v0.7.3

14 **Send refusal at Claude 'Resume session' picker** BREAKING

55

A send to a Claude seat sitting at its `Resume session` picker now refuses with `target_needs_input`.

— Error code named; no remedy described. v0.7.3

15 **Daemon start without kernel** NEW 50

`rig daemon start --no-kernel` skips kernel startup entirely.

— Flag named; no explanation of when to use it. v0.7.3

16 **Doctor check for tmux pane capture** NEW 45

`rig doctor` gains a check that verifies tmux can capture a pane.

— Minimal one-line description of the new check. v0.7.3

17 **Rex terminal support** NEW 35

Rex is now supported as a terminal alongside Herdr and cmux.

— Bare addition with no further detail. v0.7.2

└──▷ BREAKING ON UPGRADE

! Seats that relied on the daemon's own environment variables (`ANTHROPIC_BASE_URL` , `OPENAI_BASE_URL` , `CLAUDE_CODE_USE_BEDROCK`) must now declare them in the rig's or member's `env:` — the first seat on a new tmux server now gets the same scoped environment as every other seat.

! If a second person is registered, `rig gateway default <id>` must be run or the last escalation step and notice digest will stall.

! Downgrading to 0.7.2 loses the default person, shared channels, organizations, groups, and policy; a Codex seat's permission settings are kept in storage but ignored by 0.7.2 launches, which may cause broader permissions than intended.

! `rig scope --workspace <path> mission create` now refuses a path that does not exist.

! A registered person can no longer settle an ask that was directed to someone else (returns `answer_not_recorded`); from the TUI only the primary user can do so.

! A send to a Claude seat at its `Resume session` picker now refuses with `target_needs_input` .

! On Linux, if a seat's tools use the session bus, set the bus address in the rig's or member's `env:` — seats now start with an address that cannot connect, even when a shell startup file sets `XDG_RUNTIME_DIR` .

! Declare a seat's own `CLAUDE_CONFIG_DIR` or `CODEX_HOME` in the rig spec's `env:` — a one-off `rig up --env` no longer applies those two variables.

! Code from before #1914 opens the upgraded database but omits per-member Claude tool rules from launch, so a seat's denies stop applying — set equivalent native Claude rules before downgrading.

! Startup replaces OpenRig's own plugin and skill files (openrig-core 0.1.18 → 0.1.19; `delegating-work` and `openrig-user` skills changed) — reapply any local edits from the

copies saved under `$OPENRIG_HOME/backups/managed-refresh/`.

! Six database migrations run on the daemon's first start after upgrade — restart the daemon with `rig daemon stop && rig daemon start` after installing.

WAS THIS USEFUL? 👍 👎

OpenAI Codex CLI

Feedback ⓘ rust-v0.165.0-alpha.1.2 NOTES ↗

CODE ↗ RANK

OpenAI Codex CLI runs an agent in the terminal that reads, changes, and tests code in local repositories.

> OpenAI Codex CLI added a built-in Google Cloud Vertex AI provider and gpt-6.1-terra model support across its catalogs, alongside FedRAMP endpoint routing, expanded network proxy/MITM credential handling, and new trace instrumentation for turns, tools and skills.

└──▷ **WHAT SHIPPED · 12 FEATURES** most completely described first ? what's the number?

01 **prediction_preserve_parent_tools feature flag** NEW 80

Adds the `prediction_preserve_parent_tools` feature flag (default-enabled) and its configuration schema entries, controlling whether parent tool declarations are preserved in MCP-disabled prediction forks and `thread/sample` requests.

— Flag and default state given, no config file path shown. rust-v0.165.0-alpha.1.2

02 **gpt-6.1-terra model added across catalogs** NEW 70

Adds `gpt-6.1-terra` to the bundled model catalog, the GCP Vertex model picker as the preferred memory consolidation model, and Amazon Bedrock catalogs including Global and US cross-region runtime variants (`openai.gpt-6.1-terra`).

— Named across three catalogs but no usage command shown. rust-v0.165.0-alpha.1.2

03 **Network proxy and MITM hook enhancements** NEW 68

Adds AWS SigV4 signing to network proxy MITM hooks with support for bounded request bodies, enables masked AWS credentials for local proxy signing, adds scoped request-body secret substitution to MITM hooks, and adds constraint-aware standalone network proxy configuration.

— Several named mechanisms, no exact flags or commands given. rust-v0.165.0-alpha.1.1 rust-v0.166.0-alpha.1

04 **GCP Vertex AI provider support** NEW 65

Adds a built-in Google Cloud Vertex AI provider, enabling Vertex AI as a first-class backend without custom configuration, with a credential file override for GCP provider authentication and the ability to configure GCP project and location in provider settings.

— Names provider and settings but no exact config keys. rust-v0.165.0-alpha.1.1 rust-v0.166.0-alpha.1

05 **Message board upsert_channel API and subscriptions** NEW

65

Adds an `upsert_channel` API to create or replace a channel description on message boards, and paginated channel subscription queries to the local message board.

— Named API endpoint, no request/response schema shown. `rust-v0.166.0-alpha.1`

THINNER COVERAGE BELOW

06 **Windows sandbox helper SystemRoot fix** IMPROVED

55

Preserves `SystemRoot` through the Windows helper environment filter using case-insensitive matching, enabling the Windows filesystem sandbox helper process to initialize correctly.

— Clear bug cause and fix, no reproduction steps given. `rust-v0.165.0-alpha.1.2`

07 **Working directory reporting in tool hooks** IMPROVED

55

Reports the local command working directory in `PreToolUse` hooks and preserves command `cwd` in permission and post-tool hooks.

— Named hook, no usage example given. `rust-v0.166.0-alpha.1`

08 **Local stdio MCP protocol discovery and sharing** NEW

50

Adds opt-in protocol discovery for local stdio MCP servers and shares eligible local stdio MCP connections across chats.

— Opt-in behavior named, no flag or command given. `rust-v0.166.0-alpha.1`

09 **zstd and WebSocket compression for app-server client** NEW

45

Enables zstd compression for code-mode gRPC responses and negotiated WebSocket compression in the remote app-server client.

— Describes mechanism but no config toggle named. `rust-v0.166.0-alpha.1`

10 **FedRAMP endpoint routing** NEW

40

Routes FedRAMP API keys to government endpoints by default.

— One-line description, no config detail given. `rust-v0.166.0-alpha.1`

11 **Trace instrumentation for turns, tools and skills** NEW

40

Adds turn, tool, and skill attribution to traces, and adds trace instrumentation for turn startup and response streaming.

— Describes scope only, no trace format or access path. rust-v0.165.0-alpha.3

12 **Marketplace update failure warnings** NEW 30

Surfaces automatic marketplace update failures as startup warnings.

— Behavior named only, no trigger or format detail. rust-v0.165.0-alpha.3

> CVE-2025-59532 CVE ↗ STABLE
>
> `CVE-2025-59532` : Sandbox bypass due to bug in path configuration logic (affects `0.2.0 <= 0.38.0`)

WAS THIS USEFUL? 👍 👎

GitHub Copilot CLI

Feedback ⓘ v1.0.104 NOTES ↗ RANK

GitHub Copilot CLI is a terminal-based coding agent that can plan, write, and test code in local repositories.

> GitHub Copilot CLI v1.0.104 graduates Assisted Permissions out of experimental mode, refines how the Auto and fusion managed models select and fall back, and fixes background task handling during Autopilot sessions.

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

01 **Background task waiting in Autopilot sessions** IMPROVED

73

Interactive and SDK Autopilot sessions now keep waiting for attached background tasks instead of stopping after `COPILOT_TASK_WAIT_TIMEOUT_SECONDS`; the `-p --autopilot` flag still honors that timeout.

– Names env var and flag with clear before/after behavior v1.0.104

02 **Fusion model auto-starts on Auto** IMPROVED

60

The managed model `fusion` now starts sessions on Auto and applies `fusionTier` settings.

– Names config key `fusionTier` but no values or example shown v1.0.104

THINNER COVERAGE BELOW

03 **Auto preference fallback when unavailable** IMPROVED

40

When a saved Auto preference is no longer available, Auto now falls back to its default preference instead of failing each turn.

– Describes behavior change but names no config or flag v1.0.104

04 **Assisted Permissions no longer experimental** IMPROVED

35

Assisted Permissions are now available without enabling experimental mode, making the permission-assist workflow generally available by default.

– States the change but gives no mechanism or flag detail v1.0.104

CVE-2026-45033 CVE ↗ STABLE

`CVE-2026-45033`: Nested Bare Repository Can Execute Arbitrary Commands via core.fsmonitor (CVSS 7.8, affects `<= 1.0.42`)

CVE-2026-29783 CVE ↗ STABLE

`CVE-2026-29783` : Dangerous Shell Expansion Patterns Enable Arbitrary Code Execution (CVSS 7.8, affects `<= 0.0.422`)

WAS THIS USEFUL?  

swttch

Feedback ⓘ v0.34.9 NOTES ↗ RANK

Claude Code GUI plugin for JetBrains IDEs. (ex - Claude Code with GUI)

> swttch v0.34.9 adds browser and file preview panes, full-session chat search, AI-generated commit messages, and Mermaid diagram rendering in replies.

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

01 **Find-in-chat search** NEW 65

Adds find-in-chat search via `Cmd/Ctrl+F`, which searches the entire session rather than just the visible pane.

Search back through a long session to find a previous response without scrolling.

📍 In any open chat session, press Cmd+F (macOS) or Ctrl+F (Windows/Linux) to open the find-in-chat bar, then type your search term.

– Exact keystroke and scope, with a usage example v0.34.9

THINNER COVERAGE BELOW

02 **Browser and file preview panes** NEW 45

Adds browser and file preview panes with element selection and Desktop-style transcript files.

– Names the panes and transcript format but no exact UI path v0.34.9

03 **Mermaid diagram rendering in replies** NEW 40

Renders Mermaid fences in chat replies as diagrams.

– Names Mermaid fences but no further mechanism detail v0.34.9

04 **Commit-message generation** NEW 20

Adds commit-message generation for selected changes.

– Bare feature name, no mechanism or trigger given v0.34.9

WAS THIS USEFUL? 👍 👎

coffee-cli

Feedback ⓘ v3.7.4 NOTES ↗ RANK

A clean AI terminal workspace. No hacks, no extra hooks, no lock-in — with quick input, folder switching, copy & paste, and image pasting. 极致清爽的 AI 终端工作平台。0 魔改、无额外 Hook、无绑定；支持快捷输入、目录切换、复制粘贴与图片粘贴。

> Coffee CLI's latest releases add multi-agent orchestration (Claude delegating to Claude, Codex, and OpenCode), a skill management system, built-in computer and browser control, and an agent configuration page, alongside smaller UI improvements to session alerts, PR status, and statistics.

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

01 **Multi-agent orchestration across Claude, Codex, OpenCode** NEW

60

Claude can now delegate tasks to Claude, Codex, and OpenCode, with child tabs retaining task states, questions, and results alongside the parent session.

— Mechanism and participating tools named, no command shown v3.7.4

02 **Built-in computer and browser control** NEW

60

Adds built-in computer control on Windows and dedicated Chrome/Edge windows for browser control, enabled per tool without adding hooks.

— Names platforms and no-hook constraint, no setup steps v3.7.4

THINNER COVERAGE BELOW

03 **Agent configuration page** NEW

55

A new agent configuration page lets users manage favorite tools and global memory, with direct editing, conflict protection, and installation status indicators.

— UI surface named with its functions, no exact navigation path v3.7.4

04 **Skill management for installed tools** NEW

50

Discover skills from public upstream sources and install, update, or remove them per tool.

— Clear capability, lacks exact command or UI path v3.7.4

05 **Session list icon and status color refresh** IMPROVED

45

Adds text-sized tool icons beside session titles in the session list, with unified branch, PR, and CI status colors for light and dark themes.

— Visual change specified, no interaction or config detail v3.7.0

06 **One-click detection and update of favorite CLIs** NEW

40

Supports one-click detection and update of favorite CLIs, with an animated update indicator.

— Clear one-click action but no command or config detail v3.7.4

07 **Awaiting-action session alerts with unread markers** NEW

40

Adds blue awaiting-action labels and unread dots for replies completed in the background; viewing a conversation clears its unread marker.

— Behavior described, no config or command surface v3.7.1

08 **Per-conversation PR status display** NEW

40

Associates PRs with individual conversations, showing merged and closed states correctly, and hides plain branch icons.

— States behavior change, no actionable step given v3.7.1

09 **Current model ID shown in Codex welcome area** NEW

35

Shows the current model ID in the Codex welcome area so you always know which model is active.

— Simple UI addition, minimal mechanism described v3.7.0

10 **Full-year activity heatmap in statistics view** NEW

30

The statistics view now includes a full-year activity heatmap.

— Bare description of a visualization, no detail on data shown v3.7.4

WAS THIS USEFUL? 👍 👎

claude-mem

Feedback ⓘ v15.3.0 NOTES ↗

claude-mem gives coding agents persistent memory by capturing session activity, compressing it with a model, and injecting relevant context into later sessions.

Adds agent deliverables capture to My Work cards, automatic local DB snapshots with crash-safe restore, and a `GET /api/deliverables/:id/:filename` endpoint.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v15.3.0 https://github.com/thedotmack/claude-mem.git
# already have the repo? check out this version:
$ git checkout v15.3.0
```

Copies agent-produced deliverables into `~/.claude-mem/deliverables/<project>/` as content-addressed files via a new Stop hook, surfaced under Deliverables on the My Work card.

Serves each deliverable copy at `GET /api/deliverables/:id/:filename`; markdown and text open inline.

Adds `CLAUDE_MEM_DELIVERABLES_ENABLED=false` to disable deliverable copying, with 50-files-per-Stop and 512 MiB-per-file limits.

Adds `SubagentStop` hook to capture files a subagent produced from its own transcript.

Adds `npx claude-mem backup [run|status|list]` and `npx claude-mem restore` for automatic local database snapshots with crash-safe restore.

+1 more

NEW INTEGRATION Notion export integration

NEW INTEGRATION Runcrate preset for openai-compatible provider

NEW SUPPORTED PLATFORM Codex SubagentStart hook for memory timeline injection

WAS THIS USEFUL? 👍 👎

pstack-claude

Feedback ⓘ v0.9.103 NOTES ↗ RANK

Claude Code, Codex, Copilot, Pi, OpenCode, Muse, Gemini, and Prime Agent versions of Poteto's pstack. Rigorous agent workflows with Cursor primitives translated for other harnesses.

> pstack-claude's v0.9.103 release adds a `/statechart` command that turns TLA+ models into visual Mermaid diagrams with automatic tooling setup and CI verification, alongside a new `/ketchup-video` summary format, safer PR-branch handling in `orch frontier set`, and a rule-validation flag for the sync tool.

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

01 **`/statechart` TLA+ model visualization command** NEW

90

`/statechart` draws a Mermaid statechart from a TLA+ model, numbers each bug path that TLC finds, and can convert an agreed chart back into a TLA+ skeleton. It installs TLC and Java automatically via `mise` on first run (~305 MB), and a new CI job runs `/statechart`'s 150 `node:test` cases against the TLC jar pinned by `TLC_TOOL`, verified against its sha256.

— Full mechanism, install path, and CI verification detail given v0.9.103

02 **`--check-rules` validation flag for sync tool** NEW

80

`bun tools/sync.mjs --check-rules` derives every component at its pin and fails, naming each substitution rule that matches no upstream file.

— Exact runnable command and precise failure behaviour given v0.9.103

THINNER COVERAGE BELOW

03 **Safer PR branch selection in `orch frontier set`** IMPROVED

52

`orch frontier set` now prefers a branch's open PR over closed or merged ones, and refuses to proceed if there are two open PRs on one branch.

— Names the exact command and decision logic, no further scope v0.9.103

04 **`/ketchup-video` slide-based catch-up summaries** NEW

50

`/ketchup-video` renders the `/ketchup` catch-up as a short silent slide video, showing waiting calls as text below the slides.

— Clear command and output format, no further mechanism given v0.9.103

05 **Command discovery and skill integration for new commands**

IMPROVED 40

`/create-formal-verification-skill` now instructs generated skills to process TLA+ results with `/statechart`, and `/poteto-help` now routes to both `/statechart` and `/ketchup-video`.

— Thin routing/integration updates with no added mechanism v0.9.103

WAS THIS USEFUL? 👍 👎

Anthropic Claude Code

Feedback ⓘ v2.1.311 NOTES ↗

CODE ↗ RANK

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

> Claude Code v2.1.311 adds idle-hold controls for self-hosted runners, a signed model catalog for Bedrock/Vertex/Foundry deployments, and raises ultrathink to xhigh effort across Opus, Sonnet, Haiku 5.5 and Fable 5.1, while tightening JSON output escaping and restricting `DISABLE_COMPACT` to a one-way switch.

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

01 Signed model catalog for Bedrock, Vertex, Foundry NEW

90

Claude Code on Bedrock, Mantle, Vertex, and Foundry now loads model settings from a signed file on `downloads.claude.ai`; set `CLAUDE_CODE_MODEL_CATALOG=off` to disable this.

Opt Bedrock/Vertex/Foundry deployments out of pulling model settings from `downloads.claude.ai`.

```
$ CLAUDE_CODE_MODEL_CATALOG=off claude --model <model-id>
```

— Names platforms, env var, and gives runnable example. v2.1.311

02 Idle-hold flags for self-hosted-runner NEW

85

Adds `--hold-idle-file` and `--max-hold-min` flags to `self-hosted-runner` to keep an idle session alive while the host reports processes it left running.

Keep a self-hosted runner session alive while the host reports lingering processes, preventing premature idle teardown.

```
$ claude self-hosted-runner --hold-idle-file /tmp/runner-hold --
max-hold-min 10
```

— Named flags with a runnable example command. v2.1.311

03 Ultrathink effort raised to xhigh IMPROVED

70

`ultrathink` in the interactive terminal now runs at xhigh effort on Opus 5.5, Sonnet 5.5, Haiku 5.5, and Fable 5.1 via Anthropic's API; typing it while Claude is working raises

the effort for the rest of that turn.

— Explains mechanism and models but no runnable example. v2.1.311

04 **Escaped control characters in JSON output** IMPROVED

65

Changes `--output-format json` and `stream-json` output to write DEL, C1 control, and Unicode format characters inside strings as `\uXXXX` escapes; parsed values are unchanged.

— Named flags and escaping behavior, no example command. v2.1.311

05 **DISABLE_COMPACT now one-way off switch** BREAKING

65

`DISABLE_COMPACT` in settings files and in a mod's `$.env.set` can now only turn compaction off, not back on; any existing use of it to re-enable compaction will no longer work.

— Breaking change named precisely but no migration steps given. v2.1.311

THINNER COVERAGE BELOW

06 **Compaction warning in /teleport** IMPROVED

55

`/teleport` and `claude --teleport` now warn when a compacted cloud session loads from its last compaction point.

— Names command and flag but lacks further mechanism detail. v2.1.311

07 **VSCode extension UI polish** IMPROVED

50

Adds syntax coloring to code blocks in Claude's replies in light, dark, and high contrast dark themes, and adds a Running UserPromptSubmit hooks spinner status with a screen reader announcement for slow UserPromptSubmit hooks.

— Named UI features but no navigation path given. v2.1.311

08 **Posting-rule check in Slack connector** IMPROVED

45

The Slack connector now checks a channel's topic for posting rules before messaging a channel you did not name.

— Describes behavior but no config or command specifics. v2.1.311

09 **Tip for Explanatory output style** IMPROVED

25

Adds a tip for the Explanatory output style, shown when you ask Claude for explanations.

— Bare description with no mechanism or location. v2.1.311

└──▷ BREAKING ON UPGRADE

! `DISABLE_COMPACT` in settings files and in a mod's `$.env.set` can now only turn compaction off, not back on — any use of it to re-enable compaction will no longer work.

`CVE-2026-88968` `CVE ↗` `STABLE`

`CVE-2026-88968`: TOCTOU Race in Edit Tool Allows File Writes Outside the Workspace (affects `< 2.1.251`)

`CVE-2026-86063` `CVE ↗` `STABLE`

`CVE-2026-86063`: Plugin Commit SHA Pin Bypass Allows Installation of Attacker-Controlled Code (affects `< 2.1.179`)

`CVE-2026-103435` `CVE ↗` `STABLE`

`CVE-2026-103435`: Arbitrary File Write via Write-Time Symlink Following (TOCTOU) in Claude Code (affects `< 2.1.129`)

`CVE-2026-103012` `CVE ↗` `STABLE`

`CVE-2026-103012`: Managed Settings Not Applied When a Stored API Key Was Preferred Over Enterprise Sign-In (affects `>= 2.0.68, < 2.1.260`)

`CVE-2026-46406` `CVE ↗` `STABLE`

`CVE-2026-46406`: Insecure Temporary File in `/copy` Command Enables Response Disclosure and Symlink-Based File Write (CVSS 6.1, affects `>= 2.1.59, < 2.1.128`)

`CVE-2026-55607` `CVE ↗` `STABLE`

`CVE-2026-55607`: Sandbox Escape via Git Worktree Path Confusion Allows Unsandboxed Code Execution (CVSS 8.8, affects `>= 2.1.38, < 2.1.163`)

`CVE-2026-54316` `CVE ↗` `STABLE`

`CVE-2026-54316`: Out-of-Band Data Exfiltration via Pre-Approved HuggingFace Domain in WebFetch (CVSS 9.1, affects `>= 0.2.54, < 2.1.163`)

`CVE-2026-40068` `CVE ↗` `STABLE`

`CVE-2026-40068`: Trust Dialog Bypass via Git Worktree Spoofing Allows Arbitrary Code Execution (CVSS 8.8, affects `>= 2.1.63, < 2.1.84`)

`CVE-2026-35603` `CVE ↗` `STABLE`

`CVE-2026-35603` : Insecure System-Wide Configuration Loading Enables Local Privilege Escalation on Windows (CVSS 7.3, affects `< 2.1.75`)

`CVE-2026-33068` CVE ↗ STABLE

`CVE-2026-33068` : Workspace Trust Dialog Bypass via Repo-Controlled Settings File (CVSS 8.8, affects `< 2.1.53`)

`CVE-2026-25722` CVE ↗ STABLE

`CVE-2026-25722` : Command Injection via Directory Change Bypasses Write Protection (CVSS 9.1, affects `< v2.0.57`)

`CVE-2026-25725` CVE ↗ STABLE

`CVE-2026-25725` : Sandbox Escape via Persistent Configuration Injection in settings.json (CVSS 10.0, affects `< v2.1.2`)

`CVE-2026-25724` CVE ↗ STABLE

`CVE-2026-25724` : Permission Deny Bypass Through Symbolic Links (CVSS 7.5, affects `< v2.1.7`)

`CVE-2026-24887` CVE ↗ STABLE

`CVE-2026-24887` : Command Injection in find Command Bypasses User Approval Prompt (CVSS 8.8, affects `< v2.0.72`)

`CVE-2026-24053` CVE ↗ STABLE

`CVE-2026-24053` : Path Restriction Bypass via ZSH Clobber Allows Arbitrary File Writes (CVSS 6.5, affects `< v2.0.74`)

`CVE-2026-24052` CVE ↗ STABLE

`CVE-2026-24052` : Domain Validation Bypass Allows Automatic Requests to Attacker-Controlled Domains (CVSS 7.4, affects `< v1.0.111`)

`CVE-2026-21852` CVE ↗ STABLE

`CVE-2026-21852` : Malicious repo configuration can trigger data leakage via environment configuration used before trust confirmation (CVSS 7.5, affects `< v2.0.65`)

`CVE-2025-65099` CVE ↗ STABLE

> `CVE-2025-65099`: Command execution prior to Claude Code startup trust dialog (CVSS 9.8, affects `<v1.0.39`)

> `CVE-2025-59828` `CVE ↗` `STABLE`
>
> `CVE-2025-59828`: Claude Code Vulnerable to Arbitrary Code Execution via Plugin Autoloading with Specific Yarn Versions (CVSS 9.8, affects `< v1.0.39`)

WAS THIS USEFUL? 👍 👎

Diagram Design

Feedback ⓘ changes since 2026-10-12 CODE ↗

Diagram Design is a Claude Code skill that generates editorial HTML and SVG diagrams and imports draw.io or Mermaid sources.

Diagram Design: Adds service blueprint as visual type #45, with equal-column stages and required customer/frontstage/backstage rows.

Adds architecture containment scopes as a new diagram capability.

WAS THIS USEFUL? 👍 👎

zot

Feedback ⓘ v0.4.51 NOTES ↗ RANK

Yet another coding agent harness, lightweight and written in go.

> zot added a new built-in AI provider and keyless template support for its mcp-bridge integration.

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

01 **Keyless deed.page template in mcp-bridge** NEW 45

Adds keyless `deed.page` template support to the `mcp-bridge` integration.

— Names exact template and integration but no usage detail v0.4.50

02 **Cheaper Inference built-in AI provider** NEW 40

Adds Cheaper Inference as a new built-in provider, including model discovery support.

— Names provider and model discovery, no further mechanism given v0.4.51

WAS THIS USEFUL? 👍 👎

claude-council

Feedback ⓘ v2026.10.14 NOTES ↗ RANK

Claude Code plugin that asks several AI coding agents the same question and shows their answers side by side. Gemini, OpenAI, Grok, Perplexity, Kimi and any model OpenRouter routes to over API; codex, antigravity, grok, kimi and cursor CLIs on subscription auth; ollama locally.

> claude-council's single release this window focused on cutting token overhead for CLI council seats and making the `/heads-up` side-check smarter about reporting usage and pausing itself when codex can't run.

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

01 **Automatic pause/resume for heads-up checks** NEW 78

The heads-up check automatically pauses for the session when codex is out of credits, at its usage limit, or when the codex setup can't run in its own home (keyring login, no login, a `profile`, or a custom `model_provider`); `/heads-up on` resumes it.

Resume a heads-up side-check that was automatically paused because codex hit its usage limit.

```
$ /heads-up on
```

— Names exact trigger conditions and the resume command. v2026.10.14

02 **Token usage row in heads-up status** NEW 73

`/heads-up status` gains a `tokens` row reporting per-seat token counts for codex, all API seats, claude-cli, cursor-cli, copilot-cli, and opencode; grok-cli, kimi-cli, and antigravity show as `unreported`.

Check how many tokens each seat consumed in the last heads-up side-check, so you can spot unexpectedly large seat contexts.

```
$ /heads-up status
```

— Names every seat covered and gives the exact command to run. v2026.10.14

03 **Isolated temp directories for CLI seats** IMPROVED 68

Each CLI seat now starts from an empty, read-only directory at `$TMPDIR/claude-council-<uid>/seat` instead of the user's working directory, cutting input tokens substantially — e.g. grok-cli drops from 46.3k to 12.0k tokens and the codex council seat from 30.6k to 13.2k on a one-line question.

— Clear mechanism and before/after numbers, but nothing for the user to invoke.

v2026.10.14

04 **Trimmed codex home for heads-up checks** IMPROVED 62

The heads-up codex check now runs under a trimmed codex home containing only model, effort, and login, cutting the same 15.6 KB prompt from 35.3k to 17.5k input tokens.

— Precise before/after numbers but no user-facing action. v2026.10.14

WAS THIS USEFUL? 👍 👎

Superset

Feedback ⓘ desktop-v2.0.0 NOTES ↗ RANK

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

> Superset's desktop v2.0.0 adds cloud workspace launches via pdev-agent, new keyboard shortcuts for pane management, a toggle for ACP chat, and chat navigation and sign-in screen refinements.

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

01 **Keyboard shortcuts for pane management** IMPROVED 70

Adds `⌘⇧L` to toggle the right pane full screen, and reassigns the Open Changes shortcut to `⌘⇧G`.

Toggle the right pane to full screen while reviewing changes, freeing maximum horizontal space for the diff view.

📍 Press ⌘⇧L to expand the right pane to full screen; press ⌘⇧G to open the Changes pane.

— Exact key combos given, directly runnable desktop-v2.0.0

THINNER COVERAGE BELOW

02 **ACP chat toggle in Experimental settings** NEW 55

Flagged users can now turn ACP chat off via Settings > Experimental, letting them disable the feature if not needed.

Disable ACP chat for flagged accounts that do not need it, reducing attack surface on shared machines.

📍 In the desktop app, go to Settings › Experimental and toggle ACP chat off.

— Exact UI path given but limited to flagged users with no further mechanism

desktop-v2.0.0

03 **Cloud workspace launch via pdev-agent** NEW 35

Desktop v2.0.0 launches `pdev` in cloud workspaces behind `pdev-agent`.

— Names the mechanism but no setup steps or scope given desktop-v2.0.0

04 **Chat navigation improvements** IMPROVED 35

Links `/resume` to open the chat history dropdown, and shows Open Chat in the empty workspace placeholder.

— Two thin UI additions with no deeper behaviour described desktop-v2.0.0

05 **Sign-in screen refresh** IMPROVED 25

Adds a Last used badge on the sign-in provider button for faster re-authentication, and delivers a calmer sign-in screen with clearer browser hand-off.

— Visual/UX change described only in general terms desktop-v2.0.0

WAS THIS USEFUL? 👍 👎

Graphify Labs Graphify

Feedback ⓘ v0.9.96 NOTES ↗

RANK

Graphify converts code, docs, SQL schemas, configs, and PDFs into a queryable knowledge graph with explainable relationships.

> Graphify's latest release adds a configurable timeout for watch rebuilds, persists token-cost tracking for standalone extract runs, and introduces smart skip logic so `graphify update` avoids re-running unchanged pipelines.

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

01 Watch rebuild timeout via GRAPHIFY_REBUILD_TIMEOUT NEW

88

Adds the `GRAPHIFY_REBUILD_TIMEOUT` environment variable (default `600` s, `0` disables) to bound each watch rebuild, with a thread-Timer fallback used where `SIGALRM` is unavailable.

Cap a watch rebuild to 2 minutes in CI so a hung LLM pass never stalls the watcher process.

📍 `GRAPHIFY_REBUILD_TIMEOUT=120` graphify watch .

— Names exact env var, default, disable value, and fallback mechanism. v0.9.96

02 Skip unchanged pipeline in graphify update IMPROVED

72

`graphify update` now skips the detect/extract/build pipeline when no relevant input has changed, keyed on the scan file-set, config, `.graphifyrc`, extractor epoch, and an output-completeness stamp.

— Explains the caching mechanism and its keys but gives no command to trigger it. v0.9.96

03 Token cost ledger for standalone extract runs IMPROVED

70

`graphify extract` now records each run's token usage to `graphify-out/cost.json`, keeping the cost ledger current after standalone extract refreshes.

After a standalone extract refresh, inspect the updated token spend written to the cost ledger.

```
$ graphify extract . && cat graphify-out/cost.json
```

— Names exact command and output file path. v0.9.96

WAS THIS USEFUL? 👍 👎

The Open Engine Zeroshot Feedback ⓘ v10.19.0 NOTES ↗

RANK

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

> Zeroshot added a guided interview skill that helps design run workflows directly inside Codex, GitHub Copilot, Pi, and Claude Code.

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

01 Guided workflow design skill for AI coding agents NEW

70

Adds the `zeroshot-graph-design` guided interview skill to the canonical skills installed for Codex, GitHub Copilot, Pi, and Claude Code, intended to help design run workflows. Run `zeroshot update` to install it if it's missing after upgrading.

— Names exact skill and update command, but lacks mechanism detail. v10.19.0

WAS THIS USEFUL? 👍 👎

OpenQodex

Feedback ⓘ openqodex@0.15.0 NOTES ↗

Open source code review that runs in your coding agent before you push. Scanners plus a separate AI reviewer, on your own login, no key.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch openqodex@0.15.0
https://github.com/openqodex/openqodex.git
# already have the repo? check out this version:
$ git checkout openqodex@0.15.0
```

NEW SUPPORTED PLATFORM Java language support in the code graph

WAS THIS USEFUL? 👍 👎

Cline

Feedback ⓘ v4.1.26 NOTES ↗ CODE ↗ LATEST STABLE

Cline is an open-source AI coding agent for VS Code that can read and edit files, run terminal commands, and use browsers with user approval.

Adds a Local | Cloud model toggle in API Configuration for cloud sessions, plus a sandbox progress card showing repo, branch and startup steps.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v4.1.26 https://github.com/cline/cline.git
# already have the repo? check out this version:
$ git checkout v4.1.26
```

Adds a `Local | Cloud` toggle in `Settings > API Configuration` when cloud sessions are on, letting you pick the Cline model for new cloud tasks separately from your local models.

Adds a progress card when starting or resuming a cloud sandbox, showing the repo, branch and steps (start sandbox, clone repository, start agent); collapses once the agent starts.

Cloud sessions now hide the Plan/Act toggle and auto-approve bar entirely rather than showing them disabled.

Kimi K3 now receives `max` reasoning level for Extra High and Max settings.

NEW INTEGRATION CommandCode, Reflection, Liquid Inference and Munito AI providers

RETIRED INTEGRATION CrofAI, iFlow and Llama providers removed

NEW SUPPORTED PLATFORM MiniMax M3.1 Flash and DeepSeek V4.1 Flash 1017 added to ClinePass

NEW SUPPORTED PLATFORM Reflection Beam, Nemotron 3.5 Lightning and Laguna S 2.1 added to Cline free list

RETIRED SUPPORTED PLATFORM Mimo V2.6 Flash and Muse Spark 1.3 Contributor dropped from Cline free list

└──▷ BREAKING ON UPGRADE

! If you use a provider without pinning a model, expect a different default model after upgrade — most providers now default to Kimi K3.1 Pro, MiniMax M3.1 Flash or DeepSeek V4.1 Flash 1017.

▾ 1 more release in this issue · 2026-06-23 → 2026-10-26

CVE-2026-59723 CVE ↗ STABLE

`CVE-2026-59723` : Cross-Origin WebSocket Hijacking in Cline Hub Dashboard (`/browser` endpoint) (CVSS 8.8, affects `<= 3.0.24`)

WAS THIS USEFUL? 👍 👎

ai-rulez Feedback ⓘ v5.7.0 NOTES ↗ RANK

Author, generate, validate and govern AI agent rules, skills, plugins and MCP configuration across 52 coding assistants. Review with verified model evidence, evaluate skills, lock and sign dependencies, and publish bundles.

> ai-rulez v5.7.0 adds a skill-scoped fallback option for rule configuration and tightens MCP skill lookups to keep user-only skills hidden from models.

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

01 **Skill-scoped fallback for rules** NEW 70

Adds `scoped_fallback = "skill"` to rule configuration, letting a rule's fallback resolution stay restricted to skill scope instead of pulling in a broader default context.

Restrict a rule's fallback resolution to skill scope so it never pulls in broader context unintentionally.

toml

```toml
scoped_fallback = "skill"
```

— Names exact config key with runnable TOML example v5.7.0

THINNER COVERAGE BELOW

02 **MCP find_skill path scoping and privacy** IMPROVED 45

Scopes MCP `find_skill` lookups by skill paths and keeps user-only skills hidden from the model, tightening which skills an agent can discover over MCP.

— Describes behavior change but no concrete command or config shown v5.7.0

WAS THIS USEFUL? 👍 👎

ai-git-bot

Feedback ⓘ 1.29.0 NOTES ↗ RANK

The self-hosted AI workflow automation platform for Git repositories.

> ai-git-bot 1.29.0 adds Azure DevOps as a supported repository provider, introduces a backmerge-to-develop workflow and agentic review progress comments, and forwards provider-specific chat template arguments to OpenAI-compatible chat requests.

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

01 **chat_template_kwargs forwarding to chat requests** NEW

60

Forwards `chat_template_kwargs` from provider settings to OpenAI-compatible chat requests.

– Names exact config key forwarded, but not its values. 1.29.0

THINNER COVERAGE BELOW

02 **Timestamp-based Flyway migration versions** IMPROVED

40

Switches Flyway database migrations to timestamp-based versions instead of the prior scheme.

– Names the mechanism change but no migration guidance given. 1.29.0

03 **Azure DevOps repository provider support** NEW

35

ai-git-bot now supports Azure DevOps as a repository provider, alongside its existing provider integrations.

– Names the integration but no setup or scope detail. 1.29.0

04 **Agentic review progress comments** NEW

35

Adds optional agentic review progress comments posted during code review.

– Describes behaviour but not how to enable it. 1.29.0

05 **Backmerge-to-develop workflow** NEW

15

Adds a backmerge-to-develop capability, mentioned in the release summary alongside the other 1.29.0 features.

– Only named in the summary, no mechanism or config given. 1.29.0

WAS THIS USEFUL? 👍 👎

prflow

Feedback ⓘ v2.63.0 NOTES ↗ RANK

Turn feature requests into review-ready pull requests: issue, implementation, tests, a review-and-fix loop, and docs. Runs locally or on GitHub Actions, and is built for complex brownfield repos.

> prflow's v2.63.0 release refines the `/prflow:spec` and `/prflow:implement` commands: spec sessions can now exit cleanly without filing anything and ask fewer redundant questions, while implement now blocks on unresolved spec issues until they're explicitly unblocked.

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

01 **"File nothing" exit path in /prflow:spec** NEW 80

Adds a No change — file nothing option to `/prflow:spec` implementation-approach questions, ending the run without any GitHub writes and surfacing the open issue for manual triage.

Bail out of a spec session cleanly when none of the implementation approaches are right — no issue filed, no GitHub writes.

```
$ /prflow:spec Add CSV export to the reports page
# When the approach question appears, choose: No change — file
nothing
```

— Names exact option and shows runnable command with no GitHub writes. v2.63.0

02 **Blocked-issue guard in /prflow:implement** NEW 80

Adds a pre-branch guard to `/prflow:implement` that stops execution when the target issue still contains a `## 🚫 Blocked` section, naming the exact remedy required before re-running.

Unblock a stalled implementation by resolving open questions — remove `## 🚫 Blocked` from the issue before re-running.

```
$ /prflow:implement 42
# If blocked: edit issue #42, resolve each open question, remove
the '## 🚫 Blocked' section, then re-run:
/prflow:implement 42
```

— Exact section name, command, and remedy steps provided. v2.63.0

THINNER COVERAGE BELOW

03 **Fewer redundant questions in /prflow:spec** IMPROVED

50

Reduces the `/prflow:spec` question set by skipping questions made wrong or moot by a prior answer, and covers each approach option's recurring cost at selection time.

– Mechanism described but no example or exact trigger given. v2.63.0

WAS THIS USEFUL? 👍 👎

Cotool

Feedback ⓘ v0.77.0 NOTES ↗ RANK

Cotool is an AI security operations platform for investigating alerts, enriching threat intelligence, and automating response workflows.

> Cotool v0.77.0 adds an MCP server for using Cotool from popular AI clients and introduces a Cisco Umbrella integration for DNS activity and domain reputation, alongside tightened Tailscale authentication.

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

01 **MCP server for AI client access** NEW

57

Cotool MCP lets Cotool be used from ChatGPT, Claude, Cursor, Codex, VS Code, and custom AI clients, operating under the user's existing login and permissions.

– Names supported clients but no setup steps or config shown. v0.77.0

02 **Cisco Umbrella integration for DNS and domain reputation** NEW

52

Adds a Cisco Umbrella integration that surfaces DNS activity and domain reputation data, with destination list remediation gated behind an approval step.

– Describes capability and approval gate but no endpoint or config detail. v0.77.0

03 **Scoped OAuth credentials for Tailscale** IMPROVED

27

Enhances Tailscale connections to use scoped OAuth client credentials rather than broader access.

– Bare mention of scoped OAuth with no mechanism or config shown. v0.77.0

WAS THIS USEFUL? 👍 👎

rtk

Feedback ⓘ v1.1.0 NOTES ↗ RANK

CLI proxy that reduces LLM token consumption by 60-90% on common dev commands. Single Rust binary, zero dependencies

> rtk v1.1.0 adds fish shell support for wrapped commands and compacts Angular build output tables.

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

01 **Fish shell support via run command** NEW 65

rtk now supports wrapping commands in the fish shell using `rtk run --shell fish`.

— Names exact flag but no further mechanism detail v1.1.0

THINNER COVERAGE BELOW

02 **Compact Angular build tables** IMPROVED 30

The `ng` subsystem now compacts Angular build tables safely, reducing token output for `ng` build commands.

— Bare description, no mechanism or metrics given v1.1.0

> CVE-2026-55249 CVE ↗ STABLE
>
> `CVE-2026-55249`: OpenClaw Rewrite Plugin Command Injection via execSync Template String (CVSS 6.3, affects `1.0.0`)

> CVE-2026-54555 CVE ↗ STABLE
>
> `CVE-2026-54555`: Permission-gate bypass in rtk rewrite auto-allow via unsplit shell separators (CVSS 7.8, affects `<= 0.40.0`)

> CVE-2026-45792 CVE ↗ STABLE
>
> `CVE-2026-45792`: RTK improperly trusts project-local filter configuration, allowing silent tampering of command output shown to LLM (CVSS 5.5, affects `< 0.32.0`)

WAS THIS USEFUL? 👍 👎

keploy

Feedback ⓘ v3.6.163 NOTES ↗ LATEST STABLE

Open-source platform for creating safe, isolated production sandboxes for API, integration, and E2E testing.

Keploy v3.6.163 improves proven-mode outgoing call attribution by deciding and learning each call within its owning test.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v3.6.163 https://github.com/keploy/keploy.git
# already have the repo? check out this version:
$ git checkout v3.6.163
```

Improves `proven` mode to decide and learn each outgoing call within the test it belongs to, enabling more accurate per-test mock attribution.

▾ 1 more release in this issue · 2026-08-31 → 2026-10-25

> CVE-2026-82641 CVE ↗ STABLE
>
> `CVE-2026-82641` : Keploy agent binds unauthenticated control-plane HTTP server to all interfaces, exposing TLS session keys (CVSS 8.6, affects `>= 3.1.0, < 3.6.26`)

WAS THIS USEFUL? 👍 👎

flow-next

Feedback ⓘ flow-next-v9.3.0 NOTES ↗ RANK

Faster than your agent alone. And better.

> flow-next v9.3.0 overhauls `make-pr` PR bodies into a leaner, reorganized set of sections and adds a new `changeWalkthrough.shape` field to the aid artifact for surfacing public API changes — both are breaking changes for anything consuming the old formats.

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

01 make-pr PR bodies restructured into new sections

BREAKING

95

`make-pr` PR bodies are restructured with new headings — Why, Review focus, Needs you, Shape, What changes for a user or operator, Reading order, Tradeoffs, and Verified — cutting a reviewer's required reading from ~720 words to ~150. The Shape section renders a fenced sketch of any new or changed public surface (types, signatures, CLI flags, schema, config keys, call paths) drawn from the diff or spec. Review focus names the file to start with, up to three areas to scrutinise and why, and the feedback the author wants. Needs you surfaces failed and unverified checks and open items near the top, moving passed checks to the Verified section at the end. Reading order replaces the old file tree, naming one start file per group, listing requirement IDs and what to check, and flagging mechanical or generated files as safe to skim. `make-pr` no longer re-runs tests to populate the PR body; proof is drawn from receipts, task evidence, and CI, with gates lacking a recorded result shown as unverified. This is a breaking change: tools parsing the old headings (Scope, Blast radius, Verification, Open items) must be updated to the new ones.

Refresh an existing open PR to pick up the new Shape section and restructured headings without opening a new pull request.

```
$ make-pr --update
```

— Names every new heading, its mechanism, and a command to refresh a PR. flow-next-v9.3.0

02 `changeWalkthrough.shape` field in aid artifact

BREAKING

78

Adds `changeWalkthrough.shape` to the aid artifact: an optional fenced code block of up to 4,000 characters capturing types, signatures, CLI flags, schema, or config keys added or changed by a PR; the validator rejects any form other than a single fenced

block. The field is introduced at schema version 1, so consumers that reject unknown fields must be updated before they can read new artifacts.

— Names field, char limit and schema version but no usage walkthrough. `flow-next-v9.3.0`

└──▷ BREAKING ON UPGRADE

! PR body headings are renamed: tools that parse the old headings Scope, Blast radius, Verification, or Open items must be updated to follow the new headings Why, Review focus, Needs you, Shape, What changes for a user or operator, Reading order, Tradeoffs, and Verified.

! The aid artifact gains an optional `shape` field at schema version 1; consumers that reject unknown fields must be updated to accept it before reading new artifacts.

WAS THIS USEFUL? 👍 👎

harness-remote

Feedback ⓘ v3.6.0 NOTES ↗ RANK

Native-session control plane for Codex, Claude Code, OpenCode, OMP and PI. Run, resume and hand off coding sessions across your machines.

> Harness Remote 3.6.0 closes a root-boundary escape across Windows drives, UNC shares and native Session handoff, ships a new Mahurley mobile skin as default, and adds desktop update notifications alongside cross-agent sync improvements for Codex, Claude Code and OpenCode 2.

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

01 Mahurley mobile skin set as default NEW

83

Adds Mahurley as the new default mobile skin on screens up to 780 px wide, with Classic still available via Settings → Visual. Mahurley Home has a one-row top bar with search, New Session, status tabs (All / Live / Attention), agent-logo filters, a per-machine Recent group, and collapsible Projects; a New Session modal lets users pick the coding agent via agent logos with Projects ordered by recent use; and the Mahurley Conversation layout has a one-row header, a model/effort/harness card next to the attachment button, an effort slider, touch-friendly model search, and persistent error display until dismissed.

Switch back to the Classic layout on mobile if Mahurley's new default doesn't suit your workflow.

📍 In the desktop or mobile app, go to Settings → Visual, open the skin gallery, and select Classic.

— Enumerates every UI element and gives the switch-back path v3.6.0

02 Root boundary enforced on every session path BREAKING

72

Enforces `--root` on every path that starts work, including other Windows drives, UNC shares, and native Session handoff, which previously could escape the configured root boundary.

— Names the flag and exact escape paths closed, but no migration note v3.6.0

THINNER COVERAGE BELOW

03 Transparent reload for idle Claude Code sessions IMPROVED

56

Sessions idle for 30 minutes are unloaded only when Claude Code itself reports them idle, and reload transparently on the next prompt or read.

— Names exact idle threshold and reload trigger, no user action needed v3.6.0

04 **Desktop update notifications** NEW 54

The packaged desktop app checks for a newer release at startup, shows a notification when one exists, and offers a manual recheck in Settings.

Manually trigger an update check on the desktop app without waiting for the next startup.

> 📍 In the desktop app, go to Settings and select the manual recheck option in the update notifications section.

— Describes mechanism and gives manual-check location v3.6.0

05 **OpenCode 2 history paging improvements** IMPROVED 48

OpenCode 2 history paging now follows user turns and shows at least 20 complete exchanges, with reliable large history export loads.

— Gives concrete exchange count but no reproduction steps v3.6.0

06 **Codex cross-device conversation sync** IMPROVED 42

Turns continued in the Codex desktop app now appear in a conversation already open on mobile, and live refresh follows long turns spanning several pages.

— Explains behaviour but no config or trigger named v3.6.0

07 **Markdown links open in system browser** NEW 38

Markdown links in the desktop app now open in the system browser, restricted to HTTP(S) and mailto schemes.

— Names the restricted schemes but behaviour is simple and self-explanatory v3.6.0

WAS THIS USEFUL?  

Amazon Kiro Crew

Feedback ⓘ CVE-2026-112380 CVE ↗

Kiro Crew is an open-source workspace for coordinating AI coding agents.

`CVE-2026-112380`: Missing authorization in the WhatsApp channel in Kiro Crew (CVSS 9.9, affects `>=0.4.0, <0.7.2`)

WAS THIS USEFUL? 👍 👎

Block Goose

Feedback ⓘ CVE-2026-72718 CVE ↗

Goose is an open-source AI agent with a desktop app and CLI for coding and other automated tasks.

`CVE-2026-72718`: Arbitrary command execution in goose CLI via `goose review` via git core.fsmonitor (affects `< 1.44.0`)

WAS THIS USEFUL? 👍 👎

Earendil Works Pi

Feedback ⓘ CVE-2026-54325 CVE ↗

STABLE

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

`CVE-2026-54325`: Pi loads project-local extensions without approval (CVSS 4.4, affects `< 0.79.0`)

▾ 1 more release in this issue · 2026-06-08

> CVE-2026-54328 CVE ↗ STABLE
>
> `CVE-2026-54328`: Predictable temporary extension install paths allow local privilege escalation on shared Linux hosts (CVSS 7.3, affects `>= 0.74.0, < 0.78.1`)

WAS THIS USEFUL? 👍 👎

Zed

Feedback ⓘ CVE-2026-44461 CVE ↗ STABLE

Zed is an open-source code editor that runs collaborative development workflows.

CVE-2026-44461 : Remote Command Injection via Unquoted Environment Variable Keys (SSH / WSL Remote) in Zed v0.225.9 (CVSS 8.6, affects <=0.225.9)

▾ 7 more releases in this issue · 2025-08-11 → 2026-05-08

> CVE-2026-44463 CVE ↗ STABLE
>
> CVE-2026-44463 : Allowlist Bypass via Environment Variable Injection in Terminal Tool Permissions (CVSS 8.6, affects <v0.229.0)

> CVE-2026-44465 CVE ↗ STABLE
>
> CVE-2026-44465 : Zed IDE Arbitrary Code Execution via untrusted repository with poisoned .git/config (CVSS 8.6, affects <v0.227.1)

> CVE-2026-27976 CVE ↗ STABLE
>
> CVE-2026-27976 : Zed Extension Sandbox Escape via Tar Symlink Following (CVSS 8.8, affects <v0.224.4)

> CVE-2026-27800 CVE ↗ STABLE
>
> CVE-2026-27800 : Zip Slip Path Traversal in Extension Archive Extraction (CVSS 7.4, affects <v0.224.4)

> CVE-2026-25805 CVE ↗ STABLE
>
> CVE-2026-25805 : Parameter Values are not shown for MCP Tool Calls. Users cannot detect tool poisoning. (CVSS 6.4, affects 0.217.3)

> CVE-2025-68433 CVE ↗ STABLE
>
> CVE-2025-68433 : Zed IDE MCP Context Server Configuration Arbitrary Code Execution (CVSS 7.7, affects < v0.218.2-pre (including current stable release v0.217.2))

> CVE-2025-55012 CVE ↗ STABLE

> `CVE-2025-55012`: AI Agent Remote Code Execution (affects `< 0.197.3`)

WAS THIS USEFUL? 👍 👎

All Hands AI OpenHands

Feedback ⓘ CVE-2026-33718 CVE ↗

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

`CVE-2026-33718`: Command Injection in Git Diff Handler (CVSS 7.6, affects `<=1.4.0`)

WAS THIS USEFUL? 👍 👎

SST OpenCode

Feedback ⓘ CVE-2026-22813 CVE ↗ STABLE

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

`CVE-2026-22813`: Malicious website can execute commands on the local system through XSS in the OpenCode web UI (CVSS 6.1, affects `<1.1.10`)

▾ **1 more release in this issue** · 2026-01-12

> CVE-2026-22812 CVE ↗ STABLE
>
> `CVE-2026-22812`: Unauthenticated HTTP Server Allows Arbitrary Command Execution (CVSS 8.8, affects `< 1.0.216`)

WAS THIS USEFUL? 👍 👎

◆ Model Training & Fine-Tuning

Google Flax

Feedback ⓘ v0.12.11 NOTES ↗ RANK

Flax provides neural-network APIs for building and training models with JAX.

> Flax v0.12.11 adds fine-grained control over no-op capture in sow/perturb, makes nnx.Sequential iterable with per-layer kwargs, includes WeightNorm scales as trainable parameters, and adds two new example models.

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

01 No-op capture control for sow and perturb NEW

70

Adds a `capture_noop` option to `Module.sow` and `Module.perturb`, plus a global `nnx_sow_capture_noop` config flag, to control whether intermediate values are captured during no-op passes (e.g. inference) consistently across both call sites.

Suppress intermediate sow captures in a no-op pass (e.g. during inference) without changing model code, using the new per-call flag.

python

```
# During inference, skip sow side-effects cleanly
intermediate = model.sow('intermediates', 'hidden', x,
capture_noop=True)
```

— Names exact flag and option but limited mechanism detail beyond control. v0.12.11

02 Iterable nnx.Sequential with per-layer kwargs IMPROVED

60

`nnx.Sequential` is now iterable and filters `**kwargs` per layer, enabling layer-by-layer iteration and finer-grained argument passing to individual sub-modules.

Iterate over layers in an `nnx.Sequential` model to inspect or modify individual sub-modules after the container is built.

python

```
model = nnx.Sequential([nnx.Linear(4, 8, rngs=rngs),
nnx.Linear(8, 2, rngs=rngs)])
for layer in model:
    print(layer)
```

— Concrete API named with runnable example but limited mechanism detail. v0.12.11

THINNER COVERAGE BELOW

03 **WeightNorm scales included as trainable parameters** IMPROVED

50

`nnx.WeightNorm` scales are now stored as `Param` so they are included in the model's trainable parameters.

— Names the module and type change but no usage example given. v0.12.11

---

04 **New sepreformer and multimodal LLM examples** NEW

25

Adds a sepreformer example and a multimodal LLM example to the Flax examples collection.

— Bare mention of two examples with no further detail or location. v0.12.11

WAS THIS USEFUL?  

fizgig

Feedback ⓘ v8.9.92 NOTES ↗ RANK

Fizgig — LoRA & Fine-tune Studio for Klein 9B, Krea 2, MiniMax H3, Qwen Image 2.1, Anima & Z-Image Turbo and Ace-Step XL: train, fine-tune, profile, repair and extract LoRAs & LoKR for Image, Video & Music Models

> Fizgig's biggest window in a while: a full in-app mask editor with brush/eraser/SAM-click plus word-driven SAM 3.1 mask generation, sub-folder training with per-folder captions, a Continue button and persistent queue sidebar, a prior-preservation control for masked training, per-run log files, and a `FIZGIG_REF` env var to pin pods to a specific release.

└──▷ **WHAT SHIPPED · 15 FEATURES** most completely described first ? what's the number?

01 **In-app mask editor on Image Prep tab** NEW 80

Adds an Edit masks step to the Image Prep tab's masks card with Brush and Eraser tools (Size, Softness, Strength controls), SAM click add/remove interaction powered by SAM 3.1 (left-click to add, right-click/Alt+click to remove), Fill, Clear and Invert operations on the full mask, scroll-to-zoom and Space+drag pan, and Undo/Redo. Provides three view modes (picture-and-mask, mask-only, picture-only), auto-saves masks on image change/filter change/tab exit (deleting all-black masks so that image trains unmasked), and makes the Only my subject checkbox available immediately after a Look Filter scan without a tab switch.

— Many named controls but only UI path, no command v8.9.91

02 **Sub-folder training** NEW 75

Adds Also train on its sub-folders tick on the Training tab to recursively train pictures in sub-folders at any depth, with each sub-folder treated as its own set with its own captions and cache. Automatically excludes `excluded_by_look`, `originals`, hidden folders, and output/cache folders nested inside the training folder, and carries sub-folders automatically when using Upload a folder, Send to this pod, or Move a run on a pod.

— Mechanism and exclusion rules named, but only a UI checkbox v8.9.92

03 **SAM 3.1 word-driven mask generation** NEW 70

Adds `SAM 3.1: from words` mask mode on the Image Prep tab's masks card, letting you name what to include (e.g. `person`, `hair`, `red dress`) and what to exclude (e.g. `glasses`, `phone`) to generate GPU-accelerated masks with sharper edges than the person finder. Adds `Preview on 2 random pictures` sampling that cycles without repeats until every picture has been shown, and makes `Only my subject` compatible with SAM 3.1 by filtering masks to the person whose face matches the Look Filter's baselines.

— Named mode and examples but no runnable command v8.9.9

04 **`FIZGIG_REF` environment variable for pod version pinning** NEW 70

Adds `FIZGIG_REF` environment variable to pin a pod to a specific Fizgig release tag (e.g. `v8.9.92`) instead of always pulling the newest version on start.

Pin a RunPod pod to this exact release mid-job so an automatic update does not interrupt a long fine-tune.

```env
FIZGIG_REF=v8.9.92
```

— Named env var with runnable example value v8.9.92

05 **Per-run log files in `user_data/logs`** NEW 70

Saves each training run's log as a file in `user_data/logs`, named by date, time, and LoRA name; retains the newest 20 logs so failed runs can be reviewed after a restart.

— Exact path, naming scheme and retention count given v8.9.7

06 **Prior-preservation control for masked training** NEW 70

Adds 'Outside the mask, keep the model's own picture' control (Off, 0.5, 1, or 2) to the Masked training card on the Training tab, letting LoRAs learn prior preservation outside the mask so backgrounds and base-model appearance do not drift toward training photos.

— Named control with discrete values and effect explained v8.9.7

07 **Uncapped Extract target rank** IMPROVED 65

Extract target rank now accepts any integer via an `Other...` option, removing the previous 1–128 ceiling so the Profiler's `Rank holding 95%` value can be used directly.

— Names before/after limit and linked Profiler value v8.9.8

THINNER COVERAGE BELOW

08 **Continue button for queued and paused runs** NEW 55

Adds a Continue button to each run card in Runs & outputs, resuming training from the newest save with the run's original settings, and queuing automatically if another run is already active.

Resume a paused project from its last checkpoint while another run trains — it will be queued automatically.

> 📍 In Runs & outputs, locate the run whose newest save is still on disk and click Continue. If another run is currently training, the continued run is added to the queue. To adjust epoch count before it starts, click Edit in the queue.

— Step-by-step example given, clear starting point v8.9.92

---

09 **Persistent queue sidebar and run editing bar** IMPROVED 50

Moves the Queue box into the persistent sidebar so it is visible on every tab, and relocates the queue itself to the top of Runs & outputs. Adds a bottom-of-window editing bar (available on every tab) while editing a queued run, with Save to queue and Cancel edit controls.

— UI relocation and controls named, no deeper mechanism v8.9.92

---

10 **Numbered-list sample prompts with inline editing** IMPROVED 50

Samples tab now renders each prompt in its own numbered box with full-width layout, Enter/Backspace/paste editing, and hover-to-expand in the gallery and Prompt × epoch views.

— Named interactions but no config surface v8.9.8

---

11 **Independent recaption task assignment per training attempt** NEW 45

Adds First try uses and Second try uses selectors on the Training tab so auto-recaption can assign any Captions-tab task to each attempt independently.

— Named selectors but thin on mechanism v8.9.92

---

12 **Training tab reorganized with Extras divider** IMPROVED 45

Training tab reorganized so only Output and Training Parameters cards open by default; Masked training and Subjects move under a new Extras divider, with cards ordered by run frequency.

— UI layout change, no mechanism beyond reordering v8.9.8

13 **Expanded auto-download for mask helper models** IMPROVED

40

Expands Download models for me to also fetch the Make masks person-finder model alongside the face model and other helpers.

— Named feature but little mechanism detail v8.9.7

14 **ArcFace likeness score ceiling explanation** IMPROVED

35

Adds an explanatory line under ArcFace likeness scores in Samples gallery, LoRA Royale, and Look Consistency Filter noting that 72% is the ceiling (two photos of the same person), so scores are easier to interpret correctly.

— Informational tooltip only, nothing to act on v8.9.7

15 **Profiler 'Open in Studio' button removed** BREAKING

30

The Profiler's Open in Studio button has been removed.

— Bare removal notice, no replacement or migration path given v8.9.7

└──▷ **BREAKING ON UPGRADE**

! The Profiler's Open in Studio button has been removed.

WAS THIS USEFUL? 👍 👎

Unsloth AI Unsloth

Feedback ⓘ v0.1.911-beta NOTES ↗

PRE-RELEASE

Unsloth runs and fine-tunes language and diffusion models locally through a Python library and interface.

Adds scheduled prompt tasks, FLUX.3 Klein image editing, Detailer face-fix pass, multi-GPU tensor parallelism via `UNSLOTH_TP_SIZE`, skills ZIP import/export, saved Run presets, and float8 full fine-tuning.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v0.1.911-beta
https://github.com/unslothai/unsloth.git
# already have the repo? check out this version:
$ git checkout v0.1.911-beta
```

Adds `UNSLOTH_TP_SIZE` environment variable and `torchrun` launch support to split model layers across GPUs for LoRA and QLoRA training.

Adds `UNSLOTH_STUDIO_TRUST_FORWARDED=1` environment variable to enable trust of forwarded client IPs when running behind a reverse proxy.

Adds Scheduled Tasks feature to run a prompt once, daily, on weekdays, weekly, or monthly against a local model or connected API provider, with each answer saved as a new chat.

Adds FLUX.3 Klein image editing in Images > Edit, preserving the original photo's dimensions.

Adds Detailer pass on the Images page to redraw faces, hands, or named regions at higher resolution using SAM 3 and an inpainting model.

+8 more

NEW INTEGRATION MiniMax-M3.1-Flash model support

NEW INTEGRATION NVIDIA Cosmos3-Edge video generation support

└──▷ BREAKING ON UPGRADE

! Unsloth no longer trusts forwarded client IPs by default; without setting `UNSLOTH_STUDIO_TRUST_FORWARDED=1` behind a reverse proxy, all users share one login rate limit.

! Apps using an Unsloth API key can no longer approve their own Python or terminal calls to run outside the sandbox.

WAS THIS USEFUL?  

Lightning AI PyTorch Lightning Feedback ⓘ

CVE-2026-44484 CVE ↗

PyTorch Lightning organizes PyTorch training loops and scales pretraining and fine-tuning across one to thousands of GPUs without changing model code.

`CVE-2026-44484`: Compromise of PyTorch Lightning PyPi Package Versions (CVSS 9.8, affects `>=2.6.2`)

WAS THIS USEFUL?  

Google Keras

Feedback ⓘ CVE-2026-0897 CVE ↗ STABLE

Deep Learning for humans

`CVE-2026-0897` : Critical DoS via Malicious .keras Model (HDF5 Shape Bomb Causes Petabyte Allocation in KerasFileEditor) (CVSS 6.5, affects `3.0.0 to 3.12.0`)

▾ **6 more releases in this issue** · 2025-03-11 → 2026-04-29

> CVE-2026-1669 CVE ↗ STABLE
>
> `CVE-2026-1669` : Local File Disclosure via HDF5 External Storage During Keras Weight Loading (CVSS 7.1, affects `3.11.3`)

> CVE-2025-12060 CVE ↗ STABLE
>
> `CVE-2025-12060` : `keras.utils.get_file` Directory Traversal Vulnerability for TAR files (CVSS 9.8, affects `<=3.11.3`)

> CVE-2025-12058 CVE ↗ STABLE
>
> `CVE-2025-12058` : Load-Time File & Network Access via StringLookup / IndexLookup Vocabulary Paths in Keras 3.11.3 (CVSS 5.5, affects `3.11.3`)

> CVE-2025-9905 CVE ↗ STABLE
>
> `CVE-2025-9905` : The Keras `Model.load_model` method silently ignores `safe_mode=True` and allows arbitrary code execution when a `.h5` / `.hdf5` file is loaded. (CVSS 7.3, affects `3.0.0`)

> CVE-2025-8747 CVE ↗ STABLE
>
> `CVE-2025-8747` : CVE-2025-1550 bypass via reuse of internal Keras functionality (CVSS 8.8, affects `3.9.0`)

> CVE-2025-1550 CVE ↗ STABLE
>
> `CVE-2025-1550` : Arbitrary Code Execution via Crafted Keras Config for Model Loading (CVSS 9.8, affects `>= 3.0.0 < 3.9.0`)

WAS THIS USEFUL? 👍 👎

PyTorch

Feedback ⓘ CVE-2026-24747 CVE ↗ STABLE

Tensors and Dynamic neural networks in Python with strong GPU acceleration

`CVE-2026-24747` : Loading a malicious PyTorch checkpoint with `weights_only=True` can result in arbitrary code execution (CVSS 8.8, affects `<=2.9.1`)

▾ **1 more release in this issue** · 2025-04-17 → 2026-01-26

> CVE-2025-32434 CVE ↗ STABLE
>
> `CVE-2025-32434` : torch.load with `weights_only=True` RCE (CVSS 9.8, affects `<=2.5.1`)

WAS THIS USEFUL? 👍 👎

LLaMA-Factory

Feedback ⓘ CVE-2025-61784 CVE ↗ STABLE

LLaMA-Factory unifies efficient fine-tuning for over 100 language and vision-language models through a command line and web interface.

`CVE-2025-61784` : Critical SSRF and LFI Vulnerabilities in LLaMA Factory's Chat API (CVSS 7.6, affects `<=0.9.3`)

▾ **2 more releases in this issue** · 2025-04-23 → 2025-10-07

> CVE-2025-53002 CVE ↗ STABLE
>
> `CVE-2025-53002` : LLaMA-Factory Remote Code Execution (RCE) Vulnerability (CVSS 8.3, affects `<=0.9.3`)

> CVE-2025-46567 CVE ↗ STABLE
>
> `CVE-2025-46567` : Arbitrary Code Execution via Unsafe Deserialization in llamafy_baichuan2.py (CVSS 6.1, affects `<=0.9.2`)

WAS THIS USEFUL? 👍 👎

ModelScope MS-SWIFT

Feedback ⓘ CVE-2025-41419 CVE ↗

STABLE

MS-SWIFT provides training, fine-tuning, reinforcement learning, evaluation, and deployment workflows for language and multimodal models.

`CVE-2025-41419`: WEB-UI RCE (affects `ms-swift<3.7`)

▾ **1 more release in this issue** · 2025-07-31

> CVE-2025-50460 CVE ↗ STABLE
>
> `CVE-2025-50460`: Remote Code Execution in modelscope/ms-swift via unsafe PyYAML deserialization (CVSS 9.8, affects `ms-swift<3.7`)

WAS THIS USEFUL? 👍 👎

Google TensorFlow Feedback ⓘ CVE-2023-33976 CVE ↗

STABLE

An Open Source Machine Learning Framework for Everyone

`CVE-2023-33976` : Segfault in array_ops.upper_bound (CVSS 7.5, affects `< 2.13.0`)

▾ **12 more releases in this issue** · 2022-11-18 → 2024-07-30

> CVE-2023-25661 CVE ↗ STABLE
>
> `CVE-2023-25661` : Denial of Service in TensorFlow (CVSS 6.5, affects `< 2.12.0`)

> CVE-2023-25669 CVE ↗ STABLE
>
> `CVE-2023-25669` : FPE in AvgPoolGrad with XLA (CVSS 7.5, affects `< 2.12.0`)

> CVE-2023-25671 CVE ↗ STABLE
>
> `CVE-2023-25671` : Segmentation fault in tfg-translate (CVSS 7.5, affects `< 2.12.0`)

> CVE-2023-25801 CVE ↗ STABLE
>
> `CVE-2023-25801` : Double free in Fractional(Max/Avg)Pool (CVSS 7.8, affects `< 2.12.0`)

> CVE-2023-27579 CVE ↗ STABLE
>
> `CVE-2023-27579` : FPE in TFLite in conv kernel (CVSS 7.5, affects `< 2.12.0`)

> CVE-2023-25662 CVE ↗ STABLE
>
> `CVE-2023-25662` : Integer overflow in EditDistance (CVSS 7.5, affects `< 2.12.0`)

> CVE-2022-41902 CVE ↗ STABLE
>
> `CVE-2022-41902` : OOB write in grappler (CVSS 9.1, affects `< 2.11.0`)

> CVE-2022-41907 CVE ↗ STABLE
>
> `CVE-2022-41907` : Overflow in `ResizeNearestNeighborGrad` (CVSS 7.5, affects `< 2.11.0`)

> CVE-2022-41909 CVE ↗ STABLE

`CVE-2022-41909`: Segfault in `CompositeTensorVariantToComponents` (CVSS 7.5, affects `< 2.11.0`)

`CVE-2022-41911` CVE ↗ STABLE

`CVE-2022-41911`: Invalid char to bool conversion when printing a tensor (CVSS 7.5, affects `< 2.11.0`)

`CVE-2022-35935` CVE ↗ STABLE

`CVE-2022-35935`: `CHECK` failure in `SobolSample` via missing validation (CVSS 7.5, affects `< 2.11.0`)

`CVE-2022-35991` CVE ↗ STABLE

`CVE-2022-35991`: `CHECK` fail in `TensorListScatter` and `TensorListScatterV2` in eager mode (CVSS 7.5, affects `< 2.11.0`)

WAS THIS USEFUL? 👍 👎

AGENT

◆ AI Agent Frameworks

mateclaw

Feedback ⓘ v3.0.0 NOTES ↗ RANK

MateClaw — Your second brain with Multi-Agent Orchestration, MCP Protocol, Skills & Memory, Dream, and Multi-Channel Support. Built on Spring AI Alibaba.

> MateClaw 3.0.0 is a breaking release that moves execution onto AgentScope Java 2.0.4, replaces Alibaba AI with native DashScope adapters, adds opt-in per-user file isolation and a governed decision pipeline, and tightens JWT and database-migration requirements.

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

01 Opt-in member file isolation in workspace sandbox NEW

93

Opt-in strict member file isolation via `MATECLAW_WORKSPACE_SANDBOX_MEMBER_ISOLATION_ENABLED=true` restricts file roots to the authenticated user/workspace identity, limiting access to six audited builtin file/code tools; DSH, MCP/plugin tools, shared skill scripts and unauthenticated channel identities are denied.

Enforce per-user file isolation in a multi-tenant workspace so each authenticated user can only access their own file root.

```
$ MATECLAW_WORKSPACE_SANDBOX_MEMBER_ISOLATION_ENABLED=true java -jar mateclaw-server.jar
```

— Names exact env var, exact tool-count scope and denied surfaces v3.0.0

02 Mandatory minimum-length JWT_SECRET BREAKING

85

Every backend now requires a non-empty `JWT_SECRET` of at least 32 UTF-8 bytes; desktop local mode generates and persists its secret automatically, while source and server deployments must configure and retain theirs. Deployments without a compliant secret are rejected at startup.

Set a compliant JWT secret for a server deployment before starting the backend to satisfy the new mandatory minimum-length requirement.

```
$ JWT_SECRET=your-32-plus-utf8-byte-secret-here java -jar mateclaw-server.jar
```

— Exact env var, length requirement, and startup behavior given v3.0.0

03 **Native DashScope adapters replace Alibaba AI integration** 75

BREAKING

Replaces Alibaba AI integration with native DashScope chat and embedding adapters, retaining the native speech SDK; Spring AI Alibaba starter and Graph dependencies are removed, so any configuration or code referencing them will fail to start and existing Alibaba AI adapter configuration must be migrated.

— Names removed dependencies and migration need, no exact config keys v3.0.0

04 **Governed Choice/Boolean/Score decision pipeline** NEW 70

Introduces a bounded typed Choice / Boolean / Score decision pipeline with durable audit, OFF / SHADOW / ACTIVE modes and scenario overrides; SHADOW is the default and suggestions cannot bypass domain constraints.

— Modes and default named, no config key or command given v3.0.0

05 **ApplicationGuardedToolkit for tool batch delegation** NEW 70

Adds `ApplicationGuardedToolkit` to delegate tool batches to the application executor, retaining approvals, workspace permissions, audit and cancellation.

— Named class with retained guarantees, no usage example v3.0.0

06 **Goal tool checkpoints for execution identity** NEW 65

Adds Goal tool checkpoints that persist execution identity before side effects, with uncertain outcomes blocking blind replay and retaining approval recovery and session isolation.

— Behavior described in detail, no config surface named v3.0.0

07 **AgentScope Java 2.0.4 execution engine** IMPROVED 60

Integrates AgentScope Java 2.0.4 to drive ordinary ReAct and the reasoning/tool/observation loop inside each Plan-Execute step, with MateClaw retaining overall plan scheduling, Goal budgets and team coordination.

— Mechanism and division of responsibility described, no commands v3.0.0

08 **Structured MCP result presentation below final answer** 60

BREAKING

Adds structured MCP result presentation below the final answer; ordinary tool results no longer automatically become final-answer text, so integrations relying on that promotion

will need to be updated.

— Clear before/after behavior, no migration steps given v3.0.0

THINNER COVERAGE BELOW

09 **MCP SDK 0.18.3 upgrade** IMPROVED 55

Upgrades to MCP SDK 0.18.3, preserving call IDs, structured results and protocol errors; local execution capabilities stay out of wire metadata.

— Version and scope named, no concrete action for reader v3.0.0

10 **Flyway migration strictness and rollback requirements** 55

BREAKING

Flyway validation failures block startup by default; previously permissive upgrade paths that relied on automatic history repair now halt until explicit opt-in is configured. Downgrading an upgraded database is not verified — rollback requires restoring the pre-upgrade database backup rather than running the old binary against the upgraded schema.

— Operational consequences given, no exact opt-in flag named v3.0.0

11 **Voice lists for Edge and DashScope read-aloud** NEW 40

Adds provider voice lists, selection and preview for Edge and DashScope read-aloud.

— Brief feature description with named providers only v3.0.0

└──▷ **BREAKING ON UPGRADE**

! Spring AI Alibaba starter and Graph dependencies are removed; any configuration or code referencing them will fail to start.

! Native DashScope chat and embedding adapters replace the former Alibaba AI integration; existing Alibaba AI adapter configuration must be migrated.

! Every backend now requires a non-empty `JWT_SECRET` of at least 32 UTF-8 bytes; deployments without one will be rejected at startup.

! Flyway validation failures block startup by default; previously permissive upgrade paths that relied on automatic history repair will now halt until explicit opt-in is configured.

! Ordinary tool results no longer automatically become final-answer text; integrations that relied on that promotion to surface tool output will need to be updated.

! Downgrading an upgraded database is not verified; rollback requires restoring the pre-upgrade database backup rather than running the old binary against the upgraded schema.

WAS THIS USEFUL? 👍 👎

magic-context

Feedback ⓘ dashboard-v0.20.20 NOTES ↗ RANK

Unbounded context. Memory that manages itself.

> Magic-context's dashboard expanded Broca log support, adding new message-part types for provider fallbacks and refusals, and fixing a timeline truncation bug tied to tool schema pins.

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

01 **Fix for run timeline truncation on tool schema pins** IMPROVED

55

The dashboard now supports the `tool-schema-pins/v1` Broca log record type, preventing run timelines from being truncated at the first occurrence of that record.

— Names the record type and the before/after bug it fixes dashboard-v0.20.19

02 **Broca log support for provider and refusal fallbacks** NEW

52

The dashboard now reads Broca logs carrying the `provider-fallback-block/v1` message part, which records a provider's server-side switch to another model, and the `refusal-fallback/v1` message part, which records a refused model step and which model actually served the finished step.

— Names both log record types but no UI or usage detail dashboard-v0.20.20

WAS THIS USEFUL? 👍 👎

agentlas-os

Feedback ⓘ v1.2.70 NOTES ↗ RANK

Agent OS: keep specialist agents in a hub, spin up a temporary orchestrator per task. Local-first, works with any model.

> Agentlas OS v1.2.70 lets prepared agents use the host's native delegation mechanism directly and introduces package read allowlists for describing package contents.

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

01 **Native host delegation for prepared agents** NEW 40

Prepared agents can now use the current host's native delegation when it is available and authorized, allowing delegation without switching to another application.

— Describes behaviour but no flag or config name given v1.2.70

02 **Package read allowlists for package contents** NEW 30

Package read allowlists now describe package contents, while project access and tools continue to follow the host's permissions and explicit restrictions.

— Named feature but no concrete config surface or command shown v1.2.70

WAS THIS USEFUL? 👍 👎

cotal

Feedback ⓘ v0.109.6 NOTES ↗ RANK

The open standard for agent coordination

> Cotal shipped a new `cotal manager` / `cotal service` command surface for running built-in and custom services, full Windows support for `cotal up` / `cotal down`, and a wave of auth hardening — per-owner seat caps, per-actor/owner/host connect rate limiting, credential revocation, and on-demand presence reads — alongside daemon memory reductions and dozens of CLI refinements across 23 releases.

└──▷ **WHAT SHIPPED · 49 FEATURES** most completely described first ? what's the number?

01 `cotal manager` and `cotal service` command groups NEW

95

Adds `cotal manager start` to start the manager for the current folder or selected space with the pty runtime and its registered broker, plus `cotal manager stop`, `status`, `logs`, and `restart`. Adds `cotal service start <preset>` to start a built-in service from its defaults (refusing with exact flags when a required file has no default), `cotal service` to list every built-in service with its start command, and `cotal service add <name> --exec "<command>"` to register an arbitrary program as a service in one line, splitting words without a shell and refusing pipes, redirection, variables, and globs. Adds `cotal manager install`, which refuses before any effect and names a unit running `cotal manager start --foreground`. `cotal manager start` / `restart` install first-party connectors on first run as `cotal supervise` did, and after `cotal up` the manager is reported as already started rather than launching a second instance. `cotal supervise` was deprecated for removal in v1.0.0, printing a stderr line naming `cotal manager start` as the replacement unless `--execution`, `--spawn`, `--seat-cap`, or an unsupported resume flag is passed.

Start the manager for the current space, then check its status — replacing the old `cotal supervise` workflow.

```
$ cotal manager start
cotal manager status
```

Start a built-in service by preset name to bring up a first-party integration without writing a definition file.

```
$ cotal service start <preset>
```

Register a custom program as a managed service in one line without a definition file.

```
$ cotal service add my-worker --exec "node dist/worker.js --port 4000"
```

— Full subcommand set with exact flags and examples v0.108.0

02 **Per-owner seat caps on managers** NEW 95

Adds `--seat-cap <n>` flag to `cotal supervise` and `cotal up` to limit how many live seats a manager holds for one owner; defaults to `unlimited` (manager's own 50-slot ceiling applies), and the cap persists in the root so `repair`, `resume`, or a bare `up` after `down` retains it. Returns a `resource-exhausted` structured refusal when a launch exceeds the cap, carrying the owner, current count, and cap; `cotal spawn` exits 2 with `CONFLICT` and, under `--json`, exposes the three numbers in `error.cause.detail`.

Start a shared manager and cap each owner to 10 concurrent live seats so a single tenant cannot exhaust the pool.

```
$ cotal supervise --seat-cap 10
```

Bring a mesh up with a per-owner seat cap, ensuring the limit persists across `down` / `up` cycles without re-specifying it.

```
$ cotal up --seat-cap 10
```

Parse the seat-cap refusal in CI to read the owner, count, and cap programmatically when a spawn is rejected.

```
$ cotal spawn --json 2>&1 | jq '.error.cause.detail'
```

— Exact flag, default, persistence, and refusal fields detailed v0.103.1

03 **Moved/copied folder conflict detection** IMPROVED 90

Adds a `TARGET_UNAVAILABLE` (exit 2) refusal to `cotal clean store` and `cotal clean all` (including dry runs) when the stack runs from a different folder, naming the conflicting folder and step instead of silently deleting `.cotal/nats` and exiting 0. Adds a `CONFLICT` refusal to bare `cotal up` in a copied or moved folder instead of silently taking over the space name, surfacing `cotal up --space <space>` as the explicit escape hatch. Refusal messages from `ps` and `send` in such folders now

include a `Mesh: <space>` header and `target.space` field, `cotal status` stops offering the `clean store` action, and `@cotal-ai/workspace` gains `recordsElsewhereError`, `targetErrorSpace`, and `folderSpaceStopped` exports.

Recover a moved or copied project folder whose broker is unreachable — use `--space` to explicitly take the space name over to the current folder.

```
$ cotal up --space <space>
```

— Full refusal codes, exports, and escape-hatch command named `v0.109.5`

---

04 **Runtime targeting and validation in `cotal spawn --runtime`**

IMPROVED 90

Adds `--runtime <name>` flag to `cotal spawn` (also `-d` / `--on`), a `runtime` key to `cotal_spawn`, and a `runtime` option to the workflow spawn(persona, { runtime }) so each seat can target a different terminal runtime independently of the manager's own runtime. Later adds pre-launch runtime validation to `cotal spawn -d --runtime <name>`: rejects unknown, uninstalled, or unreachable runtimes with `DEPENDENCY_UNAVAILABLE` before the launch is sent (when targeting the local manager), adds `--dry-run` support reporting `result.requested.runtime`, the host checked, and a `RUNTIME_UNAVAILABLE` warning when it cannot run, and surfaces a remedy package name via `cotal ext add` for a known-but-uninstalled runtime.

Spawn a seat on a specific runtime without changing the runtime the manager was started with — useful when running seats across heterogeneous terminal environments from a single manager.

```
$ cotal spawn --runtime docker-pty myPersona
```

Verify a named runtime is reachable on the local manager before committing to a spawn, to catch missing extensions early in CI.

```
$ cotal spawn -d --runtime myruntime --dry-run
```

— Flag, dry-run output fields, and remedy path all named `v0.109.1` `v0.107.0`

---

05 **Auth-callout timeout configuration and deadline enforcement**

NEW 90

Adds `cotal up --auth-timeout <seconds>` (1–60) to set the broker's auth-callout timeout, recorded in `.cotal/broker-policy.json` as `authorization {`

`timeout: N }` and passed to the auth-service daemon as `--auth-timeout <seconds>`, with `serverConfig` exposing it as `authTimeout`. Auth callout now enforces a per-request deadline (broker timeout minus 250 ms) and returns a signed denial with reason `timeout: the auth callout could not answer inside the broker's authorization timeout` when exceeded; `startAuthCallout` takes `limits: { authTimeoutMs, maxInFlight, slowMs }` (defaults 2000, 32, 500) and `startAuthService` takes the same as `calloutLimits`.

Set a 10-second auth-callout timeout when bringing up a user-auth space, so slow callouts get a signed denial instead of a silent nats-server refusal.

```
$ cotal up --auth-timeout 10
```

— Flag, config path, and default limit values all given v0.101.6

06 **Per-actor/owner/host connect rate limiting** NEW 87

Adds connect rate limiting to the auth callout: 60 connects per minute per actor, 600 per minute per owner, and 60 per minute per client host for unverified bearers; limited connects receive a signed `RATE_LIMITED retry-after <n>s` denial, and the exchange answers HTTP 429 with `{code: "RATE_LIMITED", retryAfterSeconds, error}` and a Retry-After header. Exposes `connectLimits` on `startAuthService` so a hosted composition can configure its own actor, owner, and host limits. Rate-limit responses from the exchange now name the exhausted bucket as `limit`, `actor`, or `owner`; `cotal agent-bearer` prints `retry-after <n>` on stdout and exits 2, and endpoints respect a `retryAfterSeconds` before retrying a bearer source.

— Exact per-minute limits and HTTP 429 response shape given v0.102.0 v0.101.7

07 **`cotal manager logs` reads `cotal up` managers** IMPROVED 84

Extends `cotal manager logs` to read `<root>/.cotal/manager.<spaceKey>.log` for managers started by `cotal up`, with `--tail`, `--follow`, and `--json` flags; a `--follow` session ends automatically when the manager process exits.

Stream logs in real time from a manager that was started with `cotal up`, filtering to JSON output for machine processing.

```
$ cotal manager logs --follow --json
```

Tail the last N lines of a `cotal up` manager log without streaming, useful for a quick post-incident review.

```
$ cotal manager logs --tail 100
```

— Exact flags and log path given with two runnable commands v0.109.6

08 **`cotal auth revoke-credential` single-key revocation** NEW

84

Adds `cotal auth revoke-credential <public-key | creds-file>` to stop a single minted credential without rotating the whole space identity via `reset space`; records the key in the space account JWT `revocations` so the broker refuses it on next start, and works on a stopped space.

Revoke a single leaked minted credential by public key without rotating the entire space identity.

```
$ cotal auth revoke-credential ./leaked-creds.creds
```

— Exact command and where revocation is recorded v0.101.0

09 **On-demand presence reads via `readPresence()`** BREAKING

83

Adds CotalEndpoint.readPresence() performing batched Direct Gets of the presence bucket without binding a consumer, using the same reader rules as the watch and treating records older than the liveness window as offline. Agent seats now read presence on demand instead of watching continuously, eliminating per-heartbeat presence deliveries for idle spaces. Whole-bucket Direct Get reads, including `readOnce`, now page past 1024 keys instead of refusing the read. Agent credentials now require the `$JS.API.DIRECT.GET.KV_cotal_presence_<space>` grant; credentials minted before this release are refused with the grant named until the manager and auth service are upgraded.

— API method and exact grant string named, no CLI command v0.106.0

10 **Reused issuer sessions for auth latency** NEW

82

Adds `holdIssuerSession` export to `@cotal-ai/core` and an optional `issuerSessionTtlSeconds` parameter (5–300 s) to `startAuthService`, replacing per-connect credential minting with a reused session that cuts auth latency from ~15 ms to near-zero per issuance (over 200 ms saved under load).

Tune issuer session lifetime to match your auth plane's credential policy — set a shorter TTL to rotate sooner, or up to 300 s to minimize re-dial overhead under sustained load.

```javascript
startAuthService({ issuerSessionTtlSeconds: 120 })
```

— Export, param range, and latency numbers with code example `v0.102.0`

---

11 **Linear serve approval flow on user-auth meshes** NEW

81

`cotal linear serve` on a user-auth mesh now names the waiting proposal ID and prints the exact approval command `cotal service-endpoint approve <proposalId> --caller-ceiling linear.mcp --space <space>`, where `--caller-ceiling linear.mcp` is required to grant callers access to the five Linear commands.

Approve a pending Linear serve proposal on a user-auth mesh, granting callers access to all five Linear commands via the required ceiling flag.

```
$ cotal service-endpoint approve <proposalId> --caller-ceiling linear.mcp --space <space>
```

— Exact approval command with required flag given verbatim `v0.107.0`

---

12 **Legacy command cleanup and `cotal help` listing** DEPRECATED

80

Legacy commands (`clean`, `join`, `topology`, `endpoints`, `describe`, `invoke`, `console`, `actor`) now self-identify as legacy on their `--help` pages and name their replacements. They were then removed from `cotal help` and shell completion, each mapped to its replacement: `join` → `chat`, `endpoints` → `who --all`, `describe` → `endpoint describe`, `invoke` → `endpoint invoke`, `console` → `watch`, `topology` → `team check --graph`, `clean` → `reset` / `history clear` / `backup cleanup-fallback`, `actor` → `auth grants` / `auth grant` / `auth revoke`. `cotal help` (and `cotal --help` / `cotal -h`) now lists every command group without any flag, bare `cotal` keeps the compact entry screen pointing to `cotal help`, and `cotal help --all` still prints the same full list.

— All eight legacy commands and their replacements named `v0.108.1` `v0.105.0` `v0.101.3`

---

13 **Auth-callout connect-load harness (`pnpm smoke:callout-load`)**

IMPROVED

78

Adds `pnpm smoke:callout-load` connect-load harness for the auth callout, reporting connects attempted/succeeded/refused by reason, callout handler and client connect latency percentiles, auth-service and broker CPU per connect, and RSS — with a second

arm replaying the callout from v0.99.3 (sha256-pinned) to measure gains from native Ed25519 signing and bounded concurrency. Extends it with `--callouts N` to run multiple callout-service processes behind the existing queue group for horizontal scaling measurement, and with `--connect-limits <actor>,<owner>,<host>` to configure capacity-test rate limits.

Benchmark auth callout throughput and latency — including a comparison against the v0.99.3 baseline — to quantify gains from native Ed25519 signing and bounded concurrency.

```
$ pnpm smoke:callout-load
```

— Runnable command with metrics and tuning flags listed `v0.101.7` `v0.101.6` `v0.101.4`

---

14 **Interactive `cotal spawn` prompts for manager/folder** NEW

78

On a user-auth mesh with several live managers, a bare `cotal spawn` in a terminal now interactively asks where the agent runs (this terminal, the default, or a named manager), with `--json`, `--no-input`, and `COTAL_NO_PROMPT=1` suppressing all prompts. `cotal spawn -d` with no `--on` additionally prompts for the working folder when the picked manager is on another host, pinning the launch to that manager.

Launch an agent on a user-auth mesh without pre-specifying a manager — Cotal will interactively ask which manager to use and which folder to work in; pass `--no-input` in CI to suppress all prompts.

```
$ cotal spawn --no-input
```

— Exact flags and env var to suppress prompts `v0.101.3` `v0.101.1`

---

15 **Lazy daemon module loading cuts memory** IMPROVED

75

Adds `provide`, `provides`, and `load` to the `@cotal-ai/core` registry so each daemon entry imports only the module behind the key it requests, once, refusing keys whose module did not register them. Adds `./dispatch` export to `@cotal-ai/cli` exposing `commandWord` and `runCli` without the full command surface, and `./service-unit` export exposing the unit entry alone. Scopes lazy loading to the `supervise`, `deliver`, `feedback-intake`, `auth-service`, `__daemon-*`, `__service-*`, and `__service-unit` daemon entries. Reduces manager private memory from 148 MB to 106 MB at startup (112→75 MB settled), delivery daemon from 110 MB to 65 MB (105→60 MB settled), and user-auth service from 115 MB to 74 MB (105→62 MB settled).

— Precise memory numbers but no user-facing action `v0.109.0`

---

16 **`programHash` on run status output** NEW 75

Adds `programHash` field to `run-status` and `cotal run show --json` output, exposing the same hash value a `run-ps` row carries so a single run's program identity can be checked without scanning all runs; returns `programUnreadable` with a reason when the program record cannot be read or parsed.

Verify a specific run's program identity without scanning every run via `run-ps`.

```
$ cotal run show --json | jq '.programHash'
```

— Exact field name and jq command to read it v0.103.3

17 **Callout-only auth-service replica** NEW 75

Adds `cotal auth-service --callout-only` to run a callout-only auth-service replica that joins the space's auth callout queue group and handles connect checks, holding no HTTP listener, `auth-service.json`, plane claim, registration, barrier, scanner, or sweep.

Run a callout-only replica to keep connects answered while the full auth service restarts or to add callout capacity under load.

```
$ cotal auth-service --callout-only
```

— Exact command and what it omits clearly listed v0.101.6

18 **Windows process-service support** NEW 75

Runs `cotal up` and `cotal down` on Windows, with each background service running under a detached custodian that creates the program inside a job object (no breakaway allowed, all members terminated when the custodian exits). Reaches each custodian over a named pipe with a per-instance capability proved on every request; malformed proofs or frames over 64 KiB are dropped while the custodian keeps serving. Records machine identity using `GetSystemIdForPublisher` (sourced from TPM, UEFI, or registry) alongside the account SID and custodian pin; a service start refuses when Windows has none of those sources.

Check whether a Windows service's job still holds live child processes after the main program has exited.

```
$ cotal service status
```

— Named APIs and exact frame-size limit, little user action v0.101.1

19 **`waitsForOperator` connector screen handling** NEW 73

Allows connectors to declare `waitsForOperator` to name the screen a managed launch is waiting at; `cotal spawn` now reports the agent as waiting at that screen, attaches to it in a terminal (including with `-d`), or prints the attach command and exits 0 — instead of reporting an unknown outcome. Surfaces an `ai.cotal.goal.entity-running` detail in the `uncertain` result when the readiness window passes but the agent process is still running, with `cotal spawn` explaining how to attach.

Spawn a Claude agent that opens its session picker so an operator can attach; `cotal spawn` will report the agent as waiting and attach automatically.

```
$ cotal spawn --opt resume=
```

— Named declaration and detail field, with example command v0.101.4

---

20 **Byte-verified incremental `cotal update`** IMPROVED 72

Adds byte-for-byte verification to `cotal update` so only changed built-in connectors and the web dashboard are reinstalled, rather than all seven on every run; `cotal ext seed --force` still triggers an unconditional full reinstall.

Force a full unconditional reinstall of all built-in connectors regardless of byte verification, useful after a corrupted install.

```
$ cotal ext seed --force
```

— Names the force-reinstall escape hatch and scope of change v0.106.1

---

21 **Multi-manager `cotal ps` and `cotal who`** IMPROVED 72

`cotal ps` on a user-auth mesh now reads each live manager independently and lists agents under a `manager <id>` header with presence-reported host; `--json` reports a presence-only read as `result.kind` `presence`. `cotal ps` and `cotal who` on a user-auth mesh name unresponsive managers by instance ID in their Next, distinguishing stale registrations from live-but-silent managers.

Inspect which managers are live, silent, or left stale registrations without prompts — useful after a manager crashes without a clean stop.

```
$ cotal ps --json
```

— Named header format and --json result kind v0.102.0

---

22 **Agent name conflict handling and `cotal auth retire`** NEW

72

Adds `cotal auth retire <uid>` to release a held lifecycle on a mesh when a name collision blocks a spawn. `cotal spawn` now handles already-taken agent names: attaches in a terminal, prints the `cotal attach` command elsewhere, or releases a non-running lifecycle automatically, waiting up to 2 minutes for the mesh to confirm retirement.

Release a lifecycle whose name is blocking a new spawn — use the UID reported in the `AGENT_LIFECYCLE_HELD` conflict message.

```
$ cotal auth retire <uid>
```

— Exact command with UID argument and wait behavior `v0.101.3`

---

23 **Prepared-but-not-running space handling** BREAKING

71

Adds a `FOLDER_SPACE_NOT_RUNNING` warning to `--json` output of `cotal status --json` and `cotal space --json` when a folder holds a prepared or stopped space, surfacing the same note in `cotal chat`. A folder's space no longer captures every command run in that folder when prepared but not running; commands now target the active `cotal use` selection instead. `cotal update` and `cotal update --self` no longer refuse updates over spaces prepared but never started, and now read multiple registrations in a folder one at a time by name.

Check what space a stopped folder's spawn would target before committing to launch it.

```
$ cotal spawn -d --dry-run
```

— Named warning field and behavior change, limited command detail `v0.108.1` `v0.106.0`

---

24 **Faster `cotal agent-bearer` startup** IMPROVED

71

`cotal agent-bearer` now runs before the rest of the CLI loads, dropping per-invocation startup time from ~400 ms to ~60 ms; the command is served from `@cotal-ai/auth/agent-bearer` with parser and checks exported from `@cotal-ai/core/command`, `@cotal-ai/core/loopback`, `@cotal-ai/core/subjects`, and `@cotal-ai/workspace/liveness` without loading either package root.

— Precise latency numbers and internal module paths named `v0.101.7`

---

25 **Managed handoff file restricted to `cotal spawn`** BREAKING

71

Only `cotal spawn` reads `COTAL_MANAGED_HANDOFF_FILE` (other commands now drop it and leave the file intact, refusing an unusable file with `USAGE_INVALID` naming

the variable, path, and `unset COTAL_MANAGED_HANDOFF_FILE` as next step), and `cotal spawn` with `COTAL_MANAGED_HANDOFF_FILE` now requires `COTAL_MANAGER_INSTANCE` (`config.managerInstanceId`) before contacting any broker or exchange plane, matching `MeshAgent` 's arming of the `turn-pending` poll only when that variable is set.

— Exact env vars and error code named, no example v0.101.0

---

26 **Structured `cotal watch` refusal errors** IMPROVED 69

Adds `--json` error record output to `cotal watch` refusals, carrying the error code, message, and next step instead of a generic `OPERATION_FAILED` sentence.

Catch a refused watch target in CI by parsing the structured error code and next step from the stream instead of scraping stderr text.

```
$ cotal watch ./my-space --json 2>&1 | jq 'select(.error) |
{code: .error.code, message: .error.message, next:
.error.next_step}'
```

— Runnable jq command shows exact error shape v0.105.0

---

27 **`space` argument for `cotal_connect`** NEW 69

Adds an optional `space` argument to `cotal_connect` so personal Claude Code conversations can select a registered mesh without changing project configuration.

Connect a single Claude Code conversation to a specific mesh space without touching project config.

```
$ cotal_connect space=<your-space-name>
```

— Exact argument name and runnable example given v0.103.4

---

28 **`cotal status` output restructuring** IMPROVED 67

Adds `status --wide` flag to `cotal status` to report every row in full alongside the existing `status --all` . Adds `meshes.catalogs[].account` field to `cotal status --json` output, naming the account on error lines. `cotal status` output is restructured to show version, a single Agent applications line with a count, a shared Extensions line, and a Spaces section replacing the Brokers section.

Inspect your space at a glance without noise — see only what needs attention, then expand to every row when troubleshooting.

```
$ cotal status --wide
```

— Flag and field named, exact command to try it v0.106.0

29 **Structured refusal output for `cotal down`** IMPROVED 67

Adds `--json` and `--plain` flags to `cotal down` to carry refusal output as structured data instead of returning `UNSUPPORTED`, enabling scripts to read the same code as a plain run.

Script a `cotal down` refusal check by reading structured output instead of parsing plain text or catching `UNSUPPORTED`.

```
$ cotal down --json
```

— Both flags named with a scriptable example v0.103.4

30 **`cotal docs` category metadata and terminal rendering**

IMPROVED 66

`cotal docs` now adds a `category` field to each page in `docs --json` (Start here, Guide, Concept, Reference, Project, alongside the existing `kind` field) and renders `cotal docs <page>` for direct terminal reading with wrapping, title-styled headings, and table-to-record fallback, while piped output and `--plain` or `--json` stay as Markdown.

Read a reference doc page in the terminal with readable wrapping and headings instead of raw Markdown.

```
$ cotal docs cli
```

Retrieve docs page metadata including the new `category` field for tooling or indexing purposes.

```
$ cotal docs cli --json
```

— Named field and two runnable doc commands v0.101.0

31 **Structured auth-callout logging** NEW 65

Auth callout writes one `auth callout: level=… outcome=…` log line per answered request, including per-connect user key, broker client id, owner, actor, outcome, and milliseconds spent queued and in verify, ledger, policy, and sign; logs at warn over 500 ms or on timeout/shed.

— Exact log line fields and warn threshold named `v0.101.6`

---

32 **Expanded shell completion coverage** IMPROVED 65

Shell completion now offers offline values for `--space`, `config get` / `set` / `unset` keys, `config set color` and `config set harness` values, `docs` page names, `watch --view` views, `space add --mode` modes, and channel names from `history`, `channel show` / `members` / `set` / `remove` and `send -c`, and completes file names for every `--creds` flag (including `status`, `history`, `watch`, and `who`) and for `run start --file`, `run simulate --handler`, `auth mint --out`, and `--identity`, via the new optional `FlagSpec.completion: "files"` declaration.

— Long list of flags completed, no single example `v0.101.0`

---

33 **`cotal up` refuses missing `--user-auth` on existing state** IMPROVED 64

`cotal up` without `--user-auth` on a stopped root that has user-auth state now refuses immediately as `CONFLICT` (exit 2) before writing anything, with a Next hint to add `--user-auth`.

— Exact refusal code and hint flag named `v0.101.7`

---

34 **Accurate `cotal spawn` failure exit codes** IMPROVED 64

`cotal spawn` now reports a manager-confirmed failed launch as `OPERATION_FAILED` (exit 5) with the manager's reason, instead of the misleading `OPERATION_UNKNOWN` (exit 3) / Outcome unconfirmed.

Check whether a `cotal spawn` invocation succeeded or failed in a CI script by testing the exit code.

```
$ cotal spawn <agent>; echo "Exit: $?"  # 0 = success, 5 =
OPERATION_FAILED (manager rejected launch)
```

— Exact exit codes and example exit-check given `v0.101.5`

---

35 **Automatic bearer renewal for long-lived connections** IMPROVED 62

Long-lived connections now auto-renew their bearer at a random point between half and four-fifths of the bearer lifetime, staggered at least one second apart per process,

eliminating broker disconnects due to expired bearers; renewal failures retry on a growing randomized wait of up to one minute and are reported once.

— Detailed renewal timing but no user-facing control v0.101.7

THINNER COVERAGE BELOW

36 **Push-based turn delivery replaces polling** IMPROVED

56

Turn delivery switches from a 15-second poll to push notices: the manager publishes a data-free notice on each seat's reply rail when pending turns change, repeating every 5 seconds until the seat pulls `turn-pending`.

— Clear before/after mechanism but no user-facing control v0.102.0

37 **Accurate `PERMISSION_DENIED` on broker refusal** IMPROVED

56

Returns `PERMISSION_DENIED` (exit 2) — instead of `OPERATION_UNKNOWN` (exit 3) — when a manager command such as `cotal spawn --on` or `cotal ps --on` is refused by the broker at connect, with the broker's reason preserved in the message.

— Exact exit code change named, brief v0.101.1

38 **`COTAL_SERVERS` credential-leak validation** IMPROVED

55

A connector session now refuses a `COTAL_SERVERS` value containing a user name or password at startup, naming the variable and directing users to `COTAL_CREDS`, `COTAL_TOKEN`, or a `COTAL_LINK` instead.

— Exact env vars named, no example v0.101.0

39 **`decodeNatsJwt` export in core** NEW

53

Exports `decodeNatsJwt` from core, performing the same JWT checks as `@nats-io/jwt`'s `decode` but with a native Ed25519 signature check.

— Export named but no usage example given v0.101.5

40 **Fast-fail `cotal run --on` target resolution** IMPROVED

52

A hosted `cotal run` command whose `--on` names none of the space's running managers is now refused `TARGET_NOT_FOUND` immediately, naming the pin and available managers, rather than waiting out a ten-second deadline before refusing `TARGET_UNAVAILABLE`.

— Exact error codes named, no example command v0.101.0

41 **Explicit dial timeout and connection retry** IMPROVED

51

Manager commands gain automatic connection retry; adds `transportConnectTimeoutMs` option on an endpoint for its first connection attempt, giving callers explicit control over dial timeout.

— Option named but no example or range given v0.101.2

---

42 **`standaloneConnectOpts` creds getter and `wireCreds()`** NEW

50

Adds `standaloneConnectOpts` support for a creds getter that is read on every connection attempt, and CotalEndpoint.wireCreds() to return the endpoint's checked credential.

— Names two APIs but no usage detail or example v0.102.0

---

43 **`cotal supervise` un-deprecated with restored space targeting**

IMPROVED

50

`cotal supervise` is no longer deprecated: re-listed under its own Run a manager group, issues no deprecation warning, and will not be removed in 1.0.0. Without `--space`, it now targets the space every other command targets — `COTAL_SPACE`, then the folder's hosted/registered space, then the space selected by `cotal use` — instead of defaulting to the default space.

— Describes reversal and fallback order, no example v0.101.3

---

44 **Bounded tool-result journal and capture frontier** NEW

48

Adds a bounded tool-result journal for SPEC 13.19 in connector-core, capping how much tool output is retained across agent turns, and places the SPEC 13.19.2 capture frontier marking the boundary up to which tool results are recorded.

— Spec numbers named but no observable surface for users v0.104.2

45 **`cotal up` TLS/plaintext address validation** IMPROVED

45

`cotal up` now catches a TLS/plaintext address mismatch early, refusing before starting the broker and naming the exact address and policy file to fix, rather than failing during broker startup.

Catch a TLS/plaintext address mismatch early — cotal refuses before starting the broker and tells you exactly which address and policy file to fix.

```
$ cotal up --server nats://0.0.0.0:4222
```

— Only an example illustrates it, no bullet detail v0.105.0

46 **V8 semi-space flag trims manager memory** IMPROVED 38

Daemons now run under `--max-semi-space-size=1`, cutting manager process private memory from 268 MB to 147 MB at startup.

— Single number and flag name, no further mechanism v0.108.1

47 **Native Ed25519 signer for NATS credentials** IMPROVED

38

Uses Node's native Ed25519 signer for NATS user credentials, auth-callout responses, and authority-client renewals, reducing signing CPU cost.

— Mentions mechanism broadly, no numbers or surface v0.101.2

48 **No-op detection in `cotal channel set`** IMPROVED 37

`cotal channel set` and `channel defaults` now detect no-op writes and report Channel unchanged or Channel defaults unchanged instead of rewriting the entry.

— One-line behavior change, minimal detail v0.101.0

49 **Claude Code native approval completion** NEW 35

Enables the Claude Code connector to finish a pending native approval and report the outcome back into the conversation.

— One sentence, no mechanism or surface named v0.104.2

└──▷ **BREAKING ON UPGRADE**

! `cotal supervise` is removed in v1.0.0; migrate to `cotal manager start` for standard runs.

! Agent credentials minted before v0.106.0 lack the `$JS.API.DIRECT.GET.KV_cotal_presence_<space>` grant; their presence reads are refused with the grant named. Upgrade the manager and auth service before upgrading connectors.

! A folder's space no longer captures every command run in that folder when it is prepared but not running; commands now target the active `cotal use` selection instead.

! `DelegatedSeatLaunch.managerInstanceId` is now required; a CLI with this change refuses a managed handoff without `COTAL_MANAGER_INSTANCE`, so managers and the CLI must be upgraded together.

WAS THIS USEFUL?  

PydanticAI

Feedback ⓘ v2.62.0 NOTES ↗ RANK

PydanticAI is a framework that builds type-safe Python agents with dependency injection, model integrations, tools, and structured outputs.

> PydanticAI v2.62.0 adds batch API support, finer-grained model error categories, mid-conversation thinking-effort control, and task control for SubAgents and the coding agent, alongside a string of breaking changes to compaction, MCP tool naming, graph replay ordering and realtime settings.

└──▷ WHAT SHIPPED · 13 FEATURES most completely described first ? what's the number?

01 **SubAgents task control for delegations** NEW 80

Adds a `task_control` parameter to `SubAgents` so the model can background, steer, answer, and cancel delegated sub-tasks, and exposes Coder(task_control=True) to enable this in the coding agent.

Enable background task delegation in the coding agent so the model can run long sub-tasks without blocking the main conversation.

python

```python
from pydantic_ai import Agent
from pydantic_ai_harness import Coder

agent = Agent(
    'anthropic:claude-fable-5-1',
    capabilities=[
        Coder(task_control=True),
    ],
)
agent.to_cli_sync()
```

— Named parameter across two surfaces with a runnable example v2.62.0

02 **Finer-grained model error categories** NEW 75

Adds `ModelRequestRejectedError`, `ModelAuthenticationError`, and `ModelNotFoundError` exception classes under `pydantic_ai.exceptions` so callers can catch specific model-layer failures — auth errors, missing models, or policy rejections — instead of parsing raw exception messages.

Catch specific model-layer failures — auth errors, missing models, or policy rejections — without parsing raw exception messages.

```python
from pydantic_ai.exceptions import ModelAuthenticationError,
ModelNotFoundError, ModelRequestRejectedError

try:
    result = await agent.run('Summarize this document.')
except ModelAuthenticationError:
    print('Invalid or missing API key — check your credentials.')
except ModelNotFoundError:
    print('The requested model does not exist on this provider.')
except ModelRequestRejectedError:
    print('Request was rejected by the model provider (policy or
quota).')
```

— Named exception classes plus a runnable try/except example v2.62.0

03 **Temporal child workflow tool calls** NEW 75

Tool calls can now run as Temporal child workflows via `child_workflow` tool metadata in `TemporalDurability`, letting individual tools survive worker restarts independently.

Run tool calls as durable Temporal child workflows so individual tools survive worker restarts independently.

```python
from pydantic_ai import Agent
from pydantic_ai.capabilities.durable_execution import
TemporalDurability

@agent.tool(metadata={'child_workflow': True})
async def run_analysis(ctx, dataset: str) -> str:
    ...

agent = Agent(
    'openai:gpt-6.1-sol',
    capabilities=[TemporalDurability()],
)
```

— Named metadata key with a runnable Temporal example v2.62.0

04 **MCP tool name prefixing** BREAKING 65

Adds `tool_prefix` to the `MCP` capability, and hosted MCP capabilities now prefix tool names with the service by default; callers referencing the old unprefixed names must update them or set `tool_prefix` explicitly.

— Names the flag and the exact migration action required v2.62.0

05 **Mid-conversation thinking-effort control** IMPROVED 60

Thinking effort can now be changed mid-conversation via OpenAI's `configuration_update` and Anthropic's per-message `output_config`; the change is recorded as a `ModelSettingsDeltaPart`.

— Names mechanisms but gives no code to invoke them v2.62.0

THINNER COVERAGE BELOW

06 **Batch API support via BatchModelRequests** NEW 55

Adds a `BatchModelRequests` capability and a `get_wrapper_model` capability hook so providers' batch APIs can be used.

— Names the new capability and hook but no usage example v2.62.0

07 **Allowed-tools restriction on OpenAI Chat Completions** IMPROVED 55

Restricts `tool_choice` subsets via `allowed_tools` on OpenAI Chat Completions to preserve the prompt cache.

— Names the flag and its purpose, no example given v2.62.0

08 **CodeMode Monty sandbox configuration** DEPRECATED 55

Adds CodeMode(monty_sandbox=...) accepting either a remote Monty URL or a pool-builder function, deprecating the previous `monty_sandbox_url` argument, which will be removed in a future release.

— Names old and new arguments and the deprecation path v2.62.0

09 **SummarizingCompaction stores CompactionPart** BREAKING 55

`SummarizingCompaction` now stores its summary as a `CompactionPart` instead of a system prompt; code that reads or expects a system-prompt summary will break.

— States before/after but no migration steps v2.62.0

10 **SystemReminders request-scoped prompts** DEPRECATED 50

`SystemReminders` are now built on request-scoped system prompts, and `cache_ttl` is deprecated.

— Names the deprecated field but no replacement guidance v2.62.0

11 **Realtime shared settings validation** BREAKING 50

Realtime shared settings now warn, raise, or ignore like standard `ModelSettings`: a forced `tool_choice` raises on Gemini Live, and `parallel_tool_calls=False` warns.

— Names concrete settings and their new behavior v2.62.0

---

12 **GraphBuilder broadcast branch ordering change** BREAKING

45

`GraphBuilder` broadcast branches now start in .to() order instead of reversed, which can break replay of durable workflows paused across the upgrade.

— Explains the behavior change but not how to remediate v2.62.0

---

13 **Durable execution conversation ID export** NEW

40

Exports `conversation_id_from_run_id` from `pydantic_ai.durable_exec` for use by third-party durability engines.

— Names the function but no usage detail v2.62.0

└──▷ **BREAKING ON UPGRADE**

! `SummarizingCompaction` now stores its summary as a `CompactionPart` instead of a system prompt; code that reads or expects a system-prompt summary will break.

! `GraphBuilder` broadcast branches now start in .to() order instead of reversed, which can break replay of durable workflows paused across the upgrade.

! Hosted MCP capabilities' tool names are now prefixed with the service name by default; callers that reference the old unprefixed names must update them or set `tool_prefix` explicitly.

! CodeMode(monty_sandbox_url=...) is deprecated in favour of CodeMode(monty_sandbox=...); the old argument will be removed in a future release.

! `SystemReminders` are now built on request-scoped system prompts; `cache_ttl` is deprecated.

! Realtime shared settings now warn, raise, or ignore like standard `ModelSettings`: forced `tool_choice` raises on Gemini Live, `parallel_tool_calls=False` warns.

> CVE-2026-107286 CVE ↗ STABLE
>
> `CVE-2026-107286`: Concurrency-limited model streams can retain slots after completion or interruption (CVSS 7.5, affects `>= 2.10.0, < 2.53.0`)

> CVE-2026-107287 CVE ↗ STABLE
>
> `CVE-2026-107287`: Excessive resource use when local web fetching converts nested HTML (CVSS 6.5, affects `>= 1.77.0, < 1.107.7`)

> CVE-2026-107291 CVE ↗ STABLE

`CVE-2026-107291` : OpenTelemetry instrumentation: `include_content=False` does not redact exceptions, error statuses, instructions or the output template (CVSS 4.3, affects `>= 0.3.4, < 1.107.6`)

`CVE-2026-107290` CVE ↗ STABLE

`CVE-2026-107290` : Event loop blocked by quadratic response processing in `web_fetch` (CVSS 6.5, affects `>= 1.77.0, < 1.107.6`)

`CVE-2026-107289` CVE ↗ STABLE

`CVE-2026-107289` : SSRF cloud-metadata blocklist bypass via IPv6 zone identifiers (CVSS 6.8, affects `>= 1.56.0, < 1.107.6`)

`CVE-2026-107292` CVE ↗ STABLE

`CVE-2026-107292` : Web chat UI (Agent.to_web(), `clai web`): the local chat endpoint does not validate the `Host` header, so DNS rebinding can reach it from a website the developer visits (CVSS 6.4, affects `>= 1.34.0, < 2.0.0b1`)

`CVE-2026-107294` CVE ↗ STABLE

`CVE-2026-107294` : Unbounded memory use when downloading remote content via web_fetch or FileUrl (CVSS 6.5, affects `>= 1.77.0, < 1.107.2`)

`CVE-2026-65975` CVE ↗ STABLE

`CVE-2026-65975` : UI adapters (AG-UI, Vercel AI): a dangling client-submitted tool call can execute when a trailing message is dropped during `sanitize_messages` (CVSS 6.5, affects `>= 1.88.0, < 1.107.1`)

`CVE-2026-54249` CVE ↗ STABLE

`CVE-2026-54249` : VercelAIAdapter trusts client-controlled `providerMetadata` to construct `UploadedFile` — S3/GCS confused deputy via provider metadata injection (CVSS 6.8, affects `>= 1.65.0, < 1.106.0`)

`CVE-2026-25580` CVE ↗ STABLE

`CVE-2026-25580` : Server-Side Request Forgery (SSRF) in URL Download Handling (CVSS 8.6, affects `>= 0.0.26`)

`CVE-2026-25640` CVE ↗ STABLE

> `CVE-2026-25640`: Stored XSS via Path Traversal in Web UI CDN URL (CVSS 7.1, affects `>= 1.34.0, < 1.51.0`)

WAS THIS USEFUL? 👍 👎

OpenClaw

Feedback ⓘ v2026.9.35 NOTES ↗ CODE ↗ RANK

OpenClaw is a personal AI assistant platform that connects messaging channels, tools, and model providers through a gateway.

> OpenClaw v2026.9.35 adds support for Moonshot's Kimi K3.1 Pro models and redacts credentials from MCP command output and logs.

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

01 **Kimi K3.1 Pro model support** NEW 65

Adds `moonshot/kimi-k3.1-pro`, `kimi/k3.1-pro`, and their Highspeed variants, with K3 thinking, effort, and video controls exposed via these model IDs.

— Names exact model IDs and controls but no usage example v2026.9.35

THINNER COVERAGE BELOW

02 **Credential redaction in MCP CLI output** IMPROVED 55

Redacts credentials in `openclaw mcp list` and `openclaw mcp show` output and logs.

— Names exact commands affected, minimal mechanism detail v2026.9.35

WAS THIS USEFUL? 👍 👎

browser-use

Feedback ⓘ 0.13.14 NOTES ↗ RANK

The browser-use Python library automates web browsers for AI agents.

> browser-use 0.13.14 adds OpenAI browser tools integration, a shared Bash helper for file and data operations alongside browser actions, and support for importing Browser Use Cloud SDK V4.

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

01 **Shared Bash helper for agent tasks** NEW 35

Introduces a shared Bash helper for processing data and writing files alongside browser actions, usable during agent workflows.

— Describes purpose but no command or API shown 0.13.14

02 **OpenAI browser tools integration** NEW 25

Adds integration with OpenAI browser tools, letting agents use OpenAI's browser tooling as part of browser-use workflows.

— Named but no mechanism or usage detail given 0.13.14

03 **Browser Use Cloud SDK V4 import support** NEW 25

Adds support for importing Browser Use Cloud SDK V4.

— Bare mention of support with no usage detail 0.13.14

> CVE-2025-47241 CVE ↗ STABLE
>
> `CVE-2025-47241` : `allowed_domains` can be bypassed by putting a decoy domain in http auth username portion of a URL (CVSS 9.3, affects `<= 0.1.44`)

WAS THIS USEFUL? 👍 👎

Vercel AI SDK

Feedback ⓘ ai@7.0.154 NOTES ↗ RANK

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

> Vercel AI SDK 7.0.154 adds fine-grained error and abort handling to agentic tool loops, plus version tagging in telemetry events for multi-SDK-version environments.

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

01 **Error and abort callbacks for agent loops** NEW 75

Adds `onStepError` callback to `generateText`, `streamText`, and `ToolLoopAgent`, plus `onError` and `onAbort` callbacks to `ToolLoopAgent`, enabling fine-grained handling of per-step errors, overall agent errors, and aborted runs without stopping the entire agentic loop.

Catch per-step errors in an agent loop without aborting the whole run — useful for logging or retrying individual tool failures.

typescript

```typescript
const agent = new ToolLoopAgent({
  model: 'openai/gpt-6-astra',
  tools: { /* ... */ },
  onStepError: (error) => {
    console.error('Step failed:', error);
  },
  onError: (error) => {
    console.error('Agent error:', error);
  },
  onAbort: () => {
    console.warn('Agent run was aborted');
  },
});
```

— Names exact callbacks with a runnable code example ai@7.0.154

02 **aiVersion field in telemetry events** NEW 65

Adds an `aiVersion` field to telemetry events so tracing-channel subscribers and telemetry integrations can identify which `ai` package version emitted the event, including in bundled apps and multi-SDK-version environments.

Track which AI SDK version emitted a telemetry event when running multiple SDK versions in the same bundled application.

typescript

```typescript
// aiVersion is now automatically included in all telemetry
events
// Subscribe via tracing-channel to inspect it:
import { tracingChannel } from 'diagnostics_channel';

tracingChannel('ai.generateText').subscribe({
  asyncStart(event) {
    console.log('Emitted by ai version:', event.aiVersion);
  }
});
```

— Names the field and shows subscription code `ai@7.0.154`

> CVE-2026-64650 CVE ↗ STABLE
>
> `CVE-2026-64650` : AI SDK Codex Harness Tool Relay Authorization Bypass (affects `<= 1.0.28`)

WAS THIS USEFUL? 👍 👎

Stacklok Mecatl

Feedback ⓘ v0.0.51 NOTES ↗ RANK

Mecatl is an open-source agent harness for running the same provider-agnostic agent loop locally, remotely, or on Kubernetes, with durable state, permissions, and gRPC/HTTP-SSE APIs.

> Mecatl v0.0.51 strengthens session identity by anchoring sessions to the authenticated `sub_id`, and refines subagent delegation to be cost-aware while keeping verification in the parent agent.

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

01 **Sessions tied to sign-in sub_id** IMPROVED 43

Sessions are now owned by the sign-in's `sub_id` identity, giving each authenticated session a stable, attributable identity anchor.

— Names the `sub_id` field but no API or migration detail given. v0.0.51

02 **Cost-based subagent delegation** IMPROVED 35

Subagent delegation is now framed by cost, keeping verification in the parent agent rather than the delegated subagent.

— Describes behavior change but no concrete mechanism or config named. v0.0.51

WAS THIS USEFUL? 👍 👎

LangChain LangGraph

Feedback ⓘ cli==0.4.40 NOTES ↗

RANK

Low-level orchestration framework for building, managing, and deploying long-running, stateful AI agents.

> LangGraph CLI shipped two small improvements this window: authenticated upload header forwarding during deployment and a machine-readable JSON output mode for listing deployments.

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

01 **`--json` flag for `deploy list`** NEW 75

Adds a `--json` flag to the `langgraph deploy list` command for machine-readable JSON output, e.g. `langgraph deploy list --json`, useful for piping into tools like `jq`.

Pipe deployment listings into `jq` or other tooling for scripted inspection of active deployments.

```
$ langgraph deploy list --json
```

— Exact flag and runnable command given cli==0.4.39

THINNER COVERAGE BELOW

02 **Control-plane upload headers in CLI deploy** IMPROVED 45

The LangGraph CLI now forwards upload headers returned by the control plane when pushing deployment artifacts, enabling authenticated or signed upload flows.

— Describes mechanism but no flag or config name given cli==0.4.40

> CVE-2026-115883 CVE ↗ STABLE
>
> `CVE-2026-115883`: Namespace labels containing `.` reach other namespaces in Postgres and SQLite stores (affects `< 3.1.3`)

> CVE-2026-115885 CVE ↗ STABLE
>
> `CVE-2026-115885`: LangGraph CLI prints the registry push token in `langgraph deploy --verbose` output (affects `>= 0.4.15, <= 0.4.31`)

CVE-2026-104873 CVE ↗ STABLE

`CVE-2026-104873`: LangGraph SDK custom auth silently ignores actions= on resource decorators (affects `>= 0.1.45, <= 0.4.3`)

CVE-2026-71433 CVE ↗ STABLE

`CVE-2026-71433`: Namespace prefix matching crosses segment boundaries in Postgres and SQLite stores (CVSS 5.3, affects `< 3.1.1`)

CVE-2026-48775 CVE ↗ STABLE

`CVE-2026-48775`: Unsafe JSON deserialization in LangGraph checkpoint loading (CVSS 6.8, affects `<= 4.1.0`)

CVE-2026-27794 CVE ↗ STABLE

`CVE-2026-27794`: ZDI-CAN-28385: LangChain LangGraph BaseCache Deserialization of Untrusted Data Remote Code Execution Vulnerability (CVSS 6.6, affects `<4.0.0`)

CVE-2025-67644 CVE ↗ STABLE

`CVE-2025-67644`: SQL injection via metadata filter key in SQLite checkpointer list method (CVSS 7.3, affects `<3.0.1`)

CVE-2025-64439 CVE ↗ STABLE

`CVE-2025-64439`: RCE in `json` mode of JsonPlusSerializer (affects `<3.0`)

WAS THIS USEFUL? 👍 👎

Feynman

Feedback ⓘ v0.5.58 NOTES ↗ RANK

Feynman is a personal AI research assistant that searches papers and the web, builds literature reviews, and produces cited research reports.

> Feynman v0.5.58 adds support for routing web searches through a self-hosted 9router gateway and improves source verification, subagent reliability, and prompt caching.

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

01 **Self-hosted 9router gateway support** NEW 83

Adds `feynman search set ninerouter <base-url>` to select a self-hosted 9router gateway for web searches, useful for avoiding public bot-detection or enforcing egress policy.

Route all Feynman web searches through a self-hosted 9router instance to avoid public bot-detection or to enforce egress policy.

```
$ feynman search set ninerouter
https://9router.internal.example.com
```

— Names exact command and purpose, runnable as given. v0.5.58

THINNER COVERAGE BELOW

02 **Prompt cache preserved across subagent turns** IMPROVED 51

Prompt cache is now preserved across turns after a subagent returns, on models that accept mid-conversation system messages, reducing cost and latency.

— Explains mechanism and benefit but nothing to configure. v0.5.58

03 **Matching passage display in source_check** IMPROVED 48

`source_check` now displays the matching passages under each source, making it easier to verify citations.

— Names the feature but no concrete UI path or command. v0.5.58

04 **Reliable reporting from background research subagents**

IMPROVED 45

Background research subagents that finish work now always report their result instead of exiting silently with code 0, fixing a reliability gap in asynchronous research tasks.

— Clear before/after but no user-facing action to take. v0.5.58

WAS THIS USEFUL? 👍 👎

openexecutive

Feedback ⓘ v0.6.6 NOTES ↗ RANK

AI-powered virtual executive team — a single coherent executive persona backed by 8 specialist agents (FastAPI + Next.js).

> OpenExecutive v0.6.6 tightened frontend security with a nonce-based CSP, added support for host-origin framing, and expanded the Home dashboard with an At a glance summary.

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

01 **Nonce-based CSP on every page** IMPROVED 35

OpenExecutive now adds a nonce Content-Security-Policy header on every UI page, tightening script execution controls across the application.

— Names the mechanism but no header details or rollout notes v0.6.6

02 **Host-origin framing support** NEW 30

The UI can now be framed by host origins, with support for calling a short list of allowed hosts.

— No config key or list mechanism specified v0.6.6

03 **Expanded Home dashboard** IMPROVED 30

The Home view now includes an At a glance summary section with additional tiles, expanding the dashboard overview.

— Describes the addition but not tile contents v0.6.6

WAS THIS USEFUL? 👍 👎

claw-orchestrator

Feedback ⓘ v7.15.0 NOTES ↗ RANK

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's v7.15.0 release adds ACP agent support across fan-out, council, workflows and handoff, tightens acceptance-contract enforcement so unchecked runs fail, and extends spend-cap and ledger enforcement to inbox delivery paths, alongside smaller changes to cost reporting, council defaults and plugin manifest behaviour.

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

01 ACP agent support across fan-out, council, workflows, handoff

NEW

95

Adds an `acpAgent` tool and `acp:<preset>` CLI syntax so ACP agents can be mixed with native engines, e.g. `clawo fanout "…" -e claude,acp:kimi-code,acp:dsh --synthesize`. ACP agent permission options now accept only `plan`, `acceptEdits`, or `bypassPermissions`; ACP sessions refuse `sandboxMode: 'read-only'` and receive system prompts in their first message.

Run a fan-out query distributing work across a mix of a native engine and two ACP preset agents, then synthesize the results.

```
$ clawo fanout "Summarise the threat model for this codebase" -e claude,acp:kimi-code,acp:dsh --synthesize
```

— Exact CLI syntax, tool name, and runnable example given v7.15.0

02 Spend cap and ledger enforcement on inbox delivery

BREAKING

75

Inbox deliveries via `session_send_to`, `session_deliver_inbox`, and `team_send` now go through the normal send path: the spend cap is applied, each delivery gets a ledger row, and a busy recipient queues the next message instead of being delivered synchronously.

— Names all three affected endpoints and the new enforcement mechanism v7.15.0

THINNER COVERAGE BELOW

03 Stricter acceptance-contract enforcement

BREAKING

50

A run whose router bypassed its verifier now ends `failed` instead of `completed`, and a contract with no recognised check is refused at start rather than accepted as a no-op.

— Behaviour change described but no command to invoke it v7.15.0

---

04 **Council round defaults and abort return value** BREAKING

50

Council workflow nodes now default to 15 rounds, matching `council_start`, and `council_abort` returns `ok: false` when nothing was running in the current process.

— Names both affected calls but no example of handling the new return v7.15.0

---

05 **Plugin manifest defaultModel scoping and proxy default**

50

BREAKING

The plugin `defaultModel` setting now applies only to an engine that actually runs that declared model, leaving other engines unaffected; the plugin manifest also now declares the multi-model proxy on by default.

— Config key named but no file path or migration step given v7.15.0

---

06 **Fuller cost and token reporting** IMPROVED

45

`session_send` now returns the turn's own token count and cost, and a run's reported cost now includes tokens consumed by its fan-out and council agents.

— Names the API but no usage example v7.15.0

---

07 **Port-conflict handling for serve command** IMPROVED

35

`clawo serve` now exits with code 1 when its port is already in use.

— Exact command and exit code named, but minimal mechanism v7.15.0

└──▷ **BREAKING ON UPGRADE**

! The plugin `defaultModel` now applies only to an engine that actually runs that model — engines not running the declared model are no longer affected.

! ACP sessions refuse `sandboxMode: 'read-only'`; any workflow that passed that option to an ACP agent will fail.

! A run whose router steered past its verifier now ends `failed` instead of `completed` — pipelines that expected `completed` in that scenario will see a different terminal state.

! A contract with no recognised check is now refused at start — previously accepted no-op contracts will be rejected.

! Inbox deliveries via `session_send_to`, `session_deliver_inbox`, and `team_send` now enforce the spend cap and queue on a busy recipient — previously uncapped or synchronous delivery behaviour will change.

! The plugin manifest now declares the multi-model proxy on by default.

! `council_abort` returns `ok: false` when nothing was running in the current process — callers that assumed a truthy return in that case will need to handle the new value.

WAS THIS USEFUL? 👍 👎

llphant

Feedback ⓘ 1.0.8 NOTES ↗ RANK

LLPhant - A comprehensive PHP Generative AI Framework using OpenAI GPT 4. Inspired by Langchain

> LLPhant 1.0.8 introduces a new OpenAI-compatible provider config, adds document metadata support, and changes the default OpenAI model, which is a breaking change for existing integrations.

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

01 **Default OpenAIConfig model changed to gpt-5.6-sol** BREAKING 65

The default model in `OpenAIConfig` changes from `gpt-4-turbo` to `gpt-5.6-sol`; any code relying on the previous default will now target a different model.

— Names exact config key and old/new default values 1.0.8

THINNER COVERAGE BELOW

02 **GisooConfig for Gisoo API support** NEW 35

LLPhant adds `GisooConfig` to support the OpenAI-compatible Gisoo API.

— Names config class and API but no usage mechanism 1.0.8

03 **Document metadata support** NEW 23

LLPhant adds document metadata support via the Add document metadata change.

— No fields, API or mechanism specified 1.0.8

└──▷ **BREAKING ON UPGRADE**

! The default model in `OpenAIConfig` changes from `gpt-4-turbo` to `gpt-5.6-sol`; any code relying on the previous default will now target a different model.

WAS THIS USEFUL? 👍 👎

anton

Feedback ⓘ v2.26.10.26.1 NOTES ↗ RANK

self-improving agent

> Anton's single release this window adds an MCP client, a Landlock-based filesystem sandbox, a read-only Cortex memory layer, interactive human-in-the-loop prompts with a write-approval gate, and a generate_artifact web pipeline, alongside credential-scrubbing hardening and a breaking tool rename from open_artifact to read_text_file.

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

01 `read_text_file` tool replaces `open_artifact`

BREAKING

75

Adds the `read_text_file` tool, replacing the former `open_artifact` tool. Any agent code or host integration calling `open_artifact` must be updated to use `read_text_file`.

– Names exact tools and states migration requirement v2.26.10.26.1

02 Human-in-the-loop prompts and write-approval gate in cloud turns

NEW

65

Adds `interactive ask_user` support in cloud turns for mid-turn human-in-the-loop prompts, and a write-approval gate that holds any scratchpad cell that would mutate connected data or existing files until the user confirms.

– Explains mechanism and scope of gate but no API detail v2.26.10.26.1

03 generate_artifact web artifact pipeline NEW

65

Adds `generate_artifact`, a one-tool web artifact pipeline covering discovery, generation, and publish-or-preview.

Opt out of web search and fetch globally for a workspace where outbound HTTP is restricted.

```ini
ANTON_WEB_SEARCH_ENABLED=false
ANTON_WEB_FETCH_ENABLED=false
```

– Names tool and pipeline stages but no full usage walkthrough v2.26.10.26.1

04 Credential security hardening in scratchpad and chat IMPROVED

60

Scrubs OAuth tokens in scratchpad output and Google AQ authorization keys at chat input before they reach the model or storage, and passes Google Ads developer tokens through the cloud data vault instead of inline credentials.

— Names specific credential types and vault but no config steps v2.26.10.26.1

THINNER COVERAGE BELOW

05 **Per-call LLM usage logging via ANTON_USAGE_LOG** NEW

55

Adds the `ANTON_USAGE_LOG` environment variable to emit per-call LLM usage for evals.

— Names env var but no output format shown v2.26.10.26.1

06 **Landlock filesystem sandbox isolation** NEW 50

Adds a Landlock-based filesystem sandbox to lock agent code to its own folders during cell execution.

— Explains mechanism but no config to enable or tune it v2.26.10.26.1

07 **MCP client support (Stage 1)** NEW 40

Adds an MCP client (Stage 1) to Anton, enabling it to connect to and consume MCP servers.

— Describes capability but no connection setup detail v2.26.10.26.1

08 **Read-only memory layer in Cortex** NEW 40

Adds an optional read-only memory layer to Cortex for cross-session knowledge access.

— States purpose but no activation or scope details v2.26.10.26.1

09 **Pre-agent validation of decks and Word files** NEW 25

Validates decks and Word files before they are handed over to the agent.

— Minimal detail on validation rules or failure handling v2.26.10.26.1

└──▷ **BREAKING ON UPGRADE**

! The `open_artifact` tool is replaced by `read_text_file`; any agent code or host integration calling `open_artifact` must be updated.

WAS THIS USEFUL? 👍 👎

Graphicle AI Graphicle

Feedback ⓘ v0.1.8 NOTES ↗

RANK

Code graph and memory for your coding agents, in one Rust program. For Claude Code, Codex, Cursor and any MCP client.

> Graphicle v0.1.8 introduces a per-episode query file that lets code-graph queries decode only the portions they need, cutting median `graphicle context` latency from ~3s to ~1.3s and peak memory from 3.5GB to 0.6GB on a 30,315-file Kubernetes clone, alongside integrity checks, retry logic, per-branch graph reuse, and a wider key space covering dependencies.

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

01 **Query file for selective code-graph decoding** NEW 90

Adds a per-episode query file (`gf-cache/<project>/query/<episode>.query`) written alongside the saved graph so that `graphicle context` , the prompt hook, and `recall_project` decode only the portions of the code graph each question actually needs, instead of loading the whole graph. Setting `GRAPHICLE_QUERY_FILE=0` forces every question to read the full saved graph instead. This reduces median `graphicle context` latency from ~3 seconds to ~1.3 seconds (16 cores) and peak memory from 3.5 GB to 0.6 GB on a 30,315-file shallow clone of Kubernetes.

— Names exact file path, env var, and before/after benchmarks v0.1.8

02 **Retry logic for query-file write failures** NEW 75

Adds retry logic for query-file write failures starting at 10 minutes, doubling after each consecutive failure, capped at 6 hours, with an immediate retry triggered by `graphicle sync` , `graphicle setup` , and `graphicle doctor --fix` .

— Exact backoff schedule and trigger commands named v0.1.8

03 **Post-upgrade sync required for query-file benefits** BREAKING 75

Until `graphicle sync` is run once in each project after upgrading, questions read the full code graph from the store (median 22.5 seconds per question on a 30,315-file project) and the prompt hook gives only the files that define code names a prompt provides on projects of 2,000 files or more.

— Concrete migration step with numbers and thresholds v0.1.8

04 **Checksums for query file and saved graph integrity** NEW 67

Adds checksums over the query file's header and each page, and appends length-plus-checksum to the saved graph and file-relationships file so that `graphicle sync` detects truncation or damage and rewrites the affected file on the next sync.

— Mechanism specified but no numbers or thresholds v0.1.8

THINNER COVERAGE BELOW

05 **Per-branch saved graph and query file reuse** IMPROVED

58

Retains per-branch saved files so that switching back to a previously synced branch (`graphicle sync`) reuses that branch's saved graph and query file instead of re-reading the full code graph from the store.

— Clear mechanism but no numbers or limits given v0.1.8

06 **Query-file status reporting in graphicle doctor** IMPROVED

50

Reports query-file and saved-graph status, including pending retries and graphs the file cannot hold, under the Questions section of `graphicle doctor`.

— Names exact command and section but little mechanism v0.1.8

07 **Extended key space to cover program sources and dependencies** IMPROVED

44

Extends the saved-graph and query-file key space to cover the program's sources and its locked dependencies (not only the engine's), preventing one build's files from being read by another.

— Describes the fix but lacks concrete surfaces or numbers v0.1.8

└──▷ **BREAKING ON UPGRADE**

! Until `graphicle sync` is run once in each project after upgrading, questions read the full code graph from the store (median 22.5 seconds per question on a 30,315-file project) and the prompt hook gives only the files that define code names a prompt provides on projects of 2,000 files or more.

WAS THIS USEFUL?  

yoagent

Feedback ⓘ v0.25.15 NOTES ↗ RANK

The agent loop for Rust — stream from 7 LLM protocols, run tools, loop until done. From a laptop to a Cloudflare Worker.

> yoagent shipped two MCP-focused capabilities: detecting when a stdio MCP server process has exited, and mid-run tool-source swapping with live updates when an MCP server's tool list changes.

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

01 MCP stdio server exit detection NEW

93

`McpClient::exited() -> Option<String>` detects when a stdio MCP server has gone, returning `None` while running or a description of the exit status, signal, and last stderr once the process exits. Custom transports can participate in exit detection via the McpTransport::exit_watch() provided method and mcp::ExitWatch::new(check); `McpToolSource` now withdraws a server's tools at the next turn when the server process exits.

Poll for a crashed MCP stdio server between turns so you can surface an error or restart it externally.

rust

```rust
if let Some(reason) = mcp_client.exited() {
    eprintln!("MCP server terminated: {}", reason);
    // restart or notify user here
}
```

Implement exit detection in a custom MCP transport by supplying a check closure to `mcp::ExitWatch::new`.

rust

```rust
use yoagent::mcp;

let exit_watch = mcp::ExitWatch::new(|| {
    // return Some(reason) if the backing process has gone, None
if still alive
    check_my_process()
});
// return this from McpTransport::exit_watch() in your impl
```

— Exact function signatures plus runnable code examples for both call and implementation sides.

v0.25.15

02 Mid-run tool-source swapping with live MCP updates NEW

93

`ToolSource::generation() -> u64` enables mid-run tool-source changes: when a source's generation advances, its updated tool list is fetched before the next model request. The new `mcp::McpToolSource` lists an MCP server's tools at run start and updates them live when the server emits `notifications/tools/list_changed`, using McpClient::tools_list_changes() and McpClient::reports_notifications() to query notification state. Custom transports opt in to forwarding these notifications via `report_notifications(ServerNotifications) -> bool` on the `McpTransport` trait.

Opt a custom MCP transport in to forwarding `notifications/tools/list_changed` so `McpToolSource` picks up live updates through it.

rust

```rust
use yoagent::mcp::{McpTransport, ServerNotifications};

impl McpTransport for MyTransport {
    fn report_notifications(&self, notif: ServerNotifications) ->
bool {
        // return true to capture tool-list-changed notifications
        true
    }
    // ... other required methods
}
```

— Named trait methods and a concrete transport implementation example. v0.25.14

WAS THIS USEFUL? 👍 👎

sympozium

Feedback ⓘ v0.12.0 NOTES ↗ RANK

The Coordination Layer for Multi-Agent AI

> Sympozium v0.12.0 decouples model access from the Celln fleet, giving each Agent its own provider key and a direct Agent API, while also speeding up Celln admission and turn handling.

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

01 Agent-bound model access with per-Agent provider keys

87

BREAKING

The Celln fleet is now model-free: each Agent carries its own provider key and model configuration rather than relying on a shared gateway key, and a direct Agent API is exposed for interacting with Agents individually. This is a breaking change — existing setups that relied on a shared gateway key for all Agents beyond the initial `starter` Agent must configure per-Agent keys; an Agent manifest now sets `spec.execution.backend: celln`, `spec.execution.provider`, `spec.execution.model`, and `spec.agents.default.model` directly.

Bootstrap a Celln-backed Agent with its own per-Agent provider key and execution defaults, so it operates independently without a shared gateway key.

```yaml
apiVersion: sympozium.ai/v1alpha1
kind: Agent
metadata:
  name: my-agent
  namespace: default
spec:
  runtimeRef: json-harness
  execution:
    backend: celln
    executionLifecycle: enduring
    provider: deepseek
    model: deepseek-chat
    enduring:
      leaseSeconds: 600
      maxTurns: 8
      maxModelRequests: 24
      maxOutputTokens: 8192
    cellnSelection:
      toolRefs:
        - name: workspace-write
          revision: v1
  agents:
    default:
      model: deepseek-chat
```

— Names exact config fields and includes a runnable Agent manifest v0.12.0

THINNER COVERAGE BELOW

02 **Faster Celln admission and turn processing** IMPROVED

27

Accelerates Celln admission and turn processing, and now refuses stale turns immediately.

— No mechanism, numbers, or surface given, just a brief claim v0.12.0

└──▷ **BREAKING ON UPGRADE**

! Agent-bound model access changes how model credentials are distributed: the Celln fleet no longer holds provider keys centrally — each Agent must supply its own key. Existing setups that relied on a shared gateway key for all Agents (beyond the `starter` Agent) will need per-Agent key configuration.

WAS THIS USEFUL? 👍 👎

AWS Strands Agents

Feedback ⓘ CVE-2026-115419 CVE ↗

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.

`CVE-2026-115419`: Exposure of data element to wrong session in Strands Agents SDK A2A Server (CVSS 8.1, affects `< 1.43.0`)

WAS THIS USEFUL?  

Mem0

Feedback ⓘ CVE-2026-116575 CVE ↗

Mem0 is a memory layer for AI agents that extracts, stores, and retrieves durable facts across sessions through a hosted or self-hosted store.

`CVE-2026-116575`: Cross-tenant conversation leak via unescaped delimiter collision in session-scope key (_build_session_scope) (CVSS 6.3, affects `≤ 2.0.0`)

WAS THIS USEFUL?  

ag2

Feedback ⓘ CVE-2026-113663 CVE ↗

AG2 (formerly AutoGen): The Open-Source AgentOS.Join us at: https://discord.gg/sNGSwQME3x

`CVE-2026-113663`: Newline injection in SandboxShellTool bypasses allowed=[...] and `readonly=True`, enabling arbitrary command execution (CWE-78 / CWE-184) (CVSS 8.8, affects `>= 0.13.2, < 1.1.1`)

WAS THIS USEFUL?  

Daily Pipecat

Feedback ⓘ CVE-2026-102682 CVE ↗ STABLE

Pipecat is a framework for realtime voice and multimodal agents, orchestrating speech-to-text, model inference, and text-to-speech as a streaming pipeline.

`CVE-2026-102682` : Public exposure of Pipecat's development runner can allow unauthenticated bot session creation (CVSS 5.6, affects `>= 0.0.98, <= 1.5.0`)

▾ **3 more releases in this issue** · 2026-04-23 → 2026-09-29

> CVE-2026-54695 CVE ↗ STABLE
>
> `CVE-2026-54695` : Telephony WebSocket `/ws` Unauthenticated Call-Control Abuse via Attacker-Supplied Call SID (CVSS 7.5, affects `<= 0.0.77,<1.4.0`)

> CVE-2026-44716 CVE ↗ STABLE
>
> .
> .
> .
> .
>
> `CVE-2026-44716` : Path Traversal in Pipecat Runner `/files` Endpoint — Arbitrary File Read via `%2F` `-Encoded` Separator (CVSS 7.5, affects `>=0.0.90,<1.2.0`)

> CVE-2025-62373 CVE ↗ STABLE
>
> `CVE-2025-62373` : Remote Code Execution by Pickle Deserialization via LivekitFrameSerializer in pipecat-ai/pipecat (CVSS 9.8, affects `>= 0.0.41,<0.0.94`)

WAS THIS USEFUL? 👍 👎

Langflow Feedback ⓘ CVE-2026-105697 CVE ↗ STABLE

Langflow is a visual builder for agents and RAG workflows that exports each flow as an API or Python object for deployment.

`CVE-2026-105697` : Langflow < 1.10.3: OS command injection (RCE) via arbitrary command in MCP stdio server configuration (CVSS 9.9, affects `>= 1.1.2, < 1.10.3`)

▾ **16 more releases in this issue** · 2025-12-19 → 2026-09-28

> CVE-2026-105698 CVE ↗ STABLE
>
> `CVE-2026-105698` : Langflow < 1.10.1: Cross-user flow access and vertex execution via deprecated `/api` /v1/build/{flow_id} `/vertices` endpoints (CVSS 5.4, affects `>= 1.0.0, < 1.10.1`)

> CVE-2026-105699 CVE ↗ STABLE
>
> `CVE-2026-105699` : Authenticated Cross-Project File Disclosure via Unscoped MCP Resource Handlers (affects `>= 1.6.8, <= 1.9.0`)

> CVE-2026-105740 CVE ↗ STABLE
>
> ⋮
>
> `CVE-2026-105740` : Authenticated RCE via MCP Stdio transport allows any user to execute arbitrary OS commands on the server (CVSS 9.9, affects `1.8.3`)

> CVE-2026-105741 CVE ↗ STABLE
>
> `CVE-2026-105741` : IP Spoofing Bypass via `X-Forwarded-For` Allowing Remote Configuration Write (CVSS 7.1, affects `>= 1.5.0, < 1.10.3`)

> CVE-2026-51886 CVE ↗ STABLE
>
> `CVE-2026-51886` : Title Authenticated Remote Code Execution in validate_code via Malicious Decorators Description (CVSS 8.8, affects `>= 1.7.2, < 1.10.1`)

> CVE-2026-8446 CVE ↗ STABLE
>
> ⋮
>
> `CVE-2026-8446` : Authentication Bypass in MCP Composer OAuth Branch (CVSS 8.6, affects `1.8.0.rc1`)

CVE-2026-55447 CVE ↗ STABLE

`CVE-2026-55447`: BaseFileComponent-based nodes arbitrary file read with RCE exploit (CVSS 9.6, affects `< 1.9.2`)

CVE-2026-55255 CVE ↗ STABLE

`CVE-2026-55255`: IDOR Vulnerability in `/api/v1/responses` Endpoint Allows Authenticated Attackers to Access Another User's Flow (CVSS 9.9, affects `< 1.9.1`)

CVE-2026-55423 CVE ↗ STABLE

`CVE-2026-55423`: Logout button does not clear session (CVSS 6.1, affects `< 1.7.0`)

CVE-2026-55450 CVE ↗ STABLE

`CVE-2026-55450`: Unauthenticated file upload leads to DoS (space exhaustion) and information leak (CVSS 9.3, affects `< 1.9.1`)

CVE-2026-33760 CVE ↗ STABLE

`CVE-2026-33760`: IDOR/BOLA in Monitor API — Missing Ownership Enforcement on 7 Endpoints (CVSS 8.8, affects `<= 1.7.3`)

CVE-2026-42867 CVE ↗ STABLE

`CVE-2026-42867`: Path Traversal in Knowledge Bases API via Creation Endpoint (CVSS 6.5, affects `<= 1.8.4`)

CVE-2026-48520 CVE ↗ STABLE

`CVE-2026-48520`: Unauthenticated Shareable Playground arbitrary local or S3 file read (CVSS 6.1, affects `<1.10.0`)

CVE-2026-33017 CVE ↗ STABLE

`CVE-2026-33017`: Unauthenticated Remote Code Execution in Langflow via Public Flow Build Endpoint (CVSS 9.8, affects `<= 1.8.2`)

CVE-2025-68478 CVE ↗ STABLE

`CVE-2025-68478`: External Control of File Name or Path in Langflow (CVSS 7.1, affects `< 1.7.0`)

CVE-2025-68477 CVE ↗ STABLE

CVE-2025-68477 : SSRF in langflow-ai/langflow (CVSS 7.7, affects < 1.7.0)

WAS THIS USEFUL? 👍 👎

Anthropic Claude Agent SDK Python Feedback ⓘ

CVE-2026-96620 CVE ↗

The Claude Agent SDK for Python is a developer library for building Python applications that run Claude-powered agents with tools, file access, and autonomous multi-step task execution.

CVE-2026-96620 : Argument Injection via resume Option Allows Arbitrary Command Execution (affects < 0.2.121)

WAS THIS USEFUL? 👍 👎

yuxi Feedback ⓘ CVE-2026-97693 CVE ↗ STABLE

可私有部署的多租户知识智能体平台：统一 RAG、知识图谱、多智能体、MCP/Skills、沙盒与权限管理。Yuxi = Cloud Agents + Knowledge RAG, Self-hosted knowledge agent platform for RAG, knowledge graphs and multi-agent workflows.

CVE-2026-97693 : Yuxi sandbox isolation failure allows cross-tenant access and internal service compromise (CVSS 9.9, affects >= 0.6.0.beta1, < 0.7.1)

▾ 1 more release in this issue · 2026-05-25 → 2026-08-24

CVE-2026-50561 CVE ↗ STABLE

CVE-2026-50561 : JWT 身份认证绕过导致跨实例管理员令牌复用 (CVSS 9.4, affects < 0.6.2)

WAS THIS USEFUL? 👍 👎

LangGenius Dify Feedback ⓘ CVE-2026-105761 CVE ↗ STABLE

Dify is an LLM application platform for building agentic workflows and RAG pipelines through a visual workspace, deployable to cloud or self-hosted.

`CVE-2026-105761` : IDOR in AppMCPServer PUT Endpoint Allows Modification of Other Apps' MCP Servers (CVSS 7.1, affects `1.14.2`)

▾ **8 more releases in this issue** · 2025-04-18 → 2026-08-19

> CVE-2026-105762 CVE ↗ STABLE
> .
> .
> .
>
> `CVE-2026-105762` : Unauthenticated Server-Side Request Forgery in `/console` /api/remote-files/upload endpoint (CVSS 8.3, affects `< 1.13.0`)

> CVE-2026-34082 CVE ↗ STABLE
> .
> .
>
> `CVE-2026-34082` : IDOR in deleting someone else's chat conversation (CVSS 4.3, affects `<= 1.9.2`)

> CVE-2026-42138 CVE ↗ STABLE
> .
>
> `CVE-2026-42138` : Stored XSS via SVG-file upload (CVSS 6.1, affects `<= 1.9.2`)

> CVE-2026-28288 CVE ↗ STABLE
> .
> .
>
> `CVE-2026-28288` : User enumeration (CVSS 5.3, affects `<= 1.8.1`)

> CVE-2025-67732 CVE ↗ STABLE
>
> `CVE-2025-67732` : Plaintext API Key Exposure via Model Provider Configuration Endpoint (CVSS 6.5, affects `<= 1.10.1-fix.1`)

> CVE-2025-58747 CVE ↗ STABLE
> .
> .
>
> `CVE-2025-58747` : Dify MCP OAuth Flow Vulnerable to XSS (CVSS 6.1, affects `<=1.8.0`)

> CVE-2025-43862 CVE ↗ STABLE
> .

CVE-2025-43862 : Unauthorized Access and Modification of APP Orchestration (CVSS 7.6, affects <= v0.6.8)

CVE-2025-32795 CVE ↗ STABLE

CVE-2025-32795 : Insecure User Role Access Control for APP Editing (CVSS 6.5, affects <= v0.6.8)

WAS THIS USEFUL? 👍 👎

AutoGPT

Feedback ⓘ CVE-2026-72922 CVE ↗ STABLE

AutoGPT automates multi-step workflows with continuous AI agents that users can build and deploy.

`CVE-2026-72922` : Webhook provider path confusion bypasses generic webhook secret verification (CVSS 8.2, affects `< 0.6.70`)

▾ **12 more releases in this issue** · 2026-02-04 → 2026-08-05

> CVE-2026-56823 CVE ↗ STABLE
>
> `CVE-2026-56823` : IDOR in Webhook Ping Endpoint Allows Enumeration and Cross-User Ping Triggering (CVSS 5.4, affects `< 0.6.64`)

> CVE-2026-56663 CVE ↗ STABLE
>
> `CVE-2026-56663` : SSRF in `SendWebRequestBlock` via IPv4-mapped IPv6 / CGNAT IP-validation bypass (CVSS 8.5, affects `<= 0.6.51`)

> CVE-2026-33235 CVE ↗ STABLE
> .
> .
> .
>
> `CVE-2026-33235` : Denial of Service (DoS) via Resource Exhaustion in text templating features (CVSS 7.7, affects `>= 0.1.0, < 0.6.52`)

> CVE-2026-55237 CVE ↗ STABLE
>
> `CVE-2026-55237` : DOM-Based XSS and Open Redirect in AutoGPT SignUp Page (CVSS 8.8, affects `< 0.6.62`)

> CVE-2026-33234 CVE ↗ STABLE
>
> `CVE-2026-33234` : SendEmailBlock allows SSRF via user-controlled SMTP server, bypassing IP blocklist (CVSS 5.0, affects `>= 0.1.0`)

> CVE-2026-33232 CVE ↗ STABLE
>
> `CVE-2026-33232` : Unauthenticated DoS via Disk Space Exhaustion (CVSS 7.5, affects `>= 0.4.2`)

> CVE-2026-33233 CVE ↗ STABLE

`CVE-2026-33233` : Unsafe Redis cache deserialization (`pickle.loads`) would allow backend RCE in case of cache poisoning (CVSS 7.6, affects `>= 0.6.34`)

`CVE-2026-30950` CVE ↗ STABLE

`CVE-2026-30950` : Authenticated Session Hijacking via IDOR (CVSS 7.1, affects `>= 0.6.36`)

`CVE-2026-45023` CVE ↗ STABLE

`CVE-2026-45023` : Credit system bypassed via direct block execution in POST `/api` /blocks/{block_id} `/execute` (CVSS 5.4, affects `<= 0.6.58`)

`CVE-2026-26020` CVE ↗ STABLE

`CVE-2026-26020` : Remote Code Execution via Disabled Block Bypass (CVSS 8.8, affects `< 0.6.47`)

`CVE-2025-62615` CVE ↗ STABLE

`CVE-2025-62615` : There is an SSRF vulnerability in ReadRSSFeedBlock (CVSS 9.8, affects `Commit 57a06f7 <= 0.6.33 <=`)

`CVE-2026-22038` CVE ↗ STABLE

`CVE-2026-22038` : API Keys and Secrets Logged in Plaintext in Stagehand Integration Blocks (CVSS 8.1, affects `<= l0.6.45`)

WAS THIS USEFUL? 👍 👎

Agno (formerly Phidata) Feedback ⓘ CVE-2026-10105

CVE ↗ STABLE

Agno is an agent framework and runtime that orchestrates and runs multi-agent systems.

`CVE-2026-10105`: Improper Neutralization of Special Elements used in an SQL Command (SQL Injection) in agno (CVSS 8.3, affects `< 2.8.5`)

▾ **2 more releases in this issue** · 2025-10-31 → 2026-07-27

> CVE-2026-35002 CVE ↗ STABLE
>
> `CVE-2026-35002`: Arbitrary code execution via eval() on field_type in FunctionCall (CWE-95) (CVSS 9.8, affects `< 2.3.24`)

> CVE-2025-64168 CVE ↗ STABLE
>
> `CVE-2025-64168`: Session state overwrites between different sessions/users (CVSS 7.1, affects `>2, <2.2.2`)

WAS THIS USEFUL? 👍 👎

holmesgpt Feedback ⓘ CVE-2026-63473 CVE ↗

SRE Agent - CNCF Sandbox Project

`CVE-2026-63473`: Unauthenticated remote command execution in `/api` /chat via Host header auth bypass and forged tool approval (CVSS 9.8, affects `>= 0.27.0, < 0.32.0`)

WAS THIS USEFUL? 👍 👎

trigger.dev

Feedback ⓘ CVE-2026-73654 CVE ↗ STABLE

Trigger.dev – build and deploy durable AI agents and workflows

`CVE-2026-73654`: Prototype pollution via run metadata operations → process-wide cross-tenant DoS (CWE-1321) (CVSS 8.5, affects `>= v3.3.8, <= 4.5.5`)

▾ **2 more releases in this issue** · 2026-07-09 → 2026-07-21

> CVE-2026-73656 CVE ↗ STABLE
>
> `CVE-2026-73656`: Cross-project deployment worker registration can modify another project's deployment state (CVSS 9.9, affects `<= 4.5.0-rc.2`)

> CVE-2026-73659 CVE ↗ STABLE
>
> `CVE-2026-73659`: Cross-tenant object read/write via path traversal in packet presign API (CVSS 8.1, affects `>= 4.4.2, <= 4.4.6`)

WAS THIS USEFUL? 👍 👎

taos

Feedback ⓘ CVE-2026-67548 CVE ↗ STABLE

Self-hosted AI agent OS. Your memory, chat, agents, and files stay on hardware you own, offline by default, cloud by choice.

`CVE-2026-67548`: Non-Admin Users Can Access System Settings and Update Actions (CVSS 8.3, affects `<= 1.0.0-beta.24`)

▾ **1 more release in this issue** · 2026-07-04

> CVE-2026-67547 CVE ↗ STABLE
>
> `CVE-2026-67547`: Authenticated Non-Admin Host Code Execution via Skill Runtime (CVSS 8.8, affects `<= 1.0.0-beta.22`)

WAS THIS USEFUL? 👍 👎

Databricks Omnigent

Feedback ⓘ CVE-2026-62677 CVE ↗

STABLE

Open-source meta-harness that orchestrates Claude Code, Codex, Cursor, Pi and custom agents behind one interface, with policy controls, OS-level sandboxing and live session sharing.

`CVE-2026-62677`: Unvalidated os_env.cwd in agent bundle yields arbitrary host filesystem access on runners without `OMNIGENT_RUNNER_WORKSPACE` (CVSS 8.8, affects `< 0.3.0`)

▾ **2 more releases in this issue** · 2026-06-29

> CVE-2026-62675 CVE ↗ STABLE
>
> `CVE-2026-62675`: Uploaded Agent Bundle Allows Authenticated Runner RCE via Python Callable Tools (CVSS 8.8, affects `< 0.3.0`)

> CVE-2026-62676 CVE ↗ STABLE
>
> `CVE-2026-62676`: Guardrail policy bypass: shell-command parser fails open (CWE-184) in policies/builtins/_shell.py (CVSS 7.1, affects `< 0.3.0`)

WAS THIS USEFUL? 👍 👎

langchain4j

Feedback ⓘ CVE-2026-55405 CVE ↗

LangChain4j is an idiomatic, open-source Java library for building LLM-powered applications on the JVM. It offers a unified API over popular LLM providers and vector stores, and makes implementing tool calling (including MCP support), agents and RAG easy.

`CVE-2026-55405`: SQL injection via metadata filters in langchain4j-mariadb and langchain4j-pgvector (CVSS 7.6, affects `<= 1.2.0-beta8`)

WAS THIS USEFUL? 👍 👎

LangChain

Feedback ⓘ CVE-2026-55443 CVE ↗ STABLE

LangChain is an open-source framework that orchestrates applications powered by language models.

`CVE-2026-55443`: Path traversal and sandbox escape in LangChain file-search middleware and loaders (CVSS 5.1, affects `<= 1.3.8`)

▾ **6 more releases in this issue** · 2025-11-19 → 2026-06-12

> CVE-2026-44843 CVE ↗ STABLE
>
> `CVE-2026-44843`: Unsafe deserialization of attacker-controlled LangChain objects through overly broad load() allowlists (CVSS 8.2, affects `<=1.3.2,>=1`)

> CVE-2026-41481 CVE ↗ STABLE
>
> `CVE-2026-41481`: `HTMLHeaderTextSplitter.split`_text_from_url SSRF Redirect Bypass (CVSS 6.5, affects `<1.1.2`)

> CVE-2026-40087 CVE ↗ STABLE
>
> `CVE-2026-40087`: Incomplete f-string validation in prompt templates (CVSS 5.3, affects `<0.3.83`)

> CVE-2026-34070 CVE ↗ STABLE
>
> `CVE-2026-34070`: Path traversal in legacy `load_prompt` functions in `langchain-core` (CWE-22) (CVSS 7.5, affects `<1.2.22`)

> CVE-2025-68664 CVE ↗ STABLE
>
> `CVE-2025-68664`: LangChain serialization injection vulnerability enables secret extraction in dumps/loads APIs (CVSS 9.3, affects `< 1.2.5, >= 1.0.0`)

> CVE-2025-65106 CVE ↗ STABLE
>
> `CVE-2025-65106`: Template Injection via Attribute Access in Prompt Templates (affects `<1.0.6,>=1.0.0`)

WAS THIS USEFUL? 👍 👎

LibreChat

Feedback ⓘ CVE-2026-54040 CVE ↗ STABLE

LibreChat provides a self-hosted chat interface with multiple model providers, agents, tools, and retrieval over uploaded files.

`CVE-2026-54040` : 2FA Backup Code Regeneration Without OTP Verification Allows 2FA Bypass (CVSS 5.9, affects `<= 0.8.3`)

▾ **22 more releases in this issue** · 2025-08-04 → 2026-06-02

> CVE-2026-54030 CVE ↗ STABLE
>
> `CVE-2026-54030` : Missing Resource Parameter Validation in MCP OAuth Flow (CVSS 8.0, affects `<= v0.8.5-rc1`)

> CVE-2026-54033 CVE ↗ STABLE
>
> `CVE-2026-54033` : SSRF via User-Provided Custom Endpoint baseURL — no private IP validation on user-configured API base URLs (CVSS 7.7, affects `< v0.8.4-rc1`)

> CVE-2026-54037 CVE ↗ STABLE
>
> `CVE-2026-54037` : Incomplete Fix for CVE-2025-7105 — `/api`/convos/duplicate Lacks Rate Limiting Applied to `/api`/convos/fork (CVSS 6.5, affects `<= 0.8.3`)

> CVE-2026-44654 CVE ↗ STABLE
>
> `CVE-2026-44654` : Shared-agent editor can globally delete owner's file records — breaks owner's other private agents (CVSS 8.1, affects `<= 0.8.3`)

> CVE-2026-54027 CVE ↗ STABLE
>
> `CVE-2026-54027` : Image Upload Route Bypasses Agent Permission Check — Incomplete Fix for File Upload Authorization (CVSS 6.5, affects `<= 0.8.3`)

> CVE-2026-54029 CVE ↗ STABLE
>
> `CVE-2026-54029` : IDOR in Message Deletion — Incomplete Fix for CVE-2024-41703 Leaves deleteMessages() Without User Filter (CVSS 5.3, affects `<= 0.8.3`)

> CVE-2026-31942 CVE ↗ STABLE
>
> `CVE-2026-31942` : IDOR in API Keys Management allows any authenticated user to overwrite other users' API keys (CVSS 7.1, affects `<= 0.7.6`)

CVE-2026-44653 CVE ↗ STABLE

`CVE-2026-44653` : Shared MCP Server View Leaks Decrypted Admin Secrets (CVSS 6.5, affects `v0.8.3`)

CVE-2026-54024 CVE ↗ STABLE

`CVE-2026-54024` : Incomplete Fix for CVE-2024-11171 — Conversation Import Multer Instance Missing File Size Limits (CVSS 6.5, affects `<= 0.8.3`)

CVE-2026-54036 CVE ↗ STABLE

`CVE-2026-54036` : 2FA Re-enrollment Allows Full Account 2FA Takeover Without OTP Verification (CVSS 5.3, affects `<= 0.8.3-rc1`)

CVE-2026-54025 CVE ↗ STABLE

`CVE-2026-54025` : Stored XSS via unescaped image alt text in markdown artifact preview (CVSS 5.4, affects `<= 0.7.8`)

CVE-2026-34371 CVE ↗ STABLE

`CVE-2026-34371` : Arbitrary File Write via `execute_code` Artifact Filename Traversal (CVSS 6.3, affects `v0.8.3`)

CVE-2026-31943 CVE ↗ STABLE

`CVE-2026-31943` : SSRF protection bypass via IPv4-mapped IPv6 normalization in isPrivateIP (CVSS 8.5, affects `v0.8.3-rc2`)

CVE-2026-31945 CVE ↗ STABLE

`CVE-2026-31945` : LibreChat Server-Side Request Forgery using DNS resolution (CVSS 7.7, affects `v0.8.2-rc2, v0.8.2-rc3, v0.8.2`)

CVE-2026-31944 CVE ↗ STABLE

`CVE-2026-31944` : MCP OAuth callback does not validate browser session, allows token theft via redirect link (CVSS 7.6, affects `>= v0.8.2, <= 0.8.2-rc3`)

CVE-2026-31949 CVE ↗ STABLE

`CVE-2026-31949` : Denial of Service (DoS) via Unhandled Exception in DELETE `/api`/convos (CVSS 6.5, affects `<= 0.8.2`)

CVE-2026-22252 CVE ↗ STABLE

`CVE-2026-22252` : LibreChat MCP Stdio Remote Command Execution (CVSS 9.1, affects `0.8.2-rc1`)

CVE-2025-69221 CVE ↗ STABLE

`CVE-2025-69221` : LibreChat Insufficient Access Control on Agent Permission Queries (CVSS 4.3, affects `0.8.1-rc2`)

CVE-2025-66452 CVE ↗ STABLE

`CVE-2025-66452` : Invalid json payloads in api requests were reflected entirely in the http response - may lead to XSS risk if the user includes html/javascript code in the request body (CVSS 6.1, affects `0.8.0`)

CVE-2025-66451 CVE ↗ STABLE

`CVE-2025-66451` : Insufficient json validation on api requests to prompt creation - may expose invalid or the possibility to modify permissions of the prompt (CVSS 6.5, affects `0.8.0`)

CVE-2025-66450 CVE ↗ STABLE

`CVE-2025-66450` : JSON Injection on chat POST leading to remote resource inclusion - may lead to PXSS on image upload (CVSS 5.4, affects `0.8.0`)

CVE-2025-54868 CVE ↗ STABLE

`CVE-2025-54868` : Reading of arbitrary Chats (CVSS 7.5, affects `v0.0.6 - v0.7.7`)

WAS THIS USEFUL? 👍 👎

nanobot

Feedback ⓘ CVE-2026-48716 CVE ↗ STABLE

Nanobot is a self-hosted personal AI assistant framework with tools, memory, scheduled tasks, and integrations with chat applications.

`CVE-2026-48716` : Path traversal via unsanitized WhatsApp document fileName enables arbitrary file write (CVSS 8.7, affects `<= 0.1.5.post3`)

▾ **3 more releases in this issue** · 2026-03-11 → 2026-05-24

> CVE-2026-35589 CVE ↗ STABLE
>
> `CVE-2026-35589` : Cross-Site WebSocket Hijacking (CSWSH) in nanobot WhatsApp Bridge (CVSS 8.0, affects `<= 0.1.4.post6`)

> CVE-2026-33654 CVE ↗ STABLE
>
> `CVE-2026-33654` : Zero-Click Indirect Prompt Injection and Authentication Bypass via Email Polling (CVSS 9.8, affects `<= 0.1.4.post5`)

> CVE-2026-31977 CVE ↗ STABLE
>
> `CVE-2026-31977` : `is_allowed` Authorization Bypass via `|` `-Separated` Sender ID Injection (affects `< 0.1.4.post3`)

WAS THIS USEFUL? 👍 👎

evolver

Feedback ⓘ CVE-2026-42077 CVE ↗ STABLE

The GEP-powered self-evolving engine for AI agents. Auditable evolution with Genes, Capsules, and Events. | evomap.ai

`CVE-2026-42077`: Prototype Pollution via Object.assign() in mailbox store operations (CVSS 5.2, affects `< 1.69.3`)

▾ **2 more releases in this issue** · 2026-04-20

> CVE-2026-42075 CVE ↗ STABLE
>
> `CVE-2026-42075`: Path Traversal via `--out` flag in `fetch` command allows Arbitrary File Write (CVSS 8.1, affects `< 1.69.3`)

> CVE-2026-42076 CVE ↗ STABLE
>
> `CVE-2026-42076`: Command Injection via `execSync` in _extractLLM() function allows Remote Code Execution (CVSS 9.8, affects `< 1.69.3`)

WAS THIS USEFUL? 👍 👎

Microsoft Semantic Kernel

Feedback ⓘ CVE-2026-26030 CVE ↗ STABLE

Semantic Kernel provides an open-source SDK for building AI agents and integrating language models into C#, Python, and Java applications.

`CVE-2026-26030`: InMemoryVectorStore filter functionality vulnerable to remote code execution (CVSS 9.9, affects `< 1.39.4`)

▾ **1 more release in this issue** · 2026-02-06 → 2026-02-19

> CVE-2026-25592 CVE ↗ STABLE
>
> `CVE-2026-25592`: Arbitrary File Write via AI Agent Function Calling in .NET SDK (CVSS 9.9, affects `< 1.71.0`)

WAS THIS USEFUL? 👍 👎

Cloudflare Agents

Feedback ⓘ CVE-2026-1664 CVE ↗

The Cloudflare Agents SDK builds stateful agents on Workers and Durable Objects, with persistent state, scheduling, WebSockets, and MCP server support.

`CVE-2026-1664`: Insecure Direct Object Reference (IDOR) via Header-Based Email Routing (affects `< 0.3.7`)

WAS THIS USEFUL? 👍 👎

Khoj

Feedback ⓘ CVE-2025-69207 CVE ↗ STABLE

Khoj is a personal AI assistant that searches documents and the web, runs research tasks, and schedules automations.

`CVE-2025-69207`: IDOR in Notion OAuth Flow Enables Index Poisoning (CVSS 5.4, affects `latest`)

▾ **3 more releases in this issue** · 2024-07-08 → 2026-02-01

> CVE-2024-52294 CVE ↗ STABLE
>
> `CVE-2024-52294`: IDOR in subscription management allows unauthorized subscription modifications (CVSS 4.3, affects `<1.29.0`)

> CVE-2024-43396 CVE ↗ STABLE
>
> `CVE-2024-43396`: Stored XSS In Automate (Preview feature) (CVSS 5.4, affects `<1.15`)

> CVE-2024-25639 CVE ↗ STABLE
>
> `CVE-2024-25639`: Prompt Injection triggered XSS vulnerability in Khoj Obsidian, Desktop and Web clients (CVSS 5.9, affects `<1.13.0`)

WAS THIS USEFUL? 👍 👎

Mastra

Feedback ⓘ CVE-2025-61685 CVE ↗

Mastra is a TypeScript framework for AI applications and agents, providing workflows, agent memory, RAG, tool calling, and evals.

`CVE-2025-61685`: Mastra Docs MCP Server `@mastra/mcp-docs-server` Leads to Information Exposure (CVSS 6.5, affects `<= 0.13.8`)

WAS THIS USEFUL? 👍 👎

deepset Haystack

Feedback ⓘ CVE-2024-41950 CVE ↗

Haystack builds language-model applications with modular pipelines for retrieval-augmented generation and agents.

`CVE-2024-41950`: Insecure Jinja2 templates rendered in Haystack Components can lead to RCE (CVSS 7.5, affects `<=2.3.0`)

WAS THIS USEFUL? 👍 👎

◆ MCP Servers & Tooling

rails-ai-context

Feedback ⓘ v5.38.0 NOTES ↗ RANK

Stop AI from guessing your Rails app. 35 MCP tools give coding agents ground truth on demand: schema, models, routes, controllers, views, jobs and conventions.

> rails-ai-context v5.38.0 ships a boot-free settings reader and shared app-root resolution, a major expansion of schema/`get_schema` coverage for PostgreSQL exclusion constraints, enum types, check constraints and generated columns, sweeping extensions to credential and secret filtering across tools and formats, and a breaking change to how Ruby and YAML exclusion config combine.

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

01 Boot-free settings reader and shared app-root resolution NEW

85

Adds RailsAiContext::UnbootedSettings.read(root) to return an app's settings as its boot would apply them, without running anything, with per-file/line/key reporting of unreadable values. Adds `CLI::AppRoot.target` as the shared single answer used by the server, the hook, `doctor`, and the settings reader.

Inspect an app's effective settings without booting it — useful in CI or before a deploy to confirm what rails-ai-context will see.

ruby

```ruby
require 'rails_ai_context'
result = RailsAiContext::UnbootedSettings.read('/path/to/my_app')
puts result.inspect
```

Run a specific tool against one of several apps in a mono-repo folder from the command line, without switching directories.

```
$ rails-ai-context --app-path ./api tool rails_get_schema
```

— Runnable code and CLI examples given, mechanism explained v5.38.0

02 Schema support for exclusion constraints, enums, checks and generated columns

IMPROVED

80

Extends the schema/`get_schema` answer with exclusion constraints (`t.exclusion_constraint` in schema.rb, `ADD CONSTRAINT ... EXCLUDE` in

structure.sql, `add_exclusion_constraint` and `remove_exclusion_constraint` in migrations, and the live PostgreSQL connection), PostgreSQL enum types, check constraints, and generated columns for apps with no schema dump (replayed from `create_enum`, `add_enum_value`, `rename_enum_value`, `rename_enum`, and `drop_enum`, including on secondary databases), and foreign-key name preservation (when not the Rails default) plus `validate: false` on check constraints, all read from schema.rb, structure.sql, migrations, and the live connection.

— Mechanism and sources detailed but no direct usage example shown `v5.38.0`

---

03 **Expanded credential and secret filtering across tools and formats**

`IMPROVED` 80

Adds `.npmrc`, `.yarnrc`, and `.bundle/config` to the default `sensitive_patterns` list, and filters `rails_search_code` and the unbooted settings reader to ignore a ripgrep config named by `RIPGREP_CONFIG_PATH`. Extends `rails_get_turbo_map`, `rails_get_callbacks`, `rails_get_job_pattern`, `.ai-context.json`, and every generated file to show filtered text for the context sources they record. Extends credential filtering to cover GitHub (`gho_`, `ghr_`, `github_pat_`), AWS (`ASIA`), Stripe, Slack, JWT, and private-key tokens wherever they appear in the assembled context, including inside `Set`, `Struct`, and `Data` values; gem server credentials via `BUNDLE_<HOST>`, `bundle config <host> <credentials>`, `Bundler.settings`, `ENV[...]`, `ENV.store`, `ENV.fetch` defaults, fish `set -x`, cmd `setx`, PowerShell, Dockerfile `ENV` and exec-form `RUN`, `.bundle/config`, and a key used as a URL's user; and multi-line forms such as YAML block scalars, quotes spanning lines, heredocs, trailing-backslash continuations, and `name` / `value` key pairs in JSON, YAML, container env entries, and CI variables. Extends `sensitive_patterns` gating to cover every directory walk and constant lookup (models, controllers, jobs, migrations, base classes, concerns, route files, views, partials, Stimulus and Turbo listings), all template extensions (rabl, arb, liquid, markdown, prawn, plain html), `app/graphql` listings, template counts, and custom `app/` directories. Extends `rails_review_changes` to list a changed file under its own name when the path contains a byte outside ASCII, a quote, a backslash, or a control character, and checks that name against `sensitive_patterns`. Extends `rails_get_edit_context` to pick redaction rules by the file's own name and to identify log, YAML, and JSON files by extension in any case.

— Every named filter, token type and config surface enumerated `v5.38.0`

---

04 **Union of Ruby and YAML exclusion config** `BREAKING` 65

Assigning `sensitive_patterns`, `excluded_paths`, or `skip_tools` in Ruby no longer drops entries listed in `.rails-ai-context.yml`; the result is now the union of both sources, which may expand the effective exclusion set for apps that previously relied on a Ruby assignment to override YAML entries.

— Clear before/after behavior change, no migration command given v5.38.0

└──▷ **BREAKING ON UPGRADE**

! Assigning `sensitive_patterns`, `excluded_paths`, or `skip_tools` in Ruby no longer drops entries listed in `.rails-ai-context.yml`; the result is now the union of both sources, which may expand the effective exclusion set for apps that previously relied on a Ruby assignment to override YAML entries.

WAS THIS USEFUL?  

memwhale

Feedback ⓘ v0.23.0 NOTES ↗

Persistent, local memory for developers and their coding agents. Records commands, output, errors, and the fixes that worked into SQLite and serves them over MCP.

Adds a desktop app (macOS/Linux), `mw recall` CLI command, `/recall` dashboard page, menu-bar notifications, and schema 17 note-text storage.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v0.23.0 https://github.com/wuisabel-gif/MemWhale.git
# already have the repo? check out this version:
$ git checkout v0.23.0
```

Adds `mw recall` and `mw recall --limit 20` to list every note and warning shown to agents, with agent identity, directory, failing command, exact text, and outcome (`followed, worked`, `followed, no pass yet`, `not followed`, `then passed`, or `no pass yet`).

Adds a `/recall` page to the web dashboard, listing the same note and warning history surfaced by `mw recall`.

Ships a desktop app (macOS `.dmg`, Linux `.deb`, Linux `.AppImage`) with Home, Search, Recall, Commands, and Lessons views, reading the same local SQLite store as the CLI and agent hooks.

Adds a menu-bar presence to the desktop app that polls every five seconds and fires a system notification (e.g. Warned claude before cargo build) when a hook writes a new note or warning.

Stores the exact text of each note and warning in the SQLite store (schema 17, adding `text` to `recall_hints`); older entries display without it.

+1 more

NEW INTEGRATION Codewhale capture integration

WAS THIS USEFUL? 👍 👎

mcp

Feedback ⓘ v1.1.2 NOTES ↗ RANK

Rapidly build MCP servers for your Laravel applications.

> Laravel MCP's v1.1.2 release focuses on OAuth security hardening, adding SSRF mitigations and more transparency to the authorization flow, alongside more flexible MCP Apps configuration.

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

01 **MCP Apps without ui.domain setting** IMPROVED 40

MCP Apps can now be configured without a `ui.domain` setting.

— Names exact config key but no further mechanism given v1.1.2

02 **OAuth SSRF mitigations and authorize page transparency**

IMPROVED 35

Laravel MCP v1.1.2 hardens OAuth with SSRF mitigations and now shows the redirect URI on the OAuth authorize page.

— Mentions SSRF mitigation and UI change but no mechanism detail v1.1.2

WAS THIS USEFUL? 👍 👎

fablecut

Feedback ⓘ v1.17.0 NOTES ↗ RANK

Zero-dependency browser video editor that AI agents can drive — JSON timeline, MCP + REST, live-reloading UI

> FableCut shipped an in-browser transcription and transcript-based editing suite powered by Whisper, alongside automatic silence detection/removal, agent-driven caption generation with SRT/VTT export, new word-level text animation fields, and Apple Silicon hardware-accelerated export.

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

01 Silence detection and one-step removal NEW

93

Adds the `fablecut_silence` agent tool with a `cut:true` parameter to detect and remove silence pauses in one revision, supporting thresholds for minimum pause duration and silence level. Also adds silence detection in the UI via the Inspector's Silence section (or right-click → Find silences) on audio/video clips, marking pauses on the timeline with one-step Cut extraction across targeted tracks.

Have an AI agent strip all detected silences from a recording in one automated revision.

```javascript
fablecut_silence({ cut: true })
```

— Runnable agent command plus exact UI navigation path v1.17.0

02 In-browser transcription and transcript-based editing NEW

75

Adds the `fablecut_transcribe` agent function/tool for speech-to-text with word-level timestamps, accessible from the Transcript tab or programmatically; it supports a `cut:true` parameter to read timed transcript lines, find phrases, and cut them in one revision. The Transcript tab runs Whisper in-browser (77 MB or 250 MB model) to produce timed word-level transcripts stored per file in `analysis/transcripts/`, with click-to-seek and delete-to-cut word selection.

— Names function, models and storage path but no exact invocation shown v1.17.0

03 Caption generation with SRT/VTT export NEW

73

Adds the `fablecut_captions` agent tool to generate captions from transcripts via agents, plus a Transcript tab Add captions feature that produces captions in four styles — subtitle, karaoke, highlight, and word pop — with SRT and VTT export and subtitle file import.

— Names tool, formats and styles but no runnable command v1.17.0

04 **Word-level animation fields for text clips** NEW 65

Adds `wordTimes` and `highlightColor` fields to text clips, enabling word-pop, word-slide, karaoke, and the new `highlight` animation to follow per-word timing.

— Names fields and animations but no usage example v1.17.0

THINNER COVERAGE BELOW

05 **Hardware-accelerated export on Apple Silicon** IMPROVED 25

v1.17.0 adds Apple Silicon hardware-accelerated export, with no further mechanism or scope described.

— Mentioned only in summary, no technical detail given v1.17.0

WAS THIS USEFUL? 👍 👎

tabularis

Feedback ⓘ v0.30.0 NOTES ↗ RANK

Open-source desktop SQL workspace with 3 built-in database drivers and 24 shipped driver plugins, including SQL Server, DuckDB, ClickHouse and Redis. Built-in MCP server for Claude, Cursor and Devin, SQL notebooks and visual EXPLAIN.

> Tabularis v0.30.0 adds a self-hosted browser UI, scheduled database backups, chart views for query results, and a round of data grid, EXPLAIN, filter, history and AI command-palette improvements.

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

01 **Self-hosted browser UI via Docker** NEW 83

Tabularis can now run as a browser-hosted web UI instead of only as a desktop app, with authentication support. It's launched with `docker run -d -p 127.0.0.1:8080:8080 -v tabularis-data:/data ghcr.io/tabularisdb/tabularis:0.30.0`.

Self-host the Tabularis web UI in a browser with persistent data — no desktop app required.

```
$ docker run -d -p 127.0.0.1:8080:8080 -v tabularis-data:/data
ghcr.io/tabularisdb/tabularis:0.30.0
```

— Exact runnable docker command given v0.30.0

02 **Per-scope network proxy modes** NEW 65

Adds per-scope proxy modes — system, custom, and disabled — configurable independently for app network requests, database connections, AI, and SSH tunnels, as shown in Settings > Network with database connections on Disabled, AI on Custom via a SOCKS5 server, and SSH tunnels on Disabled.

— Named modes and scopes, UI path only v0.30.0

03 **Scheduled database content backups** NEW 60

Adds scheduled database content backups (Milestone 1), configurable per connection with a time, retention count, and scope of structure and/or data — shown configured for a SQLite connection to run daily at 02:00 local time keeping 7 backups of structure and data.

— Describes schedule and retention but milestone scope unclear v0.30.0

04 **Data grid export, freeze and resize** IMPROVED 60

The data grid gains the ability to export selected rows as CSV or JSON, freeze the first data column while scrolling horizontally, and resize columns with mouse and keyboard.

Export only the rows you care about from a result set without downloading the full table.

> 📍 In the data grid, select the target rows, then open the export menu and choose CSV or JSON.

— Three concrete grid capabilities with export formats named v0.30.0

THINNER COVERAGE BELOW

05 **Chart view for query results** NEW 55

The SQL editor now offers a chart view for query results, switchable from a `Chart` tab above the results grid, letting users visualise query output (e.g. a line chart) immediately after running a query instead of only viewing a table.

Visualise query results as a chart immediately after running a SQL query in the editor.

> 📍 In the SQL editor, run a query, then select the `Chart` tab above the results grid to switch from table view to chart view.

— Clear UI navigation but no config detail v0.30.0

06 **EXPLAIN CSV export cost and gap columns** IMPROVED 40

Adds numeric cost and gap columns to the EXPLAIN CSV export, giving more detail when exporting query plan data from the visual EXPLAIN view.

— Names the export and new columns but no mechanism v0.30.0

07 **Table filter case-sensitivity and keyboard navigation** IMPROVED 40

Improves table filters with a case-sensitivity toggle for PostgreSQL text filters and keyboard navigation support.

— Specific but brief, no UI path given v0.30.0

08 **AI generate and explain in command palette** NEW 30

Adds AI generate and explain actions to the command palette.

— Bare feature name, no detail on behaviour v0.30.0

09 **Source maps in nightly builds** IMPROVED 25

Ships source maps in nightly builds.

— One-line developer-facing note, no further detail v0.30.0

WAS THIS USEFUL? 👍 👎

api-mcp-server Feedback ⓘ v3.10.0 NOTES ↗

api-mcp-server v3.10.0 bundles agent skill definitions for headless operation workflows.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v3.10.0 https://github.com/hostinger/api-mcp-server.git
# already have the repo? check out this version:
$ git checkout v3.10.0
```

WAS THIS USEFUL? 👍 👎

mcp-server-plugin

Feedback ⓘ 0.221.v42ee125fc488 NOTES ↗

RANK

> The Jenkins MCP Server plugin added tools to cancel builds and queue items, with richer parameters and return values on the cancellation call.

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

01 **Build and queue cancellation tools** NEW 65

The MCP server now exposes a tool to cancel builds and queue items. The `cancelBuild` tool was enriched to accept a separate `queueId` parameter and to return a status value indicating the result of the cancellation.

— Names the tool and its new parameter/return but lacks usage detail. 0.221.v42ee125fc488

WAS THIS USEFUL? 👍 👎

pi-mcp-adapter

Feedback ⓘ v7.0.0 NOTES ↗ RANK

Token-efficient MCP adapter for Pi coding agent

> pi-mcp-adapter v7.0.0 ships async MCP Tasks, semantic tool search, typed value extraction, and structured deny reasons, alongside a substantial breaking overhaul of credential storage and semantic-search config naming.

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

01 Semantic search config renamed from jev to semanticSearch

90

BREAKING

`settings.jev` is renamed to `settings.semanticSearch` in config files. Sub-keys changed: `semanticCandidateLimit` → `candidateLimit`, `semanticMinProbability` → `minProbability`; `requestTimeoutMs`, `maxRetries`, `maxStateBytes`, and `maxQuestionsPerRequest` are removed, and `model` now accepts a Pi model as `"provider/id"`. `settings.jev: false` no longer disables semantic search — use `settings.semanticSearch: false` instead. `/mcp-adapter jev setup` is renamed to `/mcp-adapter semantic setup`, and the `Jev*` type exports from `pi-mcp-adapter/types` and the `@typesafe-ai/sdk` dependency are removed.

— Exhaustively names every renamed and removed config key and command v7.0.0

02 Async MCP Tasks via protocolVersion auto NEW

88

Adding a `"protocolVersion": "auto"` server config key enables async MCP Task execution so long-running tools (CI runs, report exports) return immediately and deliver their result as a new session message when complete. A new `/mcp-adapter tasks` command lists running background jobs and can stop one.

Enable async background execution for a slow MCP server (e.g. a CI runner) so long tool calls return immediately and deliver results as a new session message.

```json
{
  "mcpServers": {
    "ci-runner": {
      "command": "npx",
      "args": ["ci-mcp-server"],
      "protocolVersion": "auto"
    }
  }
}
```

— Names config key, command, and behavior with a runnable config example v7.0.0

03 **Typed value extraction from tool results** NEW 88

mcp({ tool, args, extract }) returns only named typed fields from a tool result, e.g. `extract: { total: "number", due: "date" }`, read by Cloudflare's Clairic model. It is enabled per-server via the `settings.extraction.allowedServers` config key.

Pull only the fields you need from a verbose tool result to save context — useful when a tool returns a long order page but you only care about the total and due date.

javascript

```
mcp({ tool: "get_invoice", args: { id: "INV-9912" }, extract: {
total: "number", due: "date" } })
```

Configure semantic search to use a specific Pi classifier model and restrict value extraction to one trusted server.

json

```
{
  "settings": {
    "semanticSearch": {
      "model": "cloudflare/clef",
      "candidateLimit": 20,
      "minProbability": 0.25
    },
    "extraction": {
      "allowedServers": ["invoicing-server"]
    }
  }
}
```

— Named API, config key, and runnable example of field extraction v7.0.0

04 **Semantic tool search** NEW 82

mcp({ search, searchMode: "semantic" }) ranks tools by meaning using classifier models (Jev, Clef, GPT-6 Luna, or any Pi credential) instead of word matching, with search results up to 5x smaller by showing full parameters only for high-confidence matches. The `settings.semanticSearch.model` config key (accepting a Pi model as `"provider/id"`) selects which classifier model ranks tools.

Configure semantic search to use a specific Pi classifier model and restrict value extraction to one trusted server.

```json
{
  "settings": {
    "semanticSearch": {
      "model": "cloudflare/clef",
      "candidateLimit": 20,
      "minProbability": 0.25
    },
    "extraction": {
      "allowedServers": ["invoicing-server"]
    }
  }
}
```

— Named API call and config key with example, mechanism partly explained v7.0.0

05 **Adapter key store removed, credentials via Pi login** BREAKING

81

The adapter's own key store is removed: `pi-mcp-adapter key` is now an unknown command, and `SYSTEMONE_ENDPOINT` and `SYSTEMONE_API_KEY` environment variables do nothing. Supply credentials to Pi instead via `/login` to OpenRouter, or set `TYPESAFE_API_KEY` or `OPENROUTER_API_KEY` in Pi's environment.

— Names removed command, dead env vars and exact migration path v7.0.0

06 **Structured deny reasons for tool approval** NEW

68

New `McpToolApprovalDenial` and `McpToolApprovalHandlerResult` type exports from `pi-mcp-adapter` and `pi-mcp-adapter/types` let broker handlers return `{ decision: "deny", reason }` so the model receives a structured denial reason. New `denied` and `denyReason` fields on `HostManagedMcpError` carry the reason for a denied linked-resource read in the host-managed adapter.

— Named types and fields but no usage example given v7.0.0

THINNER COVERAGE BELOW

07 **Variable expansion in config values** NEW

40

Supports `${VAR:-default}` variable expansion syntax in config values, matching Claude Code's `.mcp.json` behavior.

— Syntax named but no example or file location given v7.0.0

08 **Post-update release highlights** NEW

25

The first Pi session after an adapter update shows a summary of skipped releases once, with a link to the full changelog.

— Brief prose description, no named surface or example v7.0.0

└──▷ BREAKING ON UPGRADE

! The adapter's own key store is removed: `pi-mcp-adapter key` is now an unknown command, and `SYSTEMONE_ENDPOINT` and `SYSTEMONE_API_KEY` environment variables do nothing. Supply credentials to Pi instead via `/login` to OpenRouter, or set `TYPESAFE_API_KEY` or `OPENROUTER_API_KEY` in Pi's environment.

! `settings.jev` is renamed to `settings.semanticSearch` in config files. Sub-keys also changed: `semanticCandidateLimit` → `candidateLimit`, `semanticMinProbability` → `minProbability`; `requestTimeoutMs`, `maxRetries`, `maxStateBytes`, and `maxQuestionsPerRequest` are removed. `model` now accepts a Pi model as `"provider/id"`.

! `settings.jev: false` no longer disables semantic search; set `settings.semanticSearch: false` instead.

! `/mcp-adapter jev setup` is renamed to `/mcp-adapter semantic setup`.

! The `Jev*` type exports from `pi-mcp-adapter/types` and the `@typesafe-ai/sdk` dependency are removed.

WAS THIS USEFUL?  

robloxstudio-mcp

Feedback ⓘ v3.4.2 NOTES ↗ RANK

MCP server for Roblox Studio runtime debugging, playtest control, screenshots/input, multiplayer testing, and per-peer server/client eval from AI agents.

> robloxstudio-mcp v3.4.2 adds full MicroProfiler recording and analysis plus atomic and dry-run modes for property edits, alongside breaking schema and error-shape changes that require matching versions across Studio connections.

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

01 **MicroProfiler recording and analysis** NEW 93

Adds `capture_micro_profiler` with `action: "start"` / `"stop"` / `"analyze"` to record up to 30 s of MicroProfiler frames and analyze them on the MCP host, including re-reading saved `.gprx` dumps; results include `data_quality.render_throttled` and `process_roles` fields.

Record a MicroProfiler session around other tool calls to capture real frame data during a suspect workload.

```
$ # Start recording
capture_micro_profiler action="start"

# ... run other tool calls here ...

# Stop and analyze
capture_micro_profiler action="stop"
```

— Full mechanism, named fields, and a runnable start/stop/analyze command. v3.4.2

02 **Atomic and dry-run modes for set_properties** NEW 80

Adds `atomic: true` to `set_properties` for all-or-nothing property edits — validates every property before writing and rolls back earlier writes if any one fails. Adds `dryRun: true` to `set_properties` to preview changes without writing them.

Preview a batch of property changes before committing, then apply atomically so a single failure rolls back all writes.

```
$ # Preview only — no writes
set_properties dryRun=true ...

# Commit atomically
set_properties atomic=true ...
```

— Names both flags and their rollback/preview mechanism with example usage. v3.4.2

---

03 **capture_micro_profiler schema and version compatibility change**

75

BREAKING

`capture_micro_profiler` now returns `schema_version: 3`; the fields `max_events`, `frame_window`, `history_truncated`, and `frames_evicted` are removed, and `output_path` now writes the raw `.gprx` file instead of a processed result. MicroProfiler captures through a shared Studio connection require v3.4.2 on both the primary and secondary server; a 3.4.2 secondary refuses captures routed through an older primary with error `primary_unsupported`.

— Names exact schema/field/error changes but gives no migration command. v3.4.2

THINNER COVERAGE BELOW

---

04 **get_simulation_state error shape change** BREAKING

45

`get_simulation_state` `errors` entries changed shape to `{ code, message }`.

— Names the endpoint and new shape but no further detail or example. v3.4.2

└──▷ **BREAKING ON UPGRADE**

! `capture_micro_profiler` now returns `schema_version: 3`; the fields `max_events`, `frame_window`, `history_truncated`, and `frames_evicted` are removed, and `output_path` now writes the raw `.gprx` file instead of a processed result.

! `get_simulation_state` `errors` entries changed shape to `{ code, message }`.

! MicroProfiler captures through a shared Studio connection require v3.4.2 on both the primary and secondary server; a 3.4.2 secondary refuses captures routed through an older primary with error `primary_unsupported`.

WAS THIS USEFUL? 👍 👎

pullmd

Feedback ⓘ v3.16.0 NOTES ↗ RANK

Self-hosted URL- and file-to-Markdown service for humans and AI agents - web pages, documents, images, audio, YouTube. PWA + REST + MCP + Claude Code skill, Reddit-aware, refreshable share links.

> pullmd v3.16.0 adds opt-in image transcription for Reddit posts with per-request and per-instance controls, a startup warning when paid-tier provider keys are exposed on unauthenticated instances, and a resulting output format change for untranscribed images.

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

01 **Reddit post image transcription** NEW 95

New `PULLMD_EMBEDDED_IMAGES` env var (`off` | `on-request` | `always` , default `on-request`) controls whether a configured vision provider transcribes images inside Reddit posts; when set to `on-request` , callers must pass the `images=on` / `images=off` request parameter or `images: "on"` on the MCP `read_url` tool to opt in. `PULLMD_EMBEDDED_IMAGES_MAX` (default `4` , range `1` – `20`) caps images transcribed per post, with later gallery items marked but not downloaded. Transcribed posts get an `images_transcribed: "<done>/<total>"` frontmatter field and summed `llm_*` usage fields. Support is limited to Reddit hosts `i.redd.it` and `preview.redd.it` , single jpg/jpeg/png/webp images, non-animated gallery or inline images, capped at 20 MB and 90 seconds per post. Each transcribed image counts one unit against `PULLMD_LLM_RATE_LIMIT` , and transcripts are cached per image so the same image is never re-sent across URL variants (`old.reddit.com` , `redd.it` short links, `/s/` share links).

Enable automatic Reddit image transcription for all requests on a self-hosted instance so downstream LLMs always receive image context without callers having to opt in.

```
$ PULLMD_EMBEDDED_IMAGES=always PULLMD_EMBEDDED_IMAGES_MAX=8 node server.js
```

Cap image transcription to 2 images per post and hard-disable the feature, keeping output byte-identical to v3.15.2 regardless of what callers request.

```
$ PULLMD_EMBEDDED_IMAGES=off PULLMD_EMBEDDED_IMAGES_MAX=2 node server.js
```

— Full mechanism, scope, and all config/request surfaces named with examples v3.16.0

02 **Startup warning for open paid-tier provider access** NEW

65

A startup `[auth]` warning is now emitted when `PULLMD_AUTH_MODE` is disabled and any provider key (`PULLMD_VISION_API_KEY` , `PULLMD_STT_API_KEY` , `PULLMD_LLM_API_KEY` , `PULLMD_PDF_OCR_API_KEY`) is set, alerting operators that all reachable clients can consume those paid tiers.

— Names exact env vars and warning but no fix/remediation shown v3.16.0

THINNER COVERAGE BELOW

03 **Marker for untranscribed Reddit images** BREAKING

45

On instances with a vision provider and the default `on-request` setting, Reddit image posts now include a one-line Image not transcribed marker under each image that lacks a transcript, changing the output format relative to v3.15.2 for those posts.

— States the output change but no migration guidance given v3.16.0

└──▷ **BREAKING ON UPGRADE**

! On instances with a vision provider and the default `on-request` setting, Reddit image posts now include a one-line Image not transcribed marker under each image that has no transcript, changing the output format relative to v3.15.2 for those posts.

WAS THIS USEFUL? 👍 👎

whodb

Feedback ⓘ CVE-2026-112005 CVE ↗ STABLE

Where data access meets operational intelligence

`CVE-2026-112005` : Unauthenticated arbitrary host file read and write through a DuckDB source session, disclosing every logged-in user's database credential in plaintext (CVSS 9.4, affects `>= 0.104.0`)

▾ **2 more releases in this issue** · 2025-02-06 → 2026-10-15

> CVE-2025-24787 CVE ↗ STABLE
>
> `CVE-2025-24787` : Parameter injection in DB connection URIs leading to local file inclusion (CVSS 8.6, affects `<=0.45.0`)

> CVE-2025-24786 CVE ↗ STABLE
>
> `CVE-2025-24786` : Path traversal opening Sqlite3 database (CVSS 10.0, affects `<=0.45.0`)

WAS THIS USEFUL? 👍 👎

MCP TypeScript SDK Feedback ⓘ CVE-2026-111768 CVE ↗

STABLE

The MCP TypeScript SDK provides TypeScript clients, servers, and transports for Model Context Protocol.

`CVE-2026-111768`: One HTTP request within the default size limit could exhaust server memory during validation (CVSS 7.5, affects `>= 1.3.0, < 1.32.0`)

▾ **8 more releases in this issue** · 2025-12-02 → 2026-10-15

> CVE-2026-111770 CVE ↗ STABLE
>
> `CVE-2026-111770`: Shared transport or server instances delivered one HTTP client's messages to another (CVSS 7.7, affects `>= 2.0.0, < 2.3.0`)

> CVE-2026-111771 CVE ↗ STABLE
>
> `CVE-2026-111771`: Server bearer authentication accepted tokens issued for another service (CVSS 6.8, affects `>= 1.6.0, < 1.32.0`)

> CVE-2026-111767 CVE ↗ STABLE
>
> `CVE-2026-111767`: Experimental tasks were not tied to the session that created them (CVSS 8.6, affects `>= 1.24.0, < 1.32.0`)

> CVE-2026-111769 CVE ↗ STABLE
>
> `CVE-2026-111769`: HTTP client transports and OAuth requests followed cross-origin redirects and re-sent custom headers and request bodies (CVSS 6.5, affects `< 1.32.0`)

> CVE-2026-104850 CVE ↗ STABLE
>
> `CVE-2026-104850`: OAuth client could send credentials to an authorization server chosen by the MCP server (CVSS 7.5, affects `>= 1.12.0, < 1.31.0`)

> CVE-2026-25536 CVE ↗ STABLE
>
> `CVE-2026-25536`: Sharing server/transport instances can leak cross-client response data (CVSS 7.1, affects `>= 1.10.0, <= 1.25.3`)

> CVE-2026-0621 CVE ↗ STABLE

> `CVE-2026-0621`: Regular Expression Denial of Service (ReDoS) in UriTemplate (CVSS 7.5, affects `<1.25.2`)

> `CVE-2025-66414` CVE ↗ STABLE
>
> `CVE-2025-66414`: DNS Rebinding Protection Disabled by Default in Model Context Protocol TypeScript SDK for Servers Running on Localhost (CVSS 8.1, affects `< 1.24.0`)

WAS THIS USEFUL? 👍 👎

MCP Python SDK Feedback ⓘ CVE-2026-109527 CVE ↗ STABLE

The MCP Python SDK provides clients and servers for exposing tools, resources, and prompts through Model Context Protocol.

`CVE-2026-109527`: HTTP server transports and OAuth endpoints read request bodies with no size limit (CVSS 7.5, affects `< 1.29.1`)

▾ **8 more releases in this issue** · 2025-07-04 → 2026-09-30

> CVE-2026-59951 CVE ↗ STABLE
>
> `CVE-2026-59951`: Streamable HTTP sessions were never reclaimed by default and kept their initialize payloads (CVSS 7.5, affects `>= 1.8.0, < 1.30.0`)

> CVE-2026-115563 CVE ↗ STABLE
>
> `CVE-2026-115563`: OAuth client could send credentials to an authorization server chosen by the MCP server (CVSS 7.5, affects `>= 2.0.0a1, < 2.2.0`)

> CVE-2026-59950 CVE ↗ STABLE
>
> `CVE-2026-59950`: WebSocket server transport does not support Host/Origin validation (CVSS 8.1, affects `< 1.28.1`)

> CVE-2026-52870 CVE ↗ STABLE
>
> `CVE-2026-52870`: Experimental task handlers allow any client to access and cancel other clients' tasks (CVSS 7.6, affects `>= 1.23.0, <= 1.27.1`)

> CVE-2026-52869 CVE ↗ STABLE
>
> `CVE-2026-52869`: HTTP transports serve session requests without verifying the authenticated principal (CVSS 7.1, affects `<=1.27.1`)

> CVE-2025-66416 CVE ↗ STABLE
>
> `CVE-2025-66416`: DNS Rebinding Protection Disabled by Default in Model Context Protocol Python SDK for Servers Running on Localhost (CVSS 8.1, affects `<1.23.0`)

> CVE-2025-53366 CVE ↗ STABLE
>
> `CVE-2025-53366`: MCP SDK FastMCP Server Validation Error Leading to Denial of Service (affects `< 1.9.4`)

CVE-2025-53365 CVE ↗ STABLE

`CVE-2025-53365`: Unhandled Exception in Streamable HTTP Transport Leading to Denial of Service (affects `< 1.10.0`)

WAS THIS USEFUL? 👍 👎

agent-toolkit-for-aws Feedback ⓘ CVE-2026-87912 CVE ↗

Agent Toolkit for AWS gives AI coding agents the tools, knowledge, and guardrails they need to work with AWS services.

`CVE-2026-87912`: Missing S3 bucket ownership verification in the AWS Security Agent skills for aws-agents-for-devsecops (CVSS 5.9, affects `<=1.0.0`)

WAS THIS USEFUL? 👍 👎

Docker MCP Gateway Feedback ⓘ CVE-2026-76099 CVE ↗

STABLE

Docker MCP Gateway runs MCP servers as isolated containers and routes client requests through a Docker CLI plugin.

`CVE-2026-76099` : Unvalidated config-driven bind mount via dynamic tools (affects `>= 0.9.3`)

▾ **7 more releases in this issue** · 2025-12-03 → 2026-07-09

> CVE-2026-76097 CVE ↗ STABLE
>
> `CVE-2026-76097` : Tool-name shadowing across aggregated servers (affects `>=0.25.0`)

> CVE-2026-76095 CVE ↗ STABLE
>
> `CVE-2026-76095` : Arbitrary host file read via `file://` server reference (affects `>= 0.35.0`)

> CVE-2026-76094 CVE ↗ STABLE
>
> `CVE-2026-76094` : MCP server image signature verification off by default (affects `>=0.4.1`)

> CVE-2026-76093 CVE ↗ STABLE
>
> `CVE-2026-76093` : SSRF via attacker-controlled remote server URL (affects `>=0.13.0`)

> CVE-2026-76092 CVE ↗ STABLE
>
> `CVE-2026-76092` : Unauthenticated access to proxied tools in container mode (affects `>=0.25.0`)

> CVE-2026-55887 CVE ↗ STABLE
>
> `CVE-2026-55887` : Argument injection via OCI image label YAML in Docker MCP Gateway (affects `>= 0.21.0, < 0.42.2`)

> CVE-2025-64443 CVE ↗ STABLE

> `CVE-2025-64443`: DNS Rebinding vulnerability present when running Docker MCP Gateway in sse or streaming mode (CVSS 9.6, affects `<= v0.27.0`)

WAS THIS USEFUL? 👍 👎

windows-mcp

Feedback ⓘ `CVE-2026-48989` CVE ↗

MCP Server for Computer Use in Windows

`CVE-2026-48989`: HTTP transports expose unauthenticated PowerShell control with wildcard CORS (affects `< 0.7.5`)

WAS THIS USEFUL? 👍 👎

open-websearch

Feedback ⓘ `CVE-2026-42260` CVE ↗

Multi-engine MCP server, CLI, and local daemon for agent web search and content retrieval — skill-guided workflows, no API keys.

`CVE-2026-42260`: SSRF in `fetchWebContent` MCP tool: bracketed IPv6 literals and non-resolving hostname check bypass `isPrivateOrLocalHostname` (CVSS 8.2, affects <= 2.1.6 (current at time of report))

WAS THIS USEFUL? 👍 👎

excel-mcp-server

Feedback ⓘ `CVE-2026-40576` CVE ↗

Excel MCP server: let AI (Claude, Cursor, Codex) create and edit Excel files with charts, PivotTables, slicers and formulas. No Excel required.

`CVE-2026-40576`: Improper Limitation of a Pathname to a Restricted Directory (Path Traversal) in excel-mcp-server <=0.1.7 (CVSS 9.4, affects `<=0.1.7`)

WAS THIS USEFUL? 👍 👎

Prefect FastMCP

Feedback ⓘ CVE-2025-64340 CVE ↗ STABLE

FastMCP is a Python framework for building MCP servers and clients, handling transport, schema generation, authentication, and server composition.

`CVE-2025-64340` : Command injection via server name in subprocess-backed install commands (CVSS 6.7, affects `<3.2.0`)

▾ **4 more releases in this issue** · 2025-10-28 → 2026-03-31

> CVE-2026-27124 CVE ↗ STABLE
>
> `CVE-2026-27124` : Missing Consent Verification in OAuth Proxy Callback Facilitates Confused Deputy Vulnerabilities (CVSS 6.1, affects `<3.2.0`)

> CVE-2026-32871 CVE ↗ STABLE
>
> `CVE-2026-32871` : SSRF & Path Traversal Vulnerability in FastMCP OpenAPI Provider (CVSS 10.0, affects `< 3.2.0`)

> CVE-2025-69196 CVE ↗ STABLE
>
> `CVE-2025-69196` : Improper Resource Handling in FastMCP OAuth Proxy Enables Token Reuse Across MCP Servers (CVSS 6.5, affects `<2.14.2`)

> CVE-2025-62800 CVE ↗ STABLE
>
> `CVE-2025-62800` : Reflected XSS in client's callback page (CVSS 6.1, affects `<2.13.0`)

WAS THIS USEFUL? 👍 👎

Model Context Protocol Servers Feedback ⓘ

CVE-2026-27735 CVE ↗ STABLE

MCP Servers provides reference server implementations for connecting AI applications to external tools and data.

`CVE-2026-27735`: Path traversal in git_add allows staging files outside repository boundaries in mcp-server-git (CVSS 6.5, affects `< 2026.1.14`)

▾ **4 more releases in this issue** · 2025-07-01 → 2026-02-25

> CVE-2025-68145 CVE ↗ STABLE
>
> `CVE-2025-68145`: Missing path validation when using `--repository` flag in mcp-server-git (CVSS 9.1, affects `< 2025.12.18`)

> CVE-2025-68144 CVE ↗ STABLE
>
> `CVE-2025-68144`: Argument injection in git_diff and git_checkout functions for mcp-server-git allows overwriting local files (CVSS 7.1, affects `< 2025.12.18`)

> CVE-2025-68143 CVE ↗ STABLE
>
> `CVE-2025-68143`: Unrestricted git_init tool allows repository creation at arbitrary filesystem locations in mcp-server-git (CVSS 8.8, affects `< 2025.9.25`)

> CVE-2025-53110 CVE ↗ STABLE
>
> `CVE-2025-53110`: Path validation bypass via colliding path prefix (affects `< 0.6.3`, `< 2025.3.28`)

WAS THIS USEFUL? 👍 👎

figma-context-mcp

Feedback ⓘ CVE-2025-53967 CVE ↗

MCP server to provide Figma layout information to AI coding agents like Cursor

`CVE-2025-53967`: figma-developer-mcp vulnerable to command injection in get_figma_data tool (CVSS 7.5, affects `<=0.6.2`)

WAS THIS USEFUL? 👍 👎

DEPLOY

◆ AI / LLM Gateways

opencode-models-discovery

Feedback ⓘ v2.1.0 NOTES ↗

RANK

A universal OpenCode plugin for dynamic model discovery with flexible configuration for OpenAI-compatible providers.

> opencode-models-discovery added richer Ollama model metadata by querying an additional API endpoint.

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

01 **Enriched Ollama model metadata via /api/show** IMPROVED

55

The plugin now enriches Ollama model metadata by calling the `/api/show` endpoint, surfacing additional per-model detail beyond what the standard model list returns.

– Names the endpoint but no fields or example output v2.1.0

WAS THIS USEFUL? 👍 👎

dario

Feedback ⓘ v6.18.0 NOTES ↗ RANK

One local endpoint that routes Cursor, Cline, Aider, Claude Code and the Agent SDK through your Claude and ChatGPT subscriptions, with failover and pooling and zero runtime dependencies.

> dario's latest releases add a last-resort `credits` failover rung that routes drained seats to the Anthropic Messages API, a self-hosted analytics dashboard at `/analytics/ui`, and per-request cache-trace diagnostics for ChatGPT prompt-cache misses, alongside a configurable upstream timeout, plan-weighted pool routing, and a failover identity-sync fix.

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

01 Credits failover rung to Anthropic API NEW

95

Adds a `credits:<model>` rung to `--pool-fallback` (e.g. `opus:credits:claude-opus-5`) that routes drained or cooling requests to `api.anthropic.com` via the Messages API, authenticated by the `DARIO_CREDITS_API_KEY` environment variable; responses carry the header `x-dario-pool-fallback: credits:<model>` and are filed under a new `credits` bucket in `/analytics`, `/metrics`, and the dashboard. Status is surfaced in `dario doctor`'s Failover line and in the startup log when `DARIO_CREDITS_API_KEY` is absent.

Add a Max-plan API credits rung as the last resort in your failover chain so metered capacity is only consumed when all seats are drained.

```
$ DARIO_CREDITS_API_KEY=sk-ant-... dario --pool-fallback
'opus:credits:claude-opus-5'
```

— names flag, env var, header and endpoint with runnable example v6.17.0

02 Cache-trace diagnostics for ChatGPT cache misses NEW

93

Adds `--cache-trace` flag and `DARIO_CACHE_TRACE=1` environment variable to emit one stdout line per request explaining why a ChatGPT (codex) turn missed the prompt cache — gap in milliseconds, divergence point, and whether the seat, model, system prompt, or tool names changed. Adds a `cache` field to `/analytics` rows carrying prompt-cache-miss diagnostics (hashes and counts) per request, and matching `cache_*` fields to `--log-file` records.

Enable per-request cache-trace output to diagnose why ChatGPT turns are missing the prompt cache in a live session.

```
$ DARIO_CACHE_TRACE=1 dario --log-file ./dario.log
```

Use `--cache-trace` inline to see cache-miss reasons on stdout without setting an environment variable.

```
$ dario --cache-trace --log-file ./dario.log
```

— names exact flag, env var and fields with runnable commands v6.18.0

03 **Configurable upstream streaming timeout** NEW 90

Adds `--upstream-timeout=MS` flag (env `DARIO_UPSTREAM_TIMEOUT_MS`, default `300000`) to set the upstream streaming timeout, so long-running agent turns are no longer aborted mid-stream after 300 s.

Run dario with a 10-minute streaming timeout to prevent long agent turns from being cut off mid-response.

```
$ dario --upstream-timeout=600000
```

— exact flag, env var, default and runnable command given v6.16.0

04 **Plan-weighted pool routing strategy** NEW 90

Adds `--pool-strategy=headroom-by-plan` (env `DARIO_POOL_STRATEGY`, config key `pool.strategy`) to weight seat headroom by plan size, so a Max Nx seat counts N times toward routing decisions instead of once.

Enable plan-weighted headroom routing in a mixed Max/Pro pool so high-capacity seats absorb proportionally more traffic.

```yaml
pool:
  strategy: headroom-by-plan
```

— names flag, env var and config key with config example v6.16.0

05 **Expanded analytics API metrics** IMPROVED 80

Adds `window.perFiveMinutes` and `overage` fields to `/analytics`, giving the full hour in twelve bins for sparklines and surfacing halt state to read-only-token viewers. Adds `priced` to every row in `/analytics/ledger` as USD at the day's rate. Adds `totalTokenBurnRate` (all token kinds including cache reads and writes) to

`predictions` alongside the existing `tokenBurnRate` (input and output only), and adds `dario_burn_total_tokens_per_minute` to `/metrics` the same way.

Pull per-five-minute burn data and current overage halt state from the updated `/analytics` endpoint to drive your own dashboards or alerting.

```
$ curl -H 'Authorization: Bearer <analytics-token>'
https://<your-dario-host>/analytics | jq '{bins:
.window.perFiveMinutes, halted: .overage}'
```

— names every new field with a runnable curl example v6.17.0

06 **Self-hosted analytics dashboard** NEW 70

Adds a self-hosted analytics dashboard at `/analytics/ui` showing seats, headroom, reset timers, live request tail, timing split, spend per model and per consumer, key daily budgets, and an overage halt Resume button. Readable with the analytics token or an API key, in light or dark mode, on any screen size, with a table view for every chart.

— UI path given but no runnable command or screenshot v6.17.0

07 **Failover identity and session sync fix** IMPROVED 68

Failover between Claude pool seats now rebuilds `metadata.user_id` from the target seat's identity and syncs the `session_id` with the seat's `x-claude-code-session-id` header, preventing mismatched `device_id` and `account_uuid` on the fallback seat.

— internal mechanism named but no user-facing action v6.16.0

08 **Per-tool setup pages on landing site** NEW 60

Adds per-tool setup pages under `/tools/` on the landing site (Claude Code, Cursor, Aider, Cline, Codex CLI, the SDKs, and more), generated from `docs/tools.md` by `scripts/site-tools.mjs` at build time.

— names source files and path but no usage example v6.17.0

THINNER COVERAGE BELOW

09 **Pool-shared-state scope clarified** IMPROVED 35

Documents what `--pool-shared-state` covers (the lock and shared state) and what remains per-instance (ChatGPT seats, the ledger, the overage halt), and clarifies that routing does not read Redis on the request path.

— documentation clarification only, no new capability v6.18.0

WAS THIS USEFUL?  

BerriAI LiteLLM

Feedback ⓘ CVE-2026-114612 CVE ↗ STABLE

LiteLLM is an AI gateway that calls over 100 model APIs in OpenAI format with cost tracking, guardrails, load balancing, and logging.

`CVE-2026-114612`: Privilege Escalation to Proxy Admin via Cross-Domain Reuse of the Salt Key (CVSS 9.9, affects `>= 1.91.0, < 1.100.4`)

▾ **10 more releases in this issue** · 2026-04-20 → 2026-09-30

> CVE-2026-84377 CVE ↗ STABLE
>
> `CVE-2026-84377`: Authenticated SSRF and provider-credential exfiltration via unvalidated request-body routing parameters (CVSS 6.5, affects `<1.94.0`)

> CVE-2026-59823 CVE ↗ STABLE
>
> `CVE-2026-59823`: Server-side request forgery via the `user_config` request parameter in LiteLLM Proxy (affects `<=1.83.8`)

> CVE-2026-59821 CVE ↗ STABLE
>
> `CVE-2026-59821`: Custom Code Guardrails production endpoints bypass code safety checks (CVSS 7.2, affects `< 1.82.0-stable`)

> CVE-2026-59822 CVE ↗ STABLE
>
> `CVE-2026-59822`: MCP Authentication Bypass via OAuth2 Passthrough Fallback (CVSS 8.2, affects `< 1.84.0`)

> CVE-2026-59819 CVE ↗ STABLE
>
> `CVE-2026-59819`: Local file read via request-supplied OIDC file references (CVSS 4.9, affects `< 1.83.10-stable`)

> CVE-2026-59820 CVE ↗ STABLE
>
> `CVE-2026-59820`: Arbitrary file write via path traversal in Skills archive extraction (CVSS 6.5, affects `< 1.83.7-stable`)

> CVE-2026-40217 CVE ↗ STABLE
>
> `CVE-2026-40217`: Sandbox escape in custom-code guardrail (CVSS 8.8, affects `>= 1.81.8, < 1.83.10`)

CVE-2026-42271 CVE ↗ STABLE

`CVE-2026-42271` : Authenticated command execution via MCP stdio test endpoints (CVSS 8.8, affects `>=1.74.2, <1.83.7`)

CVE-2026-42208 CVE ↗ STABLE

`CVE-2026-42208` : SQL injection in Proxy API key verification (CVSS 9.8, affects `>=1.81.16, <1.83.7`)

CVE-2026-42203 CVE ↗ STABLE

`CVE-2026-42203` : Server-Side Template Injection in `/prompts`/test endpoint (CVSS 8.8, affects `>= 1.80.5, < 1.83.7`)

WAS THIS USEFUL? 👍 👎

OmniRoute

Feedback ⓘ CVE-2026-88062 CVE ↗

OmniRoute puts hundreds of model providers behind one endpoint, with quota-aware fallback and prompt compression.

`CVE-2026-88062` : OmniRoute ACP Custom-Agent Remote Code Execution (RCE) (affects `< 3.8.49`)

WAS THIS USEFUL? 👍 👎

◆ Agentic Infrastructure

HeyGen HyperFrames

Feedback ⓘ v0.8.207 NOTES ↗ RANK

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

> HyperFrames expanded CLI-to-app project handoff to Windows and Linux, added a CLI audio-loss check and new check-review reporting/frame-export subcommands, and introduced render encoding controls plus diagnostic and proxy tooling for the CLI.

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

01 CLI-to-app project handoff across platforms NEW

75

`hyperframes open` now hands the project to the app on Windows and Linux (previously limited to one platform); a new `hyperframes <dir>` shortcut opens a project folder directly in the HyperFrames app; and the CLI now prompts Windows users to download the HyperFrames app when running on that platform.

Jump straight into the HyperFrames app from the CLI when you want to visually edit an existing project.

```
$ hyperframes ./my-project
```

— Names two exact commands and includes a runnable example v0.8.205 v0.8.207

02 Check-review report attachment and frame export NEW

70

`check-review inspect` now attaches the inspector's report to each framed finding, and a new `check-review findingFrames` subcommand saves one plain frame per finding time.

— Names both subcommands but no usage example given v0.8.203

03 Env var to skip preview-server discovery NEW

70

The `HYPERFRAMES_NO_PREVIEW_DISCOVERY=1` environment variable skips the preview-server scan at CLI startup.

— Exact env var and value a reader can set today v0.8.207

04 **CLI check for lost or silenced audio tracks** NEW 65

A new CLI `check` command detects audio tracks the render would lose or play as silence, catching mix-down issues before a full render.

— Names the command and its purpose, no example shown v0.8.206

THINNER COVERAGE BELOW

05 **Diagnostic report for silent CLI commands** NEW 45

Writes one diagnostic report when a CLI command stays silent for 60 seconds, helping identify hangs.

— Concrete 60-second threshold but no command to invoke it v0.8.207

06 **Per-chunk timing logs on AWS Lambda** NEW 45

Logs per-chunk timing on AWS Lambda, showing where each chunk's seconds were spent.

— Names the platform but no log format or flag given v0.8.207

07 **Studio preview copy progress, retry and quality memory** IMPROVED 45

Studio preview copies now show copy progress, support retry, and remember their quality setting.

— Describes three UI behaviors with no exact navigation path v0.8.203

08 **Keyframe interval and peak bitrate cap for renders** NEW 40

Renders can now set a keyframe interval and a peak bitrate cap in the engine.

— Names the two controls but no flag or key given v0.8.207

09 **Proxy routing for CLI capture traffic** NEW 35

CLI capture traffic can now be routed through an opt-in upstream proxy.

— No flag or config key named for enabling it v0.8.207

10 **Loop count control for Studio GIF export** IMPROVED 30

Studio GIF export now supports specifying a loop count.

— Bare statement with no field name or UI path v0.8.205

11 **Hosted HyperFrames MCP listed in registry** NEW 30

The hosted HyperFrames MCP is now listed in the official MCP registry.

— Thin mention with no registry link or usage detail v0.8.207

WAS THIS USEFUL? 👍 👎

Tracer Cloud OpenSRE

Feedback ⓘ v0.1.2026.10.27 NOTES ↗

RANK

OpenSRE is an open-source toolkit for building AI SRE agents.

> OpenSRE's latest release locks shell access behind GitHub authentication and improves the onboarding and REPL configuration experience.

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

01 **GitHub connection gate before shell access** NEW 52

Shell access is now blocked behind a Connect GitHub screen until GitHub is connected, ensuring the prerequisite is in place before use. The hosted GitHub prerequisite error message is now surfaced directly in the LLM flow instead of a generic error or key prompt.

— Names the UI screen and error-flow change but no exact copy or flag v0.1.2026.10.27

02 **Editable REPL setup for remote sync** IMPROVED 40

Adds editable REPL setup to remote sync, while keeping configuration CLI flag-only.

— States the change but no concrete flag names or setup steps given v0.1.2026.10.27

03 **CI/CD automation leads startup menu** IMPROVED 26

The startup menu now leads with recommended CI/CD automation in the onboarding flow.

— Brief onboarding tweak with no specifics on options or commands v0.1.2026.10.27

WAS THIS USEFUL? 👍 👎

OpenAI TypeScript SDK

Feedback ⓘ v7.45.0 NOTES ↗

RANK

The OpenAI TypeScript SDK provides typed JavaScript and TypeScript clients for OpenAI APIs and streaming responses.

> This window adds a new `client.beta.agents.automations` namespace for delegation controls, prompt caching and scheduled automations, broadens Live Decision item types in a breaking way, retires `gpt-image-1` in favor of newer image models, and adds the `gpt-6.1-terra` model identifier.

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

01 Live Decision item types broadened (breaking)

BREAKING

90

`DecisionItemCreatedEvent.item` and `DecisionItemCreateEvent.item` are now unions including messages and function/custom-tool calls and outputs. Callers that directly access `item.content` or `item.role` may no longer compile; use if ('content' in event.item) before accessing message fields, and note message items may omit `type` so checking only `item.type === 'message'` can skip valid messages.

Safely handle the expanded Live Decision item union — guard against the new tool-call variants before accessing message fields.

typescript

```typescript
for (const event of decisionEvents) {
  if ('content' in event.item) {
    // message item
    console.log('Message content:', event.item.content);
  } else {
    // function / custom-tool call or output item
    console.log('Tool item:', event.item);
  }
}
```

— Exact type names plus runnable migration code v7.44.0

02 gpt-image-1 retired for newer image models

BREAKING

75

`gpt-image-1` was shut down on October 23, 2026; replace explicit selections with `gpt-image-2.5-sunburst` or `gpt-image-2.5-flare`. The documented default for Responses image-generation tools is now `gpt-image-2`.

— Exact replacement model names given, no code sample v7.44.0

03 **Agent automations namespace added** NEW 68

Adds `client.beta.agents.automations` with invocation endpoints and an `automations` agent-tool variant, supporting delegation controls, prompt cache lifetimes, and scheduled automations.

— Names the namespace and capabilities but gives no runnable call v7.44.0

THINNER COVERAGE BELOW

04 **gpt-6.1-terra model identifier added** NEW 52

Adds `gpt-6.1-terra` model identifier to the API type definitions, usable as the `model` parameter in Responses API calls.

Target the new gpt-6.1-terra model in a Responses API call.

typescript

```typescript
import OpenAI from 'openai';

const client = new OpenAI({
  apiKey: process.env['OPENAI_API_KEY'],
});

const response = await client.responses.create({
  model: 'gpt-6.1-terra',
  input: 'Explain zero-trust network architecture.',
});

console.log(response.output_text);
```

— Thin description but paired with runnable example v7.45.0

└──▷ **BREAKING ON UPGRADE**

! `DecisionItemCreatedEvent.item` and `DecisionItemCreateEvent.item` are now unions — callers that directly access `item.content` or `item.role` may no longer compile; use if (`content` in event.item) before accessing message fields and handle tool-item variants separately.

! Message items may omit `type`, so checking only `item.type === 'message'` can skip valid messages.

! `gpt-image-1` was shut down on October 23, 2026; replace explicit selections with `gpt-image-2.5-sunburst` or `gpt-image-2.5-flare`. The documented default for Responses image-generation tools is now `gpt-image-2`.

WAS THIS USEFUL? 👍 👎

OpenAI Python SDK

Feedback ⓘ v3.39.0 NOTES ↗ RANK

The OpenAI Python SDK provides typed Python clients for OpenAI model APIs, streaming, and asynchronous requests.

> OpenAI Python SDK v3.39.0 adds support for the new `gpt-6.1-terra` model identifier.

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

01 **gpt-6.1-terra model identifier support** NEW 55

The SDK adds the `gpt-6.1-terra` model identifier, which can be passed as the `model` parameter in calls such as client.chat.completions.create(model="gpt-6.1-terra", ...).

Target the new gpt-6.1-terra model in a chat completion request.

python

```python
from openai import OpenAI

client = OpenAI()

completion = client.chat.completions.create(
    model="gpt-6.1-terra",
    messages=[{"role": "user", "content": "Explain quantum
entanglement simply."}],
)
print(completion.choices[0].message.content)
```

– Names exact model id and usage, but no mechanism detail v3.39.0

WAS THIS USEFUL? 👍 👎

Hugging Face Transformers

Feedback ⓘ v5.21 NOTES ↗

RANK

Hugging Face Transformers is a model-definition framework for machine learning models in text, vision, audio, and multimodal tasks, for both inference and training.

> Transformers v5.21 adds five new models spanning OCR, TTS and mixture-of-experts architectures, introduces an FP8 backend for quantized KV caches and FP4 expert quantization, and ships breaking changes to tensor-parallel embedding handling and MoE expert sharding.

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

01 **FP8 backend for quantized KV cache** IMPROVED 86

`QuantizedCache` gains a new backend selectable via `cache_implementation="quantized"` and `cache_config={"backend": "fp8"}`, storing KV states in `torch.float8_e4m3fn` with a single scale per tensor, alongside the existing `quanto` and `hqq` backends.

Reduce KV-cache memory footprint during inference by storing keys and values in FP8 rather than FP16/BF16.

python

```python
from transformers import AutoModelForCausalLM, GenerationConfig

model = AutoModelForCausalLM.from_pretrained("meta-llama/Meta-Llama-3-8B-Instruct")

output = model.generate(
    input_ids,
    cache_implementation="quantized",
    cache_config={"backend": "fp8"},
    max_new_tokens=512,
)
```

— Exact config flags with a runnable generate() example v5.21

02 **FP4 expert quantization for MoE models** NEW 81

`FineGrainedFP8Config` gains an `expert_dtype` option accepting `"fp8"` (default) or `"fp4"`, letting DeepSeek-V4.1's FP4 experts be quantized correctly while other layers stay in FP8.

Run DeepSeek-V4.1 with FP4 experts to cut expert-weight memory while keeping other layers in FP8.

```python
from transformers import AutoModelForCausalLM,
FineGrainedFP8Config

quant_config = FineGrainedFP8Config(expert_dtype="fp4")
model = AutoModelForCausalLM.from_pretrained(
    "deepseek-ai/DeepSeek-V4.1",
    quantization_config=quant_config,
    device_map="auto",
)
```

— Exact config key and values with runnable example v5.21

03 **MoE tensor/expert parallel plans now shard experts** BREAKING

78

Every MoE TP and EP plan now shards its experts across ranks by default; custom or downstream MoE parallel plans must add expert sharding or they will fail the new `test_moe_parallel_plans_shard_experts` test and waste memory keeping all experts on every rank.

— Names the exact test and consequence of not migrating v5.21

04 **Tensor-parallel embedding sharding changes** BREAKING

75

Tensor parallelism now uses a dedicated embedding style and shards input embeddings by default; `ColwiseParallel` and `RowwiseParallel` now apply only to linear layers, so custom TP plans that applied them to embeddings must switch to the new embedding style.

— Names exact classes and required migration step v5.21

05 **Pocket TTS CPU-capable voice cloning model** NEW

73

Adds support for Pocket TTS, Kyutai's 100M-parameter CPU-capable TTS model with voice cloning from recordings, precomputed latents, or cached KV states, plus `PocketTTSAudioTokenizer`, a continuous 32-dimensional VAE-GAN codec usable as the `audio_tokenizer` of `PocketTTSProcessor` or standalone.

— Named classes and usage pattern but no runnable example v5.21

06 **UltraVBert visual document understanding model** NEW

65

Adds support for UltraVBert, a visual document understanding and retrieval model combining an UltraBERT text encoder with a SigLIP vision encoder, with a

`max_image_tiles` setting (default 27) in its image processor capping the number of 512-pixel tiles per page.

— Named config key but no usage example given v5.21

07 **Flint mixture-of-experts language model** NEW 60

Adds support for Flint, Recursive's mixture-of-experts LM with hybrid sliding-window and full attention, learned attention sinks, and sigmoid-routed sparse experts (6 of 256) plus a shared dense expert.

.

— Rich architecture detail but no runnable example v5.21

THINNER COVERAGE BELOW

08 **Gradium Phonon on-device TTS model** NEW 55

Adds support for Gradium Phonon TTS, an on-device TTS model emitting 32-dimensional audio latents at 12.5 Hz decoded to 24 kHz audio via a one-step flow head.

— Specific technical parameters but no usage example v5.21

09 **New accessor methods for linear-attention caches** IMPROVED 55

Adds `is_recording_past`, `get_recurrent_state`, and conv-state getters to linear-attention caches, replacing direct internal-state access.

— Names exact methods but no usage example v5.21

10 **MiniCPM-OCR document understanding model** NEW 40

Adds support for MiniCPM-OCR, a document-understanding model pairing a SigLIP vision encoder with a Qwen3.5 LM, specialized for text-dense images such as scanned pages, screenshots, receipts, and tables.

— Describes architecture but no config or example v5.21

11 **Tensor-parallel plan for Nemotron models** NEW 25

Adds a tensor-parallel plan for Nemotron models.

— One-line addition with no mechanism detail v5.21

└──▷ **BREAKING ON UPGRADE**

! Tensor parallelism now uses a dedicated embedding style and shards input embeddings by default; `ColwiseParallel` and `RowwiseParallel` apply only to linear layers — custom TP plans that applied them to embeddings must switch to the new embedding style.

! Every MoE TP and EP plan now shards its experts across ranks; custom or downstream MoE parallel plans must add expert sharding or models will fail the new `test_moe_parallel_plans_shard_experts` test and waste memory keeping all experts on every rank.

WAS THIS USEFUL? 👍 👎

Google Gemini Enterprise Agent Platform

Feedback ⓘ October 26, 2026 (2026-10-26) BLOG ↗ RANK

Gemini Enterprise Agent Platform provides Google's managed model training, inference, evaluation, and agent deployment services, formerly offered under Vertex AI.

> Gemini Enterprise Agent Platform's Workbench now generally supports custom organization policies for controlling instance creation and updates.

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

01 **Custom org policies for Workbench instances** NEW 65

Agent Platform Workbench now supports custom organization policy for `notebooks.googleapis.com/Instance` resources, enabling allow/deny controls on instance creation and updates, with dry-run mode support. This capability is now generally available (GA).

— Names resource type and dry-run mode but no config example

October 26, 2026 (2026-10-26)

WAS THIS USEFUL? 👍 👎

skillhub

Feedback ⓘ dsh-v0.4.0 NOTES ↗ RANK

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 shipped a breaking install-inventory v2 migration spanning DSH 0.4.0 and CLI 0.2.0 with per-location isolation, alongside new in-conversation skill lifecycle management in DSH, CLI support for skill suites, OAuth Device Flow login, and DSH/ZCode/Pi agent integrations, plus server-side additions including public GitHub Skill discovery with a review queue, Silo/RustFS storage overlays, a `downloadableOnly` version-list filter, and preserved `.skillhub/examples.json` authoring.

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

01 downloadableOnly filter on version-list API NEW

86

Adds a `downloadableOnly=true` filter to the version-list API, returning only publicly available versions in publication-time order before pagination, while preserving the default management-history view for callers that omit the parameter.

Fetch only publicly downloadable Skill versions in publication-time order — useful in a marketplace or CI context where draft/unpublished versions should be hidden.

```
$ curl 'https://<skillhub-host>/api/skills/<skill-id>/versions?downloadableOnly=true'
```

— Named endpoint, parameter, and runnable curl example given v0.2.24

02 Public GitHub Skill discovery and sync worker NEW

83

Adds public GitHub repository discovery and scheduled synchronization via an optional Go worker with a separate service credential; administrators configure admission limits and enable discovery explicitly, and imported packages go through the normal scanning and review queue. Before upgrading, resolve any existing `github` namespace or `github-sync-publisher` account collision; the Server applies Flyway migrations `V76` – `V119`, and database rollback requires restoring the matching pre-upgrade PostgreSQL and local/object storage backup. The optional GitHub worker was updated to Go 1.27.3 and must be upgraded together with Java before enabling collection.

— Names migrations and worker version but no exact command v0.2.24

03 Install inventory v2 with per-location isolation BREAKING

75

The shared install inventory format is upgraded to v2 across DSH 0.4.0 and CLI 0.2.0: existing v1 data becomes read-only and is not auto-migrated, so v1 must be backed up before the first write. Users on old CLI versions must upgrade to CLI 0.2.0 before using DSH 0.4.0, and mixed writes from old and new consumers are unsupported. CLI 0.2.0 adds per-location inventory isolation with upgrade isolation and rollback protection, so installs at different locations are managed and restored independently; rolling back requires restoring both the inventory backup and all subsequently modified skill files and operation records — restoring the inventory backup alone is not sufficient.

— Explains migration mechanism and rollback steps, no exact command `cli-v0.2.0` `dsh-v0.4.0`

---

04 **Authored examples.json preserved through publish/download**

70

BREAKING

Preserves authored `.skillhub/examples.json` files through package publishing and downloading, and reports recognized example status and counts in publishing preflight without rejecting an otherwise valid package for an unrecognized example format. Packages containing `.skillhub/examples.json` require CLI and DSH consumers that support the new authored-content and fingerprint rules — update consumers before distributing these packages, since existing published fingerprints are not recalculated.

— Names the exact file and fingerprint caveat, no command shown `v0.2.24`

---

05 **Silo and RustFS storage deployment overlays** NEW

62

Adds optional Silo and RustFS single-node deployment overlays with storage diagnostics and separate internal upload and public download endpoints; existing local and external S3-compatible configurations are unaffected.

— Names overlay options and endpoint separation, no config keys given `v0.2.24`

THINNER COVERAGE BELOW

06 **Conversational skill usage and lifecycle in DSH** NEW

50

DSH 0.4.0 adds in-conversation skill usage with multilingual case selection, filling drafts without auto-sending, plus conversational querying, update checking, upgrade confirmation, and uninstallation of skills, with version management scoped by install location.

— Describes behavior but no commands or config surfaces given `dsh-v0.4.0`

---

07 **Skill suite install and management in CLI** NEW

40

CLI 0.2.0 adds skill suite install and management support, requiring a Registry that exposes the Suite API.

— Names the Suite API requirement but no commands shown cli-v0.2.0

08 **OAuth Device Flow login for CLI** NEW 40

CLI 0.2.0 adds OAuth Device Flow login for authenticating clients that cannot host a browser redirect.

— Names the auth mechanism but no flag or command given cli-v0.2.0

09 **Agent client integrations for DSH, ZCode, and Pi** NEW 38

CLI 0.2.0 adds client integrations for the DSH agent, ZCode agent, and Pi agent, alongside bundled usage examples and an install core shipped with the package.

— Lists integrated agents but no setup detail cli-v0.2.0

└──▷ **BREAKING ON UPGRADE**

! The shared install inventory is upgraded to v2: existing v1 data is read-only and will not be migrated — back up v1 before the first write. Users on the old CLI must upgrade to CLI 0.2.0 before using DSH 0.4.0; mixed writes from old consumers are not supported. Rolling back requires restoring both the inventory backup AND all subsequently modified skill files and operation records — restoring the inventory backup alone is not sufficient.

! The shared install inventory format is upgraded to v2: existing v1 inventory is read-only and not auto-migrated — a backup of v1 is taken before the first write. Users on old CLI versions must upgrade to CLI 0.2.0 before using DSH 0.4.0; mixed writes from old and new consumers are not supported. Rolling back by restoring only the inventory backup is unsupported if skill files or operation records were modified after upgrade.

! Resolve any existing `github` namespace or `github-sync-publisher` account collision before upgrading; database rollback requires restoring the matching pre-upgrade PostgreSQL and local/object storage backup. The Server applies Flyway migrations V76–V119.

! Packages containing `.skillhub/examples.json` require CLI and DSH consumers that support the new authored-content and fingerprint rules. Update consumers before distributing these packages; existing published fingerprints are not recalculated.

! Upgrade Java and the optional GitHub worker together before enabling collection; the Go worker was updated to Go 1.27.3.

WAS THIS USEFUL?  

AWS Amazon Bedrock AgentCore

Feedback

AWS Agent Registry now available in US East (Ohio), Europe (Frankfurt), Europe (London),

BLOG ↗

Amazon Bedrock AgentCore provides managed agent runtimes, memory, identity, gateways, observability, and evaluation services.

NEW SUPPORTED PLATFORM AWS Agent Registry regional expansion

WAS THIS USEFUL?

Anyscale Ray

Feedback ⓘ CVE-2026-57516 CVE ↗ STABLE

Ray is a distributed compute engine for AI workloads, bundling Ray Train, Ray Serve, Ray Data, and Ray Tune on a common runtime.

`CVE-2026-57516`: Arbitrary code execution via `ray.data.read` _webdataset default decoder: pickle.loads(value) and torch.load(weights_only=False) (CVSS 8.8, affects `< 2.56.0`)

▾ **3 more releases in this issue** · 2025-11-26 → 2026-07-01

> CVE-2026-41486 CVE ↗ STABLE
>
> `CVE-2026-41486`: Remote Code Execution via Parquet Arrow Extension Type Deserialization (CVSS 8.8, affects `2.54.0`)

> CVE-2026-27482 CVE ↗ STABLE
>
> `CVE-2026-27482`: Dashboard DELETE endpoints allow unauthenticated browser-triggered DoS (Serve shutdown / job deletion) (CVSS 5.9, affects `< 2.54.0`)

> CVE-2025-62593 CVE ↗ STABLE
>
> `CVE-2025-62593`: Critical RCE Vulnerability against Ray Devs exploitable via Browser (Safari & Firefox) due to DNS Rebinding Attack (CVSS 8.8, affects `< 2.52.0`)

WAS THIS USEFUL? 👍 👎

Hugging Face Diffusers

Feedback ⓘ CVE-2026-45804 CVE ↗

Diffusers provides pretrained diffusion pipelines, schedulers, and models for image, video, and audio generation in PyTorch.

`CVE-2026-45804`: TOCTOU Trust Remote Code Bypass (CVSS 7.5, affects `<0.37.1`)

WAS THIS USEFUL? 👍 👎

Anthropic TypeScript SDK

Feedback ⓘ CVE-2026-41686

CVE ↗ STABLE

The Anthropic TypeScript SDK provides typed JavaScript and TypeScript clients for Claude APIs, streaming, and tool use.

`CVE-2026-41686`: Claude SDK for TypeScript has Insecure Default File Permissions in Local Filesystem Memory Tool (CVSS 4.4, affects `>= 0.79.0`)

▾ **1 more release in this issue** · 2026-03-31 → 2026-04-24

> CVE-2026-34451 CVE ↗ STABLE
>
> `CVE-2026-34451`: Memory Tool Path Validation Allows Sandbox Escape to Sibling Directories (CVSS 5.4, affects `>= 0.79.0, < 0.81.0`)

WAS THIS USEFUL? 👍 👎

Anthropic Python SDK

Feedback ⓘ CVE-2026-34452 CVE ↗

STABLE

The Anthropic Python SDK provides typed Python clients for Claude APIs, streaming responses, and tool use.

`CVE-2026-34452`: Memory Tool Path Validation Race Condition Allows Sandbox Escape (CVSS 5.3, affects `>= 0.86.0, < 0.87.0`)

▾ **1 more release in this issue** · 2026-03-31

> CVE-2026-34450 CVE ↗ STABLE
>
> `CVE-2026-34450`: Insecure Default File Permissions in Local Filesystem Memory Tool (CVSS 4.4, affects `>= 0.86.0, < 0.87.0`)

WAS THIS USEFUL? 👍 👎

NVIDIA NeMo Speech

Feedback ⓘ CVE-2024-0081 CVE ↗

STABLE

NeMo Speech provides model training and inference tools for speech recognition, speech synthesis, audio processing, and speech language models.

`CVE-2024-0081`: Unicode use in a user-controlled filename may cause a server-side DoS (CVSS 8.6, affects `NVIDIA Neural Modules 1.22.0`)

▾ **1 more release in this issue** · 2022-01-06 → 2024-03-27

> CVE-2022-22821 CVE ↗ STABLE
>
> `CVE-2022-22821`: Arbitrary file deletion in NeMo ASR webapp (CVSS 4.4, affects `<=1.5.1`)

WAS THIS USEFUL? 👍 👎

◆ Local LLM Runtimes

Comfy Org ComfyUI

Feedback ⓘ v0.43.0 NOTES ↗ RANK

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

> ComfyUI v0.43.0 ships a new Branch node for conditional graph routing, a Pad Image node, ACEScg color space support, MOSS-Image editing upgrades, a wave of new partner model integrations (HeyGen Voice, ClarityAI, Fibo 2.0, Kling 4.0, Mirelo SFX and more), plus AMD and Qwen 3.x performance improvements.

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

01 ACEScg color space support NEW

70

Adds ACEScg color space option to the Convert Image Color Space node and the Save Image (Advanced) EXR output, enabling compositing-ready EXR exports in ACEScg for VFX pipelines.

Save a compositing-ready EXR in ACEScg color space for downstream VFX pipelines.

📍 In the graph editor, add a Save Image (Advanced) node, set format to EXR, and select ACEScg from the color space dropdown.

— Names both nodes and the color space; UI steps given but not a command. v0.43.0

02 Pad Image node with mirror padding NEW

65

Adds a Pad Image node supporting mirror padding and pad-to-multiple options, useful for preparing images for tiled upscaling without border artifacts.

Pad an image to the nearest multiple of a tile size with mirror padding before feeding it to a tiled upscaler to avoid border artifacts.

📍 In the graph editor, right-click › Add Node › search Pad Image › connect your image output, set pad mode to mirror and specify the target multiple.

— Names node and both options; usage steps given but no exact command. v0.43.0

03 MOSS-Image editing and merge enhancements IMPROVED

60

Adds a MOSS-Image model-specific merge node, and improves MOSS-Image with edit LoRA support including attention bias conditioning and image strength control, plus adds color-depth control and rope scaling support.

— Lists several named capabilities but no usage path or command. v0.43.0

THINNER COVERAGE BELOW

04 **Branch node for conditional workflow routing** NEW 55

Adds a Branch node to the graph editor for conditional workflow routing, letting a workflow follow different node paths based on a condition.

Use the new Branch node to route a workflow down one of two paths based on a condition, enabling dynamic graph control without duplicating nodes.

📍 In the graph editor, right-click › Add Node › search Branch › connect a boolean condition input and wire each output to the appropriate downstream node path.

— Names the node but not its internal logic or condition types. v0.43.0

05 **New partner model integrations across generation types** NEW 55

Adds a wave of new partner model integrations: Fibo 2.0 image generation and editing support, HeyGen Voice instant clone and text-to-speech partner nodes, ClarityAI Crystal Video Upscaler partner node, FLUX 3 Max partner nodes, Bria FIBO 2.0 Image Generate/Edit partner nodes plus 3 additional Bria nodes, VEED Fabric 1.0, Lipsync 2.0, and additional VEED partner nodes, Mirelo SFX partner nodes for sound-effect generation, Meshy T2.5 Smart Topology partner nodes, PixVerse V6-Max model partner node, Ideogram Vector Generate SVG partner nodes, Kling 4.0 partner nodes, FLUX 3 Video nodes with 2K and 4K resolution options for video generation, and a Krea 3 Image node with an ultra thinking mode option.

— Names every integration but gives no usage mechanism for any. Names resolution options but no further mechanism or usage path. Names the node and mode but no explanation of behaviour. v0.43.0

06 **PyTorch attention enabled in VAE on AMD GPUs** IMPROVED 35

Enables PyTorch attention in the VAE for AMD GPUs, improving performance on AMD hardware.

— States the change but no benchmark or scope detail. v0.43.0

07 **Adaptive tree MTP for Qwen 3.x performance** IMPROVED 35

Applies Adaptive tree MTP optimizations for Qwen 3.x models to improve performance.

— Names the technique and model family but no measured impact. v0.43.0

CVE-2026-56673 CVE ↗ STABLE

`CVE-2026-56673`: Path traversal in LoadImage via the `/prompt` API allows arbitrary file existence probing and image exfiltration (CVSS 7.5, affects `< 0.28.0`)

CVE-2026-56672 CVE ↗ STABLE

`CVE-2026-56672`: Stored XSS via `/userdata`/{file} due to missing Content-Type sanitization (CVSS 8.2, affects `< 0.28.0`)

CVE-2026-56670 CVE ↗ STABLE

`CVE-2026-56670`: Stored XSS via SVG file upload on the `/view` endpoint (CVSS 8.2, affects `< 0.28.0`)

WAS THIS USEFUL? 👍 👎

llamastash Feedback ⓘ v0.10.0 NOTES ↗ RANK

Zero-overhead, terminal-native local-LLM orchestrator. Run any engine, any model via the same proxy.

> LlamaStash v0.10.0 extends the proxy beyond chat with STT/TTS/image mode routing and raw engine route forwarding, adds alias-state diagnostics, per-request KV-cache tracking, a decision-model route, and a much larger recipes library.

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

01 **STT, TTS, and image modes for proxy routing** NEW 95

Adds `modes: [stt]`, `modes: [tts]`, and `modes: [image]` to `backend.generic.servers` entries in `config.yaml`, routing OpenAI audio and image routes (`/v1/audio/transcriptions`, `/v1/audio/translations`, `/v1/audio/speech`, `/v1/images/generations`, `/v1/images/edits`, `/v1/images/variations`) on the proxy port `:11435` to local engines such as whisper.cpp, Kokoro, and sd.cpp. Media requests routed by mode fallback now carry response headers `x-llamastash-served-by` and `x-llamastash-fallback-reason: mode_match`, and the same reason is logged in the request row so substitutions are never silent.

Route a Whisper transcription request through the proxy to a local whisper.cpp server declared with `modes: [stt]` — no client-side changes needed.

```
$ curl -s localhost:11435/v1/audio/transcriptions -F
model=whisper-1 -F file=@audio.wav
```

Declare a stable-diffusion.cpp server with `modes: [image]` and expose its A1111 API alongside OpenAI image routes using `routes:`.

```yaml
backend:
  generic:
    servers:
      - name: moss-image
        binary: /path/to/stable-diffusion.cpp/build-hip/bin/sd-
server
        args: [--diffusion-model, /path/to/MOSS-Image-SFT-
Q8_0.gguf, --listen-ip, "{host}", --listen-port, "{port}"]
        ready: /v1/models
        modes: [image]
        routes: [/sdapi/v1]
        memory_gib: 14
```

— Names exact endpoints, port, engines and headers with a runnable curl example. `v0.10.0`

02 **Alias state reporting in `llamastash list --json`** NEW

86

Adds a top-level `aliases` key to `llamastash list --json` output reporting every `proxy.aliases` entry with its `state` (`ok`, `shadowed`, `dead`, `ambiguous`, or `chain`) and the reason a dead alias cannot route.

Audit all proxy aliases for broken routing before a deployment — a `dead` or `ambiguous` state here is exactly what a live client would receive.

```
$ llamastash list --json | jq '.aliases'
```

— Exact flag, output schema and a runnable audit command. `v0.10.0`

03 **Custom route forwarding via `routes:` key** NEW

75

Adds `routes:` to `backend.generic.servers` entries in `config.yaml`, forwarding an engine's own path prefixes (e.g. `/sdapi/v1`, `/inference`) through the proxy alongside `modes` `-covered` routes, with full query-string and body pass-through.

Declare a stable-diffusion.cpp server with `modes: [image]` and expose its A1111 API alongside OpenAI image routes using `routes:`.

```yaml
backend:
  generic:
    servers:
      - name: moss-image
        binary: /path/to/stable-diffusion.cpp/build-hip/bin/sd-server
        args: [--diffusion-model, /path/to/MOSS-Image-SFT-Q8_0.gguf, --listen-ip, "{host}", --listen-port, "{port}"]
        ready: /v1/models
        modes: [image]
        routes: [/sdapi/v1]
        memory_gib: 14
```

— Named config key with example but mechanism briefly described. v0.10.0

04 **`rewrite_model` config key for vLLM backend** NEW 65

Adds `backend.vllm.rewrite_model` config key (default `false`) to stamp the vLLM launch's served name into outbound requests, matching the behaviour of a generic entry's `rewrite_model`.

— Named config key and default, but no usage example given. v0.10.0

05 **Per-request KV-cache hit ratio tracking** NEW 65

Adds a `CACHE` column (and `cache_hit_ratio` field on row JSON) to the TUI Requests tab and `llamastash requests` output, showing each individual request's KV-cache hit ratio as `cached_tokens` over `prompt_tokens`.

— Describes the metric and surfaces but only points to a UI tab. v0.10.0

06 **Expanded RECIPES.md with 11 tested recipes** NEW 60

Expands `RECIPES.md` into five groups (Chat, Image generation, Speech to text, Text to speech, Classifiers and decision models) with 11 new tested recipes covering stable-diffusion.cpp, ComfyUI, LocalAI, Speaches, Fish Speech S2-Pro, Piper, Kokoro-FastAPI, Chatterbox TTS API, Kev, Laya, LANCET, and FastFlowLM Whisper on the NPU.

— Lists specific engines covered but is documentation, not a config surface. v0.10.0

THINNER COVERAGE BELOW

07 **`/v1/systemone` route for decision models** NEW 55

Adds `/v1/systemone` to the proxy as a forwarded route resolved by model name, enabling decision models (e.g. Kev-0.8B) to answer on the proxy port.

— Names the route and an example model but lacks a usage example. v0.10.0

WAS THIS USEFUL?  

Mintplex Labs AnythingLLM

Feedback ⓘ v1.19.0

NOTES ↗ RANK

AnythingLLM is a local-first desktop and server application that runs agents, RAG workspaces, and MCP tools against local or hosted models.

> AnythingLLM v1.19.0 adds a built-in Lemonade/llama.cpp LLM engine for Docker deployments so no external Ollama or LM Studio is needed, alongside new UI color palettes and a batch of smaller workspace, transcription, and localization improvements.

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

01 **Built-in Lemonade/llama.cpp LLM engine** NEW 93

Set `ENABLE_BUILT_IN_LLM=true` in the Docker container to activate a built-in local inference engine (use `--gpus all` for GPU acceleration), removing the need for an external Ollama or LM Studio instance. It auto-selects the fastest backend — CUDA for NVIDIA GPUs, Vulkan for AMD and Intel GPUs, or CPU fallback — on both `amd64` and `arm64` hosts, and is available as a first-class provider for the system LLM, workspace and agent LLMs, the chat window model picker, and the Model Router.

Run AnythingLLM Docker with the built-in LLM enabled and GPU acceleration so no external inference server is needed.

```
$ docker run -e ENABLE_BUILT_IN_LLM=true --gpus all -p 3001:3001 mintplexlabs/anythingllm
```

— Names exact env var, flags, backends, and integration points v1.19.0

THINNER COVERAGE BELOW

02 **Instance color palettes for branding** NEW 45

Adds instance-level color palettes (Ocean, Forest, Plum, Ember) under tabbed branding settings for full-UI thematic tinting.

— Names palettes and settings location but no deeper mechanism v1.19.0

03 **Mid-thread workspace model switching for agents** IMPROVED 42

Applies workspace model changes to agent sessions mid-thread without requiring a restart.

— States before/after behaviour but no config key named v1.19.0

04 **Max output token limit notice in chat** IMPROVED 35

Shows a notice in the chat window when a response hits the max output token limit.

— Describes behaviour but no config or threshold details v1.19.0

05 **PDF and file link parsing in bulk scraper** IMPROVED 34

Parses PDF and file links in the bulk link scraper.

— Names surface but gives no further mechanism v1.19.0

06 **Per-file embed progress in context overflow modal** IMPROVED 34

Shows per-file embed progress in the context overflow modal.

— Names UI surface but lacks further mechanism detail v1.19.0

07 **.mov file transcription support** NEW 30

Adds `.mov` file support for transcription.

— Bare feature name, no detail on scope or mechanism v1.19.0

08 **Hungarian UI translation** NEW 26

Adds Hungarian (`hu`) translation to the UI.

— Bare localization addition with language code v1.19.0

> CVE-2026-88055 CVE ↗ STABLE
>
> `CVE-2026-88055` : Stored XSS Due to Unescaped Server-Side HTML Concatenation in MetaGenerator (CVSS 5.5, affects `<=1.16.1`)

> CVE-2026-72917 CVE ↗ STABLE
>
> `CVE-2026-72917` : Password recovery accepts one recovery code twice after whitespace normalization (CVSS 5.9, affects `>= 1.0.0, <= 1.15.0`)

> CVE-2026-48789 CVE ↗ STABLE
>
> `CVE-2026-48789` : Windows path containment bypass in document folder route (CVSS 4.3, affects `unknown - reproduced on latest tested version`)

> CVE-2026-48116 CVE ↗ STABLE

`CVE-2026-48116` : RCE via ripgrep `--pre` argument injection in filesystem-search-files agent skill (CVSS 7.5, affects `<= 1.12.1`)

`CVE-2026-42456` `CVE ↗` `STABLE`

`CVE-2026-42456` : ## Cross-User TTS Audio Disclosure via Chat ID (IDOR) (CVSS 4.3, affects `<= 1.11.1`)

`CVE-2026-41318` `CVE ↗` `STABLE`

`CVE-2026-41318` : Stored DOM XSS in chart caption renderer - LLM-driven prompt injection produces executable HTML via unsanitized renderMarkdown(content.caption) in Chartable component (CVSS 5.4, affects `<= 1.12.0`)

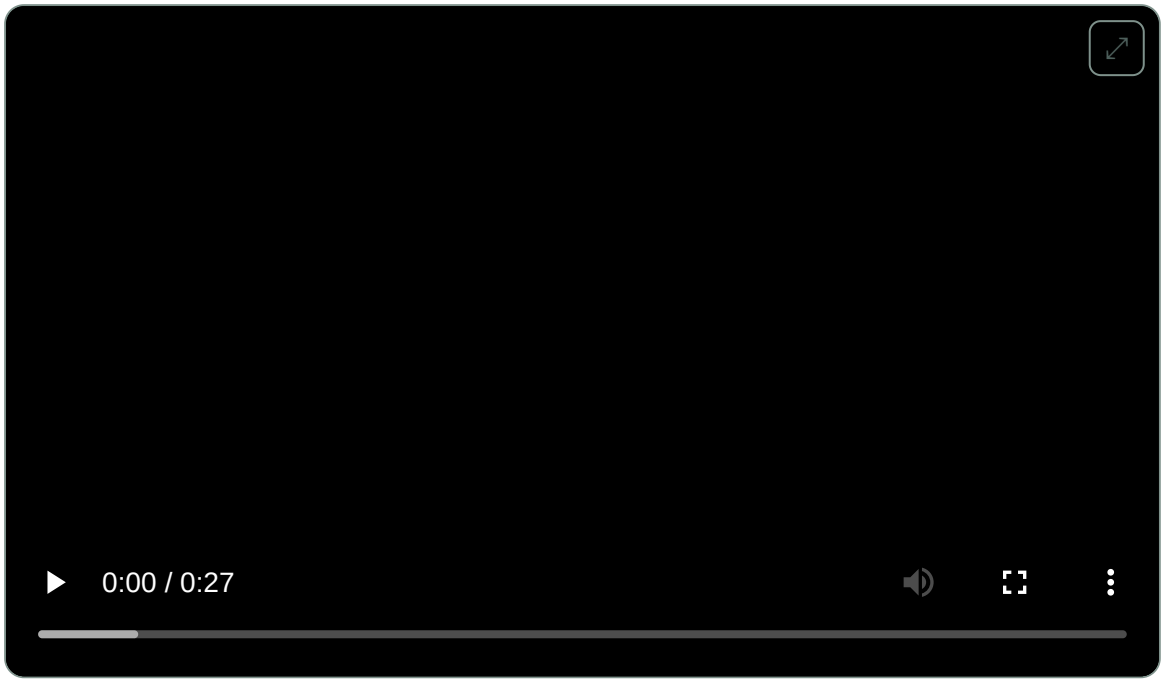
`CVE-2026-32719` `CVE ↗` `STABLE`

`CVE-2026-32719` : Zip Slip Path Traversal and Code Execution via Community Hub Plugin Import (CVSS 4.2, affects `<= 1.6.7`)

`CVE-2026-32628` `CVE ↗` `STABLE`

`CVE-2026-32628` : SQL Injection in Built-in SQL Agent Plugin via Unsanitized table_name Parameter (CVSS 8.8, affects `≤ 1.11.1`)

`CVE-2026-32626` `CVE ↗` `STABLE`

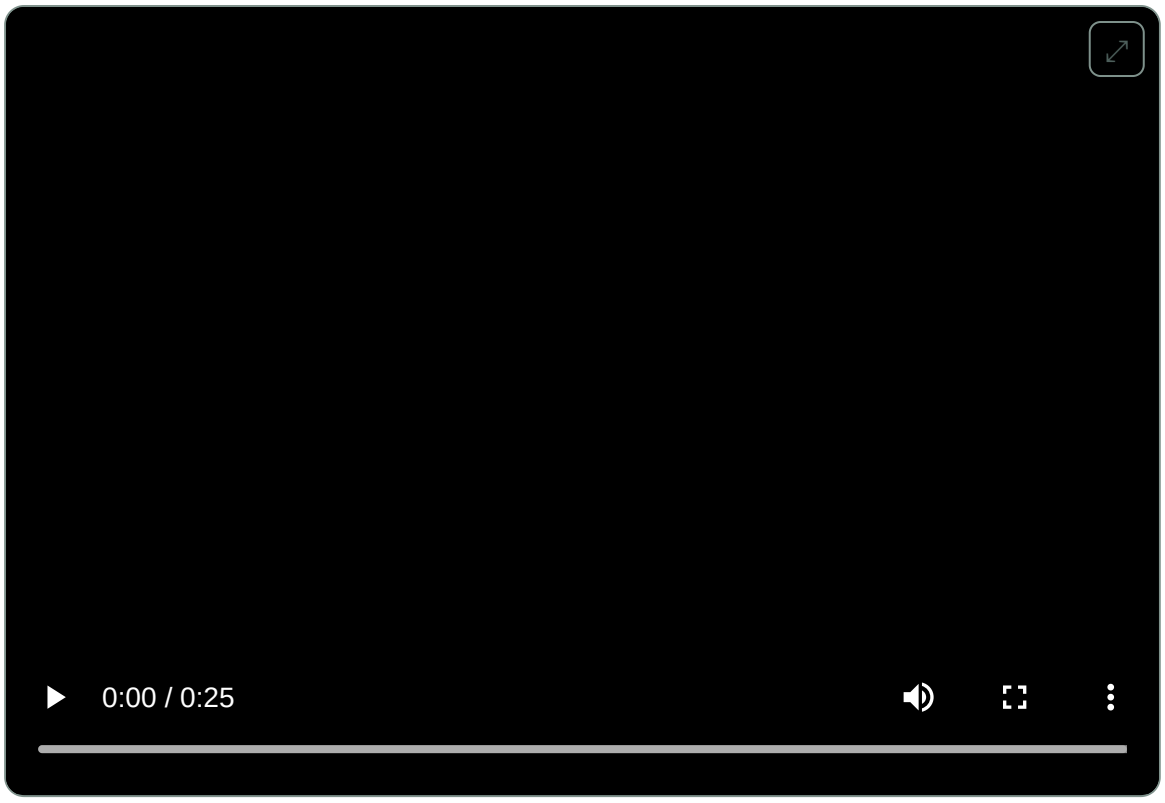


`CVE-2026-32626` : Streaming Phase XSS to RCE via LLM Response Injection (CVSS 9.6, affects `<=1.11.1`)

CVE-2026-32617 CVE ↗ STABLE

`CVE-2026-32617` : Permissable CORS policy (CVSS 7.1, affects `1.11.1`)

CVE-2026-24478 CVE ↗ STABLE



`CVE-2026-24478` : Improper Limitation of a Pathname to a Restricted Directory (Path Traversal) in anything-llm (CVSS 7.2, affects All versions including latest (as of report date))

CVE-2026-24477 CVE ↗ STABLE

`CVE-2026-24477` : Key leak in `systemSettings.js` (CVSS 7.5, affects `<1.10.0`)

CVE-2026-21484 CVE ↗ STABLE

`CVE-2026-21484` : Username Enumeration w/ Password Recovery (CVSS 5.3, affects `<= latest`)

CVE-2024-22422 CVE ↗ STABLE

`CVE-2024-22422` : Unauthenticated Denial of Service (DOS) attack in File Exports (CVSS 7.5, affects `latest`)

WAS THIS USEFUL? 👍 👎

Open WebUI

Feedback ⓘ CVE-2026-88006 CVE ↗ STABLE

Open WebUI is a self-hosted chat interface for local and remote models, with RAG over uploaded documents, tool calling, and multi-user access control.

`CVE-2026-88006` : Users denied by the OAuth role policy can still sign in via token exchange (CVSS 6.5, affects `>= 0.8.0, < 0.11.1`)

▾ **11 more releases in this issue** · 2026-08-30 → 2026-09-09

> CVE-2026-87012 CVE ↗ STABLE
>
> `CVE-2026-87012` : Any authenticated user can suppress calendar alerts instance-wide via a non-numeric alert value (CVSS 4.3, affects `>= 0.9.0, < 0.11.1`)

> CVE-2026-87013 CVE ↗ STABLE
>
> `CVE-2026-87013` : Any authenticated user can start a non-terminating request via a folder parent cycle (CVSS 4.3, affects `>= 0.10.0, <= 0.11.0`)

> CVE-2026-87014 CVE ↗ STABLE
>
> `CVE-2026-87014` : Admin demoted through SSO role sync keeps read and write access to all users' notes (CVSS 6.5, affects `>= 0.9.0, < 0.11.1`)

> CVE-2026-87015 CVE ↗ STABLE
>
> `CVE-2026-87015` : A user's session cookies are sent to tool servers configured for bearer authentication (CVSS 6.8, affects `>= 0.6.27, < 0.11.1`)

> CVE-2026-87016 CVE ↗ STABLE
>
> `CVE-2026-87016` : Sign-in as another user via wildcard characters in the OAuth subject claim on SQLite (CVSS 8.1, affects `>= 0.6.41, < 0.11.1`)

> CVE-2026-87017 CVE ↗ STABLE
>
> `CVE-2026-87017` : Inaccessible knowledge bases are exposed through the built-in knowledge tool on most vector backends (CVSS 4.3, affects `>= 0.7.0, <= 0.11.0`)

> CVE-2026-87994 CVE ↗ STABLE
>
> `CVE-2026-87994` : Channel members can overwrite another member's message via the chat completions endpoint (CVSS 4.3, affects `>= 0.9.5, < 0.11.1`)

CVE-2026-87995 CVE ↗ STABLE

`CVE-2026-87995` : Same-origin XSS to account takeover via terminal port-preview iframe hardcoding allow-same-origin (CVSS 8.7, affects `>= 0.8.11, <= 0.11.0`)

CVE-2026-87996 CVE ↗ STABLE

`CVE-2026-87996` : SSRF into internal services via DNS rebinding in the Playwright web loader (CVSS 7.7, affects `>= 0.9.6, < 0.11.1`)

CVE-2026-87997 CVE ↗ STABLE

`CVE-2026-87997` : Any authenticated user can inject chats into another user's folder via chat completions (CVSS 4.3, affects `>= 0.10.0, < 0.11.1`)

CVE-2026-87999 CVE ↗ STABLE

`CVE-2026-87999` : Any authenticated user can reach the Azure platform channel via server-side web fetch (CVSS 7.1, affects `< 0.11.1`)

WAS THIS USEFUL? 👍 👎

oobabooga's Text Generation WebUI (textgen)

Feedback ⓘ CVE-2026-35483 CVE ↗ STABLE

Oobabooga Textgen runs local language models in a desktop app with chat, vision, tool-calling, web search, and an OpenAI-compatible API.

`CVE-2026-35483` : CWE-22 Path Traversal in load_template() — .jinja/.yaml/.yml file read without authentication (CVSS 5.3, affects `< 4.3`)

▾ **3 more releases in this issue** · 2025-10-13 → 2026-04-03

> CVE-2026-35486 CVE ↗ STABLE
>
> `CVE-2026-35486` : CWE-918 SSRF in superbooga/superboogav2 extensions — no URL validation (CVSS 7.5, affects `< 4.3`)

> CVE-2026-35050 CVE ↗ STABLE
> .
> .
> .
>
> `CVE-2026-35050` : Remote Code Execution (RCE) through Path Traversal at "Session -> Save extention settings to user_data/settings.yaml". (CVSS 9.1, affects `<= 4.1`)

> CVE-2025-62364 CVE ↗ STABLE
> .
> .
> .
>
> `CVE-2025-62364` : text-generation-webui allows arbitrary file read via symbolic link upload (CVSS 7.5, affects `<= 3.13`)

WAS THIS USEFUL? 👍 👎

llama.cpp Feedback ⓘ CVE-2026-34159 CVE ↗ STABLE

Llama CPP is a C/C++ library and set of command-line tools that runs large language models locally.

`CVE-2026-34159` : Unauthenticated RCE via `GRAPH_COMPUTE` `buffer=0` bypass in llama.cpp RPC backend (CVSS 9.8, affects `<= b7991`)

▾ **8 more releases in this issue** · 2024-07-22 → 2026-03-26

> CVE-2026-33298 CVE ↗ STABLE
>
> `CVE-2026-33298` : Heap Buffer Overflow via Integer Overflow in GGUF Tensor Parsing (CVSS 7.8, affects `< b7437`)

> CVE-2026-21869 CVE ↗ STABLE
>
> `CVE-2026-21869` : Out-of-bounds Write in llama.cpp llama-server (CVSS 9.8, affects `<= 55d4206c8`)

> CVE-2025-53630 CVE ↗ STABLE
>
> `CVE-2025-53630` : Integer Overflow in GGUF Parser can lead to Heap Out-of-Bounds Read/Write in gguf (affects `<26a48ad`)

> CVE-2025-52566 CVE ↗ STABLE
>
> `CVE-2025-52566` : Tokenizer signed vs. unsigned heap overflow (CVSS 8.6, affects `<b5721`)

> CVE-2025-49847 CVE ↗ STABLE
>
> `CVE-2025-49847` : Buffer Overflow in llama.cpp via Malicious GGUF Model – Exploitable via Vocabulary Loading (`llama_vocab::impl::token_to_piece`) (CVSS 8.8, affects All versions before patch)

> CVE-2024-42479 CVE ↗ STABLE
>
> `CVE-2024-42479` : Write-what-where in rpc_server::set_tensor (CVSS 9.8, affects `< b3561`)

CVE-2024-42477 CVE ↗ STABLE

`CVE-2024-42477` : global-buffer-overflow in ggml_type_size (CVSS 5.3, affects `< b3561`)

CVE-2024-41130 CVE ↗ STABLE

.

`CVE-2024-41130` : Null pointer dereference in gguf_init_from_file (CVSS 5.4, affects `<=b3426`)

WAS THIS USEFUL? 👍 👎

Apple MLX Feedback ⓘ CVE-2025-62609 CVE ↗ STABLE

MLX is Apple's array framework for machine learning on Apple silicon, with a NumPy-like API, unified memory, and lazy evaluation.

`CVE-2025-62609` : Wild Pointer Dereference in load_gguf() (CVSS 7.5, affects `<=0.29.3`)

▾ 1 more release in this issue · 2025-11-21

CVE-2025-62608 CVE ↗ STABLE

`CVE-2025-62608` : heap-buffer-overflow in load() (CVSS 9.1, affects `<=0.29.3`)

WAS THIS USEFUL? 👍 👎

◆ Inference Providers & GPU Clouds

Eigen Labs Darkbloom

Feedback ⓘ desktop-v0.1.0 NOTES ↗

CODE ↗ RANK

Darkbloom is a for the innovators pushing the frontier.

> Darkbloom shipped its first signed, notarized Mac desktop app for Apple Silicon with background auto-update and a local OpenAI-compatible API, bundled alongside Darkbloom CLI v0.10.2's additions for per-model SSD checkpoint quantization control, cache-routed prefix caching, logprobs, GLM OpenRouter reasoning fields and expanded CLI diagnostics.

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

01 Per-model SSD checkpoint quantization control

BREAKING

95

Adds the `engine_v2_kv_checkpoint_quantization_by_model` config key and `DARKBLOOM_CBV2_CHECKPOINT_QUANTIZATION` environment variable to control per-model SSD checkpoint quantization, with `native` restoring exact (non-quantized) state. Nemotron 3.5 Super and GLM-5.3-Flash now write 4-bit `balanced` SSD checkpoints under new cache namespaces by default; existing exact SSD files for these models miss once after updating then age out via the normal TTL sweep, so operators who need exact checkpoints should set `engine_v2_kv_checkpoint_quantization_by_model = { "<model id>" = "native" }` in provider config or `DARKBLOOM_CBV2_CHECKPOINT_QUANTIZATION=native`.

— Names config key, env var, migration path, and default behaviour change v0.10.2

02 Signed Mac desktop app for Apple Silicon NEW

77

Ships a signed and notarized Mac desktop installer (`Darkbloom-0.1.0-arm64.dmg` , Developer ID `SLDQ2GJ6TL`) for Apple Silicon running macOS 27 or later, including a one-click CLI install step. Supports background auto-update, where later versions download automatically and install on quit, and serves a local OpenAI-compatible API alongside the existing network provider mode.

— Names installer file, Developer ID, OS version, and auto-update behaviour

desktop-v0.1.0

03 SSD prefix-cache write queue increased IMPROVED

65

Raises the SSD prefix-cache write queue from 2 to 6 concurrent jobs via `SSDPrefixCachePolicy.writeQueueMaxJobs = 6` , applied to both attention-block

and checkpoint writers, reducing prefix recomputation under load.

– Exact config value and effect given, but no usage walkthrough `desktop-v0.1.0`

04 **Expanded diagnostics in darkbloom logs and status** IMPROVED

65

Extends `darkbloom logs` to surface named refusal reasons before inference and for client errors found during GLM-5.3-Flash prompt preparation (e.g. Inference failure: invalid request (unanswered tool call)). Extends `darkbloom status` to show models served outside the enabled model filter, Autopilot placement holds with their expiry, and what a model load cannot evict.

– Names both commands and the specific new output they surface `v0.10.2`

05 **Cache-routing allowlist for prefix caching** NEW

60

Adds the `EIGENINFERENCE_CACHE_ROUTING_ALLOWED_ARTIFACTS` coordinator config key to allowlist (model_id, model_aggregate_sha256, prompt_contract_id) tuples for cache-routed prefix caching.

– Names the config key and tuple format but no usage example `v0.10.2`

06 **GLM OpenRouter reasoning fields** NEW

60

Adds support for GLM-5.3-Flash OpenRouter reasoning fields: `reasoning.context`, `reasoning.mode`, hidden reasoning via `reasoning.exclude: true` / `include_reasoning: false`, Anthropic's `thinking` object, and null `reasoning_effort`. Also bundled into the desktop app via Darkbloom CLI v0.10.2.

– Lists every named field but no example request/response `v0.10.2` `desktop-v0.1.0`

THINNER COVERAGE BELOW

07 **Nemotron Super prefix bank reservation removed** BREAKING

50

Removes Nemotron 3.5 Super's default 16 GiB resident prefix bank reservation, changing default memory allocation for that model.

– States before/after number but no mitigation steps given `v0.10.2`

08 **Logprobs for Nemotron Super and GLM-5.3-Flash** NEW

45

Enables logprobs support for Nemotron 3.5 Super and GLM-5.3-Flash on coordinator requests once the model's catalog row lists `logprobs`; this was also bundled into the desktop app via Darkbloom CLI v0.10.2.

– Names the two models and activation condition, no request example `v0.10.2` `desktop-v0.1.0`

09 **MLX stack upstream sync** IMPROVED 38

Updates the MLX stack to the 2026-10-22 upstream sync, recompiling provider kernels on first start after update.

— Names the sync date but no scope of changes or impact v0.10.2

└──▷ **BREAKING ON UPGRADE**

! Nemotron 3.5 Super and GLM-5.3-Flash now write 4-bit `balanced` SSD checkpoints under new cache namespaces by default; existing exact SSD files for these models miss once after updating, then age out through the normal TTL sweep. Set `engine_v2_kv_checkpoint_quantization_by_model = { "<model id>" = "native" }` in provider config or `DARKBLOOM_CBV2_CHECKPOINT_QUANTIZATION=native` to retain exact checkpoints.

! Nemotron 3.5 Super no longer reserves its 16 GiB resident prefix bank by default.

WAS THIS USEFUL? 👍 👎

AWS Amazon Bedrock Feedback ⓘ

Distribute private offers from AWS Marketplace (2026-10-22) BLOG ↗ RANK

Amazon Bedrock provides managed foundation-model inference, customization, knowledge bases, and APIs for generative AI applications.

> Amazon Bedrock now lets organizations distribute private offers for third-party models directly from the AWS Marketplace purchase confirmation page.

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

01 **Private offer distribution from AWS Marketplace** NEW 42

Bedrock can distribute private offers for third-party Bedrock models to your organization directly from the AWS Marketplace purchase confirmation page, immediately after accepting the offer.

— Describes the workflow but no API, flag, or config detail.

Distribute private offers from AWS Mark…

WAS THIS USEFUL? 👍 👎

◆ Inference Serving & Optimization

Niko1221 Strata

Feedback ⓘ v0.2.6 NOTES ↗ RANK

Qwen3.8-Flash-Next on any consumer hardware: one-click install for Windows / Linux. Strata inference engine, OpenAI/Anthropic API on localhost, optional image input.

> Strata's v0.2.6 release adds automatic RAM-based DeepSeek expert caching and AMD GPU detection under WSL, while tightening OpenAI API error handling, refining GPU monitor accuracy, and fixing the Windows uninstaller.

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

01 **Validation error for empty chat requests** IMPROVED 60

Chat completion requests with no message now return HTTP 400 with the error `messages must be a non-empty list`, matching OpenAI API behavior.

– Names exact status code and error string, but no endpoint path given v0.2.6

THINNER COVERAGE BELOW

02 **GPU monitor excludes integrated graphics** IMPROVED 55

The GPU monitor now excludes integrated Radeon or Intel graphics from totals and card count when a discrete card is present, and assigns each card a stable ID.

– Describes behavior change clearly but no config or command to act on v0.2.6

03 **Working Windows uninstaller** IMPROVED 30

Strata now ships a working Windows uninstaller, fixing a prior issue.

– Clear fix but no detail on what was broken or how it works now v0.2.6

04 **Automatic RAM-based DeepSeek expert caching** NEW 20

Strata now automatically caches DeepSeek experts in RAM, as introduced in this release.

– Mentioned only in summary, no mechanism or numbers given v0.2.6

05 **AMD GPU detection under WSL** NEW 20

Strata now detects AMD GPUs when running under WSL (Windows Subsystem for Linux).

– Only named in summary, no detail on detection mechanism v0.2.6

WAS THIS USEFUL? 👍 👎

vLLM Feedback ⓘ CVE-2026-115913 CVE ↗ STABLE

The vLLM inference engine runs large language models with high-throughput, memory-efficient GPU execution.

`CVE-2026-115913` : Phi-4 multimodal audio encoder materializes and retains full attention matrices: a single max-length audio OOM-crashes the vLLM engine (CVSS 6.5, affects `0.31.0`)

▾ **21 more releases in this issue** · 2026-08-11 → 2026-10-20

> CVE-2026-115914 CVE ↗ STABLE
>
> `CVE-2026-115914` : DeepSeek-OCR family: an image whose normalized pixels sum to exactly 0 makes embed_multimodal return None, tripping the encoder-output sanity assert — one unauthenticated request kills the whole vLLM engine (default config) (CVSS 6.5, affects `0.30.0`)

> CVE-2026-115915 CVE ↗ STABLE
>
> `CVE-2026-115915` : Untrusted EPS inputs can trigger denial of service or Ghostscript RCE in vLLM (CVSS 8.1, affects `0.30.0`)

> CVE-2026-115912 CVE ↗ STABLE
>
> `CVE-2026-115912` : Public multimodal SHM handle forgery reaches worker-side pickle.loads() with insecure serialization disabled (CVSS 7.5, affects `<= 0.25.1`)

> CVE-2026-94624 CVE ↗ STABLE
>
> `CVE-2026-94624` : vllm p2p never connected zmq_exhaustion (CVSS 6.5, affects `0.26.0`)

> CVE-2026-115916 CVE ↗ STABLE
>
> `CVE-2026-115916` : Out-of-vocabulary prompt token id passes input validation and triggers a CUDA device-side assert (engine crash / DoS) (CVSS 6.5, affects `<= 0.31.0`)

> CVE-2026-115911 CVE ↗ STABLE
>
> `CVE-2026-115911` : DoS via unbounded chat_template Jinja2 rendering exhausts the renderer thread pool (CVSS 6.5, affects `>= 0.19.0`)

CVE-2026-114050 CVE ↗ STABLE

`CVE-2026-114050`: Forged multimodal UUID + P0/P1 processor-cache eviction drift: unauthenticated remote crash of the whole vLLM engine (CVSS 6.5, affects `< 0.31.0`)

CVE-2026-114052 CVE ↗ STABLE

`CVE-2026-114052`: MOSS-Audio request-controlled processor cache permits remote memory exhaustion (CVSS 7.5, affects `> 0.24.0, < 0.31.0`)

CVE-2026-105753 CVE ↗ STABLE

`CVE-2026-105753`: Mirrored multimodal IPC caches desync after a rejected request — a later request reusing the same media hash trips a receiver assertion in the engine core (CVSS 6.5, affects `< 0.28.0`)

CVE-2026-105758 CVE ↗ STABLE

`CVE-2026-105758`: Qwen2-VL / Qwen3-VL video samplers bound on request-controlled max_frames, which the num_frames ceiling does not reach (CVSS 5.3, affects `>= 0.24.0`)

CVE-2026-105759 CVE ↗ STABLE

`CVE-2026-105759`: Unbounded Prometheus label cardinality from attacker-controlled HTTP method tokens in the vLLM Rust frontend metrics middleware (unauthenticated denial of service) (CVSS 5.9, affects `< v0.30.0`)

CVE-2026-105760 CVE ↗ STABLE

`CVE-2026-105760`: GLMGA video sampling permits request-driven CPU and memory exhaustion (CVSS 5.3, affects `>= 0.23.0rc2, < 0.30.0`)

CVE-2026-105754 CVE ↗ STABLE

`CVE-2026-105754`: Scale-out disaggregated multimodal transport trusts caller-supplied features — shared EngineCore denial of service, encoder-cache poisoning, and transport integrity loss (5 sites) (CVSS 6.5, affects `< 0.30.0`)

CVE-2026-105757 CVE ↗ STABLE

`CVE-2026-105757`: Structured-output request errors escape the request boundary and terminate the shared EngineCore — engine-fatal denial of service (3 sites) (CVSS

6.5, affects `< 0.30.0`)

CVE-2026-105755 CVE ↗ STABLE

`CVE-2026-105755`: Flash late-interaction scoring caches query embeddings under a caller-controlled request id — cross-request integrity break and induced errors on `/score` and `/rerank` (CVSS 4.2, affects `< 030.0`)

CVE-2026-105756 CVE ↗ STABLE

`CVE-2026-105756`: Loose `cache_salt` validation lets a single request kill EngineCore on LMCache-MP deployments — uncaught downstream `ValueError` denial of service (CVSS 6.5, affects `< 0.30.0`)

CVE-2026-100648 CVE ↗ STABLE

`CVE-2026-100648`: Uncontrolled resource consumption: multimodal chat audio decoding ignores `VLLM_MAX_AUDIO_CLIP_FILESIZE_MB` (CVSS 4.3, affects `< 0.29.0`)

CVE-2026-100647 CVE ↗ STABLE

`CVE-2026-100647`: vLLM DoS via unbounded `cache_salt` length: multi-hundred-MB salt stalls the single EngineCore scheduler thread (CPU exhaustion) (CVSS 5.3, affects `<0.29.0`)

CVE-2026-93592 CVE ↗ STABLE

`CVE-2026-93592`: Denial of service: a negative token id in `/v1/embeddings` or `/pooling` input kills the vLLM engine via a CUDA device-side assertion (CVSS 7.5, affects `< 0.28.0`)

CVE-2026-90553 CVE ↗ STABLE

`CVE-2026-90553`: LlavaOnevision2 processor loader executes attacker model code with `trust_remote_code=False` (inert `trust_remote_code` kwarg to `transformers.get_class_from_dynamic_module`) — RCE from a malicious model (CVSS 7.8, affects `< 0.28.0`)

CVE-2026-78684 CVE ↗ STABLE

`CVE-2026-78684`: vLLM Unauthenticated Requests Exploit DeepStream Backend Confusion for DoS (CVSS 5.3, affects `< 0.27.0`)

WAS THIS USEFUL? 👍 👎

LMCache

Feedback ⓘ CVE-2026-105192 CVE ↗

LMCache reuses and shares language-model key-value caches across requests and serving instances to reduce repeated prefill work.

`CVE-2026-105192`: Remote code execution via unsafe deserialization in multiprocess mode transport (CVSS 9.8, affects `>= 0.3.9, < 0.5.6`)

WAS THIS USEFUL? 👍 👎

InternLM LMDeploy

Feedback ⓘ CVE-2026-33625 CVE ↗

STABLE

LMDeploy compresses, deploys, and serves language and vision-language models with its TurboMind engine, offering quantization and an OpenAI-compatible API.

`CVE-2026-33625` : Arbitrary code execution via eval() of untrusted quant_dtype in model config loading (CVSS 8.8, affects `>= 0.12.1, < 0.12.3`)

▾ **5 more releases in this issue** · 2025-12-26 → 2026-09-16

> CVE-2025-59953 CVE ↗ STABLE
>
> `CVE-2025-59953` : Remote code execution via unsafe pickle deserialization in the LMDeploy ZMQ RPC server (CVSS 9.8, affects `>= 0.9.1, < 0.10.2`)

> CVE-2026-46432 CVE ↗ STABLE
>
> `CVE-2026-46432` : Arbitrary code execution via hardcoded `trust_remote_code=True` in lmdeploy model initialization (CVSS 7.8, affects `<= 0.12.3`)

> CVE-2026-46517 CVE ↗ STABLE
>
> `CVE-2026-46517` : Hardcoded `trust_remote_code=True` is an implicit unsafe remote-code load path with no user opt-out (CWE-1188 default-insecure) (CVSS 7.8, affects `<=0.12.3`)

> CVE-2026-33626 CVE ↗ STABLE
>
> `CVE-2026-33626` : Server-Side Request Forgery (SSRF) in Vision-Language Image Loading (CVSS 7.5, affects `<= 0.12.0`)

> CVE-2025-67729 CVE ↗ STABLE
>
> `CVE-2025-67729` : Arbitrary Code Execution via Insecure Deserialization in torch.load() (CVSS 8.8, affects `<=0.11`)

WAS THIS USEFUL? 👍 👎

BentoML

Feedback ⓘ CVE-2026-40610 CVE ↗ STABLE

BentoML packages models and AI applications into inference APIs, job queues, and multi-model pipelines, and deploys them to Kubernetes or BentoCloud.

⋮

`CVE-2026-40610`: Information Disclosure in `bentoml build` via symlink traversal in the build context (CVSS 5.5, affects `<= 1.4.38`)

▾ **8 more releases in this issue** · 2025-04-04 → 2026-05-07

> CVE-2026-44346 CVE ↗ STABLE
>
> `CVE-2026-44346`: Dockerfile command injection via envs[*].name in bentofile.yaml (sibling fix-bypass of CVE-2026-33744 and CVE-2026-35043) (CVSS 8.8, affects `<= 1.4.38`)

> CVE-2026-44345 CVE ↗ STABLE
>
> `CVE-2026-44345`: BentoML Dockerfile command injection via docker.base_image (sister of pending GHSA-w2pm-x38x-jp44 / CVE-2026-33744 / CVE-2026-35043) (CVSS 8.8, affects `<= 1.4.38`)

> CVE-2026-35044 CVE ↗ STABLE
>
> `CVE-2026-35044`: Server-Side Template Injection via unsandboxed Jinja2 Environment in Dockerfile generation allows host code execution from malicious bento archives (CVSS 8.8, affects `<= 1.4.37`)

> CVE-2026-35043 CVE ↗ STABLE
>
> `CVE-2026-35043`: Incomplete fix for CVE-2026-33744: command injection in cloud deployment setup script (deployment.py:1648) (CVSS 7.8, affects `<= 1.4.37`)

> CVE-2026-27905 CVE ↗ STABLE
>
> `CVE-2026-27905`: Arbitrary File Write via Symlink Path Traversal in Tar Extraction (CVSS 7.8, affects `<= 1.4.8`)

> CVE-2025-54381 CVE ↗ STABLE
>
> `CVE-2025-54381`: SSRF in File Upload Processing (CVSS 9.9, affects `>=1.4.0, <1.4.19`)

CVE-2025-32375 CVE ↗ STABLE

`CVE-2025-32375` : Insecure Deserialization leads to RCE in BentoML's runner server (CVSS 9.8, affects `>=1.0,<1.4.8`)

CVE-2025-27520 CVE ↗ STABLE

`CVE-2025-27520` : Remote Code Execution (RCE) Caused by Insecure Deserialization (CVSS 9.8, affects `>=1.3.4,<1.4.3`)

WAS THIS USEFUL? 👍 👎

XGrammar Feedback ⓘ CVE-2026-25048 CVE ↗ STABLE

XGrammar constrains language-model output to grammars and schemas with optimized token-mask generation.

`CVE-2026-25048` : Multi-layer nesting causes DoS (CVSS 7.5, affects `0.1.31`)

▾ **2 more releases in this issue** · 2025-04-08 → 2026-03-05

CVE-2025-58446 CVE ↗ STABLE

`CVE-2025-58446` : Denial of service by huge enum grammar (CVSS 7.5, affects `0.1.23`)

CVE-2025-32381 CVE ↗ STABLE

`CVE-2025-32381` : Denial of Service by abusing xgrammar unbounded cache in memory (CVSS 6.5, affects `<0.1.18`)

WAS THIS USEFUL? 👍 👎

DATA

◆ Vector DBs & RAG

liliang-cn CortexDB

Feedback ⓘ v2.139.1 NOTES ↗ RANK

AI memory and a knowledge graph in one SQLite file. Pure Go: vectors, RAG, agent memory, RDF/SPARQL, Cypher, 80+ MCP tools.

> CortexDB shipped a small enhancement to its graph module, making entity upserts aware of the chunk collections they reference.

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

01 **Collection tracking in graph upsert_entities** IMPROVED

40

The `upsert_entities` function in the graph module now records the collections of the chunks it names, linking graph entities to their source chunk collections.

– Names the exact function but lacks mechanism detail or usage example. v2.139.1

WAS THIS USEFUL? 👍 👎

Docling

Feedback ⓘ v2.143.0 NOTES ↗ RANK

Docling parses PDFs, Office files, and HTML into structured documents with layout, tables, and reading order preserved for downstream RAG pipelines.

> Docling's v2.143.0 release focuses on expanding ASR capabilities with offline transcription and new presets, alongside mixed-format ZIP batch processing and finer control over VLM output stopping.

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

01 **Offline ASR via WhisperS2T presets** NEW 85

Docling can now run WhisperS2T ASR presets fully offline using the `--artifacts-path` flag, enabling air-gapped audio transcription without network access to fetch models.

Transcribe audio files in an air-gapped environment by pointing the ASR pipeline at a local model artifact directory instead of fetching from the network.

```
$ docling --artifacts-path /path/to/local/models recording.wav
```

— Names exact flag and gives a runnable command v2.143.0

THINNER COVERAGE BELOW

02 **GenerationStopper validity control for VLM output** NEW 55

A `GenerationStopper` can now declare a valid end of VLM output, giving custom stopping logic control over whether truncated output is accepted.

— Names the interface but no code sample shown v2.143.0

03 **Mixed-format ZIP archive processing** NEW 50

Docling now supports mixed-format ZIP inputs, allowing a single archive to contain documents of different file types for batch conversion. It also adds an optional `filename` field to `ZipArchiveResult` in the service layer, exposing the source filename per archive entry.

— Names config field but lacks a usage example v2.143.0

04 **Transformers and MLX ASR presets** NEW 40

Adds `transformers` and `MLX` presets for ASR pipelines, expanding the on-device speech-to-text backend options available.

— Names presets but no usage detail v2.143.0

CVE-2026-105743 CVE ↗ STABLE

`CVE-2026-105743`: SSRF guard bypass in remote resource fetching (DNS rebinding / multi-record resolution; no IP validation in HTML render mode) (CVSS 4.0, affects `>= 2.91.0, < 2.132.0`)

CVE-2026-105744 CVE ↗ STABLE

`CVE-2026-105744`: Arbitrary file read/write (and command execution when shell-escape is enabled) when rendering untrusted TikZ with the opt-in Tectonic engine (CVSS 7.5, affects `>= 2.94.0, < 2.132.0`)

CVE-2026-105748 CVE ↗ STABLE

`CVE-2026-105748`: Crafted DoclingDocument JSON embeds local image files into converted output (CVSS 4.3, affects `>= 2.16.0, < 2.131.0`)

CVE-2026-105747 CVE ↗ STABLE

`CVE-2026-105747`: METS-GBS archive member limit enforced after full member enumeration (memory exhaustion during format detection) (CVSS 4.3, affects `>= 2.45.0, < 2.131.0`)

CVE-2026-105749 CVE ↗ STABLE

`CVE-2026-105749`: Unbounded table rowspan/colspan in HTML, JATS, ODS and BoxNote backends causes CPU/memory exhaustion (CVSS 6.5, affects `>= 2.0.0, < 2.131.0`)

CVE-2026-105745 CVE ↗ STABLE

`CVE-2026-105745`: Plugin entry points are imported before the allow_external_plugins check (CVSS 6.7, affects `>= 2.27.0, < 2.131.0`)

CVE-2026-105750 CVE ↗ STABLE

`CVE-2026-105750`: `enable_local_fetch` is not enforced in HTML browser-rendering mode (CVSS 5.9, affects `>= 2.82.0, < 2.118.1`)

CVE-2026-105751 CVE ↗ STABLE

`CVE-2026-105751` : Arbitrary local file read via draw:image xlink:href in the OpenDocument backend (CVSS 7.5, affects `>= 2.107.0`)

CVE-2026-44016 CVE ↗ STABLE

`CVE-2026-44016` : Unsafe Playwright-based HTML Rendering (CVSS 8.2, affects `>= 2.82.0, < 2.91.0`)

CVE-2026-44022 CVE ↗ STABLE

`CVE-2026-44022` : Potential Path Traversal via LaTeX \includegraphics and \input Commands (CVSS 5.5, affects `>=2.73.0, <2.91.0`)

WAS THIS USEFUL? 👍 👎

EverMind AI EverOS Feedback ⓘ CVE-2026-115875 CVE ↗

EverOS provides a local-first memory layer that stores and evolves context for AI agents.

`CVE-2026-115875` : Path traversal in EverOS knowledge document upload via unvalidated filename (CVSS 8.2, affects `< 1.1.4`)

WAS THIS USEFUL? 👍 👎

lightrag Feedback ⓘ CVE-2026-114507 CVE ↗ STABLE

LightRAG is a lightweight knowledge-graph RAG framework combining dual-layer graph and vector retrieval for efficient document question answering.

`CVE-2026-114507`: Unauthenticated cache poisoning of query responses: `conversation_history` is missing from the query cache key (CVSS 6.5, affects `< 1.5.5`)

▾ **10 more releases in this issue** · 2026-04-02 → 2026-07-29

> CVE-2026-86062 CVE ↗ STABLE
>
> `CVE-2026-86062`: Stored Cross-Site Scripting (XSS) in the LightRAG WebUI chat/answer renderer via ingested content (CVSS 6.1, affects `<= 1.5.4`)

> CVE-2026-85740 CVE ↗ STABLE
>
> `CVE-2026-85740`: SSRF via IPv6-transition address bypass (NAT64, IPv4-compatible, 6to4) of the native-markdown image-download guard (CVSS 7.1, affects `< 1.5.5`)

> CVE-2026-85734 CVE ↗ STABLE
>
> `CVE-2026-85734`: No Rate Limiting on `/login` Endpoint Allows Brute-Force Attacks (CVSS 9.1, affects `<= 1.5.3`)

> CVE-2026-85725 CVE ↗ STABLE
>
> `CVE-2026-85725`: Plaintext Passwords Compared Without Constant-Time Function (CVSS 5.9, affects `<= 1.5.4`)

> CVE-2026-61808 CVE ↗ STABLE
>
> `CVE-2026-61808`: Missing Authentication for Critical API Functions in Default Configuration (CVSS 7.3, affects `<= 1.5.4`)

> CVE-2026-85709 CVE ↗ STABLE
>
> `CVE-2026-85709`: Sensitive Information Exposure Through Raw Exception Messages in API Error Responses (CVSS 5.3, affects `<= 1.5.4`)

> CVE-2026-61740 CVE ↗ STABLE

> `CVE-2026-61740` : Authentication bypass: hardcoded `DEFAULT_TOKEN_SECRET` and public `/auth-status` defeat `LIGHTRAG_API_KEY` protection (affects `<= 1.4.15`)

> `CVE-2026-61736` CVE ↗ STABLE
>
> `CVE-2026-61736` : CORS Wildcard + Credentials Enables Any-Origin Credentialed Requests (CVSS 9.3, affects `<= 1.5.3`)

> `CVE-2026-39413` CVE ↗ STABLE
>
> `CVE-2026-39413` : JWT Algorithm Confusion Vulnerability in LightRAG API (CVSS 4.2, affects `<= 1.4.13`)

> `CVE-2026-30762` CVE ↗ STABLE
>
> `CVE-2026-30762` : Hardcoded JWT Signing Secret Allows Authentication Bypass (CVSS 7.5, affects `<= 1.4.12`)

WAS THIS USEFUL? 👍 👎

Unstructured

Feedback ⓘ CVE-2026-104250 CVE ↗ STABLE

Unstructured is an open-source ETL library that converts PDFs, images, and office documents into chunked, structured elements for LLM ingestion.

`CVE-2026-104250`: Stored XSS via unescaped HTML generation (CVSS 6.1, affects `<= 0.24.0`)

▾ **2 more releases in this issue** · 2026-02-03 → 2026-07-27

> CVE-2026-71428 CVE ↗ STABLE
>
> `CVE-2026-71428`: Server-Side Request Forgery in unstructured URL-based partitioning (CVSS 9.3, affects `>= 0.4.7, < 0.24.0`)

> CVE-2025-64712 CVE ↗ STABLE
>
> `CVE-2025-64712`: Path Traversal via Malicious MSG Attachment Allows Arbitrary File Write (CVSS 9.8, affects `<= 0.18.17`)

WAS THIS USEFUL? 👍 👎

InfiniFlow RAGFlow

Feedback ⓘ CVE-2026-45312 CVE ↗

STABLE

RAGFlow is a retrieval-augmented generation engine built on deep document understanding, combining chunk-level citations with agentic workflows.

`CVE-2026-45312`: Server-Side Template Injection in Prompt Generator leads to Remote Code Execution (CVSS 9.9, affects `<=0.24.0`)

▾ **4 more releases in this issue** · 2025-02-21 → 2026-05-09

> CVE-2025-69286 CVE ↗ STABLE
>
> `CVE-2025-69286`: Predictable Token Generation Leading to Authentication Bypass Vulnerability (CVSS 9.8, affects `< 0.20.5`)

> CVE-2025-68700 CVE ↗ STABLE
>
> `CVE-2025-68700`: RAGFlow Remote Code Execution Vulnerability (CVSS 8.8, affects v0.22.0-31-g13e212c8 (and possibly earlier versions))

> CVE-2025-27135 CVE ↗ STABLE
>
> `CVE-2025-27135`: There is an SQL injection vulnerability in ragflow. (CVSS 9.8, affects `<= v0.15.1`)

> CVE-2025-25282 CVE ↗ STABLE
>
> `CVE-2025-25282`: Potential Insecure Direct Object Reference (IDOR) vulnerability in ragflow (CVSS 8.1, affects `v0.13.0`)

WAS THIS USEFUL? 👍 👎

Docling Graph

Feedback ⓘ CVE-2026-44520 CVE ↗

Docling Graph converts documents into Pydantic objects and builds directed knowledge graphs with explicit semantic relationships using LLM or VLM backends.

`CVE-2026-44520`: SSRF via Missing Internal IP Validation in URLInputHandler (CVSS 5.7, affects `<=1.5.0`)

WAS THIS USEFUL? 👍 👎

Zep Graphiti

Feedback ⓘ CVE-2026-32247 CVE ↗

Graphiti builds temporally-aware knowledge graphs from conversations and documents so agents can query facts that change over time.

`CVE-2026-32247`: Cypher Injection via unsanitized node_labels in search filters (CVSS 8.1, affects `<= 0.28.1`)

WAS THIS USEFUL? 👍 👎

Tencent WeKnora

Feedback ⓘ CVE-2026-30861 CVE ↗ STABLE

Open-source LLM knowledge platform: turn raw documents into a queryable RAG, an autonomous reasoning agent, and a self-maintaining Wiki.

`CVE-2026-30861` : Remote Code Execution (RCE) via Command Injection in MCP Stdio Configuration Validation (CVSS 9.9, affects `>= 0.2.5 < 0.2.10`)

▾ **5 more releases in this issue** · 2026-03-05 → 2026-03-06

> CVE-2026-30859 CVE ↗ STABLE
>
> `CVE-2026-30859` : Broken Access Control - Cross-Tenant Data Exposure (CVSS 5.3, affects `<= 0.2.11`)

> CVE-2026-30858 CVE ↗ STABLE
>
> `CVE-2026-30858` : DNS Rebinding Vulnerability in web_fetch Tool Allows SSRF to Internal Resources (CVSS 6.5, affects `<= 0.2.14`)

> CVE-2026-30857 CVE ↗ STABLE
>
> `CVE-2026-30857` : Unauthorized Cross-Tenant Knowledge Base Cloning (CVSS 5.3, affects `<= 0.2.14`)

> CVE-2026-30856 CVE ↗ STABLE
>
> `CVE-2026-30856` : Tool Execution Hijacking via Ambigous Naming Convention In MCP client and Indirect Prompt Injection (CVSS 5.9, affects `<= 0.2.14`)

> CVE-2026-30247 CVE ↗ STABLE
>
> `CVE-2026-30247` : SSRF via Redirection (CVSS 6.5, affects `<= 0.2.11`)

WAS THIS USEFUL? 👍 👎

Milvus

Feedback ⓘ CVE-2026-26190 CVE ↗ STABLE

Milvus is a high-performance, cloud-native vector database built for scalable vector ANN search

`CVE-2026-26190`: Unauthenticated Access to Restful API on Metrics Port (9091) Leads to Critical System Compromise (CVSS 9.8, affects `< 2.5.27`)

▾ **1 more release in this issue** · 2025-11-10 → 2026-02-11

> CVE-2025-64513 CVE ↗ STABLE
>
> `CVE-2025-64513`: Critical Authentication Bypass Vulnerability in Milvus Proxy (affects `<2.4.24, <2.5.21, <2.6.5`)

WAS THIS USEFUL? 👍 👎

Qdrant

Feedback ⓘ CVE-2026-25628 CVE ↗

Qdrant - High-performance, massive-scale Vector Database and Vector Search Engine for the next generation of AI. Also available in the cloud https://cloud.qdrant.io/

`CVE-2026-25628`: Arbitrary file write via `/logger` endpoint (CVSS 8.5, affects `>= 1.9.3, < 1.15.6`)

WAS THIS USEFUL? 👍 👎

EVALUATE

◆ AI Observability & Evals

LangChain LangSmith

Feedback ⓘ snapshot-20261027 NOTES ↗

RANK

LangSmith provides tracing, evaluation, and deployment tools for LLM applications.

> LangSmith's October 2026 release replaces its single broad organization permission with a granular model, adds an AI-powered run search, and ships a string of deployment, thread, and run management improvements alongside several breaking changes to deployment settings and the Teams API.

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

01 Metadata exclusion filters for thread search NEW 90

Adds `exclude_metadata` parameter to `POST /threads/search` and `POST /threads/count` to filter out threads whose metadata matches any listed object (e.g. automation threads).

Exclude automation threads (e.g. Open SWE sidebar threads) when searching for human-initiated threads in a deployment.

```
$ curl -X POST 'https://api.smith.langchain.com/threads/search' \
  -H 'x-api-key: <your-api-key>' \
  -H 'Content-Type: application/json' \
  -d '{"exclude_metadata": [{"source": "automation"}]}'
```

— Endpoints, parameter, and runnable curl example all given. snapshot-20261027

02 Granular organization permissions model BREAKING 85

Replaces the single broad organization `manage` permission with granular permissions covering members, access control, authentication, security policies, billing, usage controls, retention, data planes, workspace creation, OAuth applications, and GitHub integrations separately; `PATCH /orgs/current/info` now checks the permission for each field sent. Integrations relying on the old `manage` permission for billing, data planes, or integrations read access must be updated.

— Endpoint and full permission list given, no migration steps. snapshot-20261027

03 ABAC settings page tabs: Attributes, Policies, History NEW 70

Adds Attributes, Policies, and History tabs to the ABAC settings page, each with search and filters persisted in the URL and a CSV export; the History tab lists policy create,

update, delete, attach, and detach events for a chosen time range.

— Clear UI path and scope, no API endpoint named. snapshot-20261027

04 **Find-in-pane search for run details** NEW 60

Adds Find (Cmd/Ctrl+F) within the active run-detail pane to search loaded input/output messages, object fields, tags, and metadata, with match count and next/previous navigation.

— UI shortcut with described scope, no API surface. snapshot-20261027

THINNER COVERAGE BELOW

05 **CSV export of organization members and API keys** NEW 50

Adds CSV export of organization members and API keys from the Settings page, covering every row matching the current tab, search, and workspace filters.

— Thin UI feature, no endpoint or flag named. snapshot-20261027

06 **Engine permissions in custom roles** NEW 45

Adds Engine permissions to custom roles in Settings > Roles, with all four Engine permissions visible in the role editor.

— Names UI path and permission count, no further specifics. snapshot-20261027

07 **Cancel reason field on cancelled runs** NEW 35

Adds a `cancel_reason` field to cancelled runs so callers can read why a run was cancelled.

— Field named but no endpoint or example of retrieval. snapshot-20261027

└──▷ **BREAKING ON UPGRADE**

! The separate Revisions tab on deployment pages is removed; bookmarks to tabs after Overview now open the tab to their right.

! Custom deployment headers are no longer shown in the Settings dialog — they now live exclusively in the Connect dialog.

! The Teams APIs return 403 unless the organization has the `can_use_teams` setting enabled.

! The single broad organization `manage` permission is replaced by granular dedicated permissions; integrations relying on `manage` for billing, data planes, or integrations read access must be updated.

! Deleting a deployment now requires typing `permanently delete` to confirm.

WAS THIS USEFUL? 👍 👎

overmind

Feedback ⓘ v0.1.18 NOTES ↗ RANK

The platform for continuously improving AI agents.

> Overmind v0.1.18 adds Cursor sign-in support and broadens MCP protected-resource metadata availability.

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

01 **MCP protected-resource metadata path handling** IMPROVED

40

Serves MCP protected-resource metadata at both the trailing-slash and non-trailing-slash URL paths.

— Names the exact behaviour but no endpoint path given v0.1.18

02 **Cursor sign-in support** NEW

35

Enables authentication for Cursor and prefers the client's own sign-in flow once authentication is enabled.

— Describes behaviour but no concrete flow or config detail v0.1.18

WAS THIS USEFUL? 👍 👎

Langfuse

Feedback ⓘ v4.70.1 NOTES ↗ RANK

Langfuse provides tracing, evaluation, and monitoring for LLM applications.

> Langfuse shipped a small mobile usability improvement to the Topics view in this release.

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

01 Mobile usability improvements for Topics view IMPROVED

20

The Topics view has been improved for mobile devices.

— No specifics on what changed in the UI. v4.70.1

> CVE-2026-41487 CVE ↗ STABLE
>
> `CVE-2026-41487` : Improper role-based-access control in Langfuse LLM connection management allowed users of role `member` to retrieve stored LLM provider API keys (CVSS 5.4, affects `>= 3.68.0, < 3.167.0`)

> CVE-2026-24055 CVE ↗ STABLE
>
> `CVE-2026-24055` : Unauthenticated Slack OAuth Install Allows Linking Slack to Arbitrary Project Improper Access Control) (CVSS 5.3, affects `>=3.89.0, <= v3.146.0`)

> CVE-2025-65107 CVE ↗ STABLE
>
> `CVE-2025-65107` : SSO Account Takeover via CSRF or phishing attack (CVSS 6.5, affects `>= 3.17.0, <3.131.0`)

> CVE-2025-64504 CVE ↗ STABLE
>
> `CVE-2025-64504` : Cross-organization enumeration of member & invitation lists via project membership APIs (CVSS 5.0, affects `>=2.70.0, <3.124.1`)

WAS THIS USEFUL? 👍 👎

openlit

Feedback ⓘ openlit-2.1.1 NOTES ↗ LATEST STABLE

OpenLIT is the open-source agent harness engineering platform: trace, evaluate, guard, and improve everything around the model in your AI agents, on OTEL.

Adds NATS JetStream realtime threshold alerts (`NATS_URL`), a trace-level telemetry view, API key scope controls, and an Enterprise license boot-sync hook.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch openlit-2.1.1
https://github.com/openlit/openlit.git
# already have the repo? check out this version:
$ git checkout openlit-2.1.1
```

Adds realtime threshold alerts powered by a NATS JetStream engine, enabled by setting `NATS_URL` , with new alert pages and APIs in the dashboard.

Adds a trace-level telemetry view showing one row per trace with end-to-end duration, span count, and error count, while keeping the existing per-span list in a separate spans tab.

Adds API key feature-access extension points: key management now requires a signed-in session, key verification returns scopes, and community keys keep full access.

Adds a CE-safe license boot-sync hook so Enterprise can sync active licenses at startup; Community Edition remains a no-op with status logging and error isolation.

NEW INTEGRATION OrcaRouter provider

NEW INTEGRATION Trustabl scanner connectors

NEW INTEGRATION Memcode memory connector

NEW INTEGRATION LangGraph memory connector

NEW INTEGRATION TypeSafe Jev evaluation provider

NEW INTEGRATION OpenCode coding-agent support

UPDATED INTEGRATION First-party OTLP receiver replacing bundled OTel collector

▾ 1 more release in this issue · 2026-02-25 → 2026-10-26

> CVE-2026-27941 CVE ↗ STABLE
>
> `CVE-2026-27941` : Remote Code Execution and Secret Exposure via Misuse of `pull_request_target` in GitHub Actions Workflows (CVSS 9.9, affects

> `f0fc8fa9f2a14eff70750aa81005076fe96e861b`)

WAS THIS USEFUL? 👍 👎

evalscope

Feedback ⓘ v1.14.1 NOTES ↗ RANK

A streamlined and customizable framework for efficient large model (LLM, VLM, AIGC) evaluation and performance benchmarking.

> evalscope v1.14.1 adds zero-shot SafetyBench evaluation, a configuration dry-run mode, Bayes@N aggregation for repeated trials, and a shared Task Workspace in the Web UI, alongside several breaking changes to dependency installation, cache reuse and benchmark versioning.

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

01 SafetyBench zero-shot evaluation NEW

85

Adds zero-shot SafetyBench evaluation supporting `zh`, `en`, and `zh_subset` subsets, reporting overall and seven-category accuracies, with missing or unparsed predictions counted as incorrect.

Run zero-shot SafetyBench evaluation across all three subsets to get overall and seven-category safety accuracies.

```
$ evalscope eval --model my-model --api-url $OPENAI_API_BASE_URL
--api-key $OPENAI_API_KEY --eval-type openai_api --datasets
safetybench --dataset-args '{"safetybench": {"subset_list":
["zh", "en", "zh_subset"]}}'
```

— Names subsets, scoring rule and gives a runnable command v1.14.1

02 CLI dry-run flag NEW

75

Adds `--dry-run` to the CLI to parse and print configuration without running inference, letting users validate a run before firing any requests.

Validate your evalscope configuration before a long evaluation run, without firing any inference requests.

```
$ evalscope eval --model my-model --api-url $OPENAI_API_BASE_URL
--api-key $OPENAI_API_KEY --eval-type openai_api --datasets
safetybench --dry-run
```

— Flag and behavior named, runnable command shown v1.14.1

03 LiteLLM no longer a core dependency BREAKING

70

LiteLLM is no longer installed as a core dependency; users selecting `eval_type='litellm'` must now run `pip install litellm` explicitly.

— Exact config key and install command given v1.14.1

04 **Bayes@N binary aggregation** NEW 65

Adds binary uniform-prior Bayes@N aggregation reporting posterior mean and posterior standard deviation for complete fixed-repeat `accuracy` / `exact_match` trials.

— Mechanism detailed but no usage example or command v1.14.1

05 **Task Workspace in Web UI** NEW 65

Adds a Task Workspace in the Web UI where evaluation and performance testing share model connections, with independent drafts, reusable templates, import/export, and active-task discovery that reconnects monitoring on page reload.

— Describes behavior but no exact UI navigation or command v1.14.1

06 **Rerun-review cache reuse blocking** BREAKING 65

Interrupted or failed review reruns (with `rerun_review=True` / `--rerun-review`) now block ordinary cache reuse until the rerun is completed or explicitly resumed.

— Names the flags but no resume procedure given v1.14.1

THINNER COVERAGE BELOW

07 **Benchmark scoring and prompt changes** BREAKING 55

Input or scoring changes affect several benchmarks; check `evaluation_version`. Prompt changes (e.g. ScienceQA, SPaRC) require new predictions to be generated.

— Names affected benchmarks and version field, no migration steps v1.14.1

└──▷ BREAKING ON UPGRADE

! LiteLLM is no longer installed as a core dependency; users selecting `eval_type='litellm'` must now run `pip install litellm` explicitly.

! Interrupted or failed review reruns (with `rerun_review=True` / `--rerun-review`) now block ordinary cache reuse until the rerun is completed or explicitly resumed.

! Input or scoring changes affect several benchmarks; check `evaluation_version`. Prompt changes (e.g. ScienceQA, SPaRC) require new predictions to be generated.

WAS THIS USEFUL? 👍 👎

Harbor Framework Harbor

Feedback ⓘ v0.25.0 NOTES ↗

RANK

Framework for evaluating and improving agents

> Harbor v0.25.0 expands sandbox provider support with Sail, Kernel and Docker Sandboxes, adds a static task linter, and extends RewardKit grading and network policy controls.

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

01 **Static task validation via harbor lint** NEW 66

Adds the `harbor lint` command for static task validation, later improved with a `--select` option for targeted rule selection and support for lint rule configuration directly in task and dataset TOML files.

— Names command and flag but not the rule set v0.25.0

02 **Default skills directory relocated** BREAKING 64

The default skills directory has moved to `/runtime/skills`; existing setups referencing the old path will need to be updated.

— Gives exact new path and migration need v0.25.0

03 **New sandbox environment providers** NEW 63

Adds Sail (cloud sandbox), Kernel, and Docker Sandboxes as new environment providers, with Docker Sandboxes supporting both cloud and local runs. Sail also gains a Sailbox lifetime limit and migrates to the public SDK API.

— Names three providers and a lifetime limit, no usage steps v0.25.0

THINNER COVERAGE BELOW

04 **.harborignore for upload control** NEW 55

Adds `.harborignore` support for controlling uploads and publication.

— Names config file, lacks syntax details v0.25.0

05 **RewardKit sandbox and judge enhancements** NEW 45

Adds `reasoning_effort` forwarding to agent judges in RewardKit, and adds a sandbox Worker to RewardKit to grade submissions as an unprivileged user.

— Names the field but no mechanism detail Explains purpose but lacks config or usage specifics

v0.25.0

06 **Failure policy classification for trial results** NEW

38

Adds a failure policy table and accessors to trial results, classifying trial failures.

— Describes behaviour but no field or API names v0.25.0

07 **Verifier allowlist flag for network policy** NEW

36

Adds a verifier allowlist flag to network policy.

— Flag exists but not named or detailed v0.25.0

08 **MCP HTTP header forwarding** NEW

30

Forwards HTTP headers to supported agents via MCP.

— Bare statement, no mechanism or config detail v0.25.0

└──▷ **BREAKING ON UPGRADE**

! The default skills directory has moved to `/runtime/skills`; existing setups referencing the old path will need to be updated.

WAS THIS USEFUL? 👍 👎

MLflow Feedback ⓘ CVE-2026-115886 CVE ↗ STABLE

MLflow tracks, evaluates, and monitors machine learning and LLM applications, covering experiment tracking, tracing, prompt management, and a model registry.

`CVE-2026-115886`: SSRF on unauthenticated MLflow servers via third-party scorer kwargs in the scorer invoke, register, and GenAI evaluate endpoints (CVSS 6.5, affects `>= 3.12.0rc0, < 3.17.0`)

▾ **4 more releases in this issue** · 2023-03-24 → 2026-10-22

> CVE-2026-69148 CVE ↗ STABLE
>
> `CVE-2026-69148`: CreateModelVersion source validation does not check READ permission on referenced run_id (CVSS 6.5, affects `< 3.15.0`)

> CVE-2026-69146 CVE ↗ STABLE
>
> `CVE-2026-69146`: LogInputs endpoint bypasses per-run UPDATE authorization in MLflow basic-auth (CVSS 4.3, affects `< 3.16.0`)

> CVE-2026-64849 CVE ↗ STABLE
>
> .
> .
> .
> .
>
> `CVE-2026-64849`: Unauthenticated full-read SSRF in MLflow webhook delivery: _validate_webhook_url bypassed via unvalidated HTTP redirects (and DNS rebinding) (CVSS 8.6, affects `>= 3.10.0, < 3.15.0`)

> CVE-2023-1177 CVE ↗ STABLE
>
> `CVE-2023-1177`: Remote file access vulnerability in `mlflow server` and `mlflow ui` CLIs (CVSS 9.8, affects `<= 2.2.0`)

WAS THIS USEFUL? 👍 👎

Braintrust

Feedback ⓘ snapshot-20261027 NOTES ↗ RANK

Braintrust provides evaluation, tracing, and improvement workflows for AI applications.

> Braintrust's October release adds async workflow support to Go evaluations, introduces SQL-based dashboard charts, surfaces Gateway inference usage on the billing page, reorganizes Settings around Models and Permissions, and ships new tracing and logging hooks for Cursor and TypeScript apps.

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

01 Structured application logging in TypeScript SDK NEW

75

The TypeScript SDK exposes `logger` and `wrapConsole` from `braintrust`: calling wrapConsole() forwards existing `console.log` / `console.error` calls, and `logger.info` / `logger.error` emit structured messages (e.g. logger.info('Prompt rendered', { tokens: 512 })) that appear alongside traces in the Braintrust UI.

Emit structured application log messages from TypeScript that appear alongside your traces in the Braintrust UI.

typescript

```
import { logger, wrapConsole } from 'braintrust';

wrapConsole(); // forward existing console.log/error calls
logger.info('Prompt rendered', { tokens: 512 });
logger.error('Tool call failed', { tool: 'search', reason:
'timeout' });
```

— Names exact functions and shows a runnable code sample snapshot-20261027

02 Reorganized Settings with Models and Permissions tabs

IMPROVED

70

Settings is reorganized into a single sidebar: Models holds AI provider and Loop model settings, and Permissions holds permission groups, project permissions, and SCIM as tabs, with added sidebar and command-bar search. Permission groups are redesigned so name, description, permissions, and members are editable in one sheet, the list shows member count, and project group access can be granted from the permission group itself; editing name/description/permissions requires the Enterprise plan.

— Details exact tab structure and edit scope, but no API/config keys snapshot-20261027

03 Cursor agent session tracing NEW

70

Running `bt trace enable cursor` sets up full Cursor agent session tracing — turns, model calls, tool calls, and subagents — so every coding session is inspectable in Braintrust.

Set up full Cursor agent session tracing — turns, model calls, tool calls, subagents — so every coding session is inspectable in Braintrust.

```
$ bt trace enable cursor
```

— Gives an exact runnable command and what it traces `snapshot-20261027`

---

04 **SQL charts in dashboards** NEW 65

Dashboards can now include SQL charts: build a chart from a SQL query or a visual builder, save it from the SQL sandbox (renamed `Query`), and paste it into any dashboard. Available on Pro and Enterprise plans and requires data plane v2.1.1+.

— Clear plan/version requirements and workflow, but no exact UI path `snapshot-20261027`

THINNER COVERAGE BELOW

05 **Inference activity section on billing page** NEW 50

The billing page adds an Inference activity section showing Gateway token usage over the last 7 days, split by feature (Online scoring, Topics, Loop) and by top projects.

— Describes what's shown but not how to interact with it `snapshot-20261027`

06 **Workflow evaluation support in Go SDK** NEW 45

Go evaluations (SDK v0.18.0+) gain workflow evaluation support, letting tasks and scorers call async provider APIs such as batch APIs.

— Names SDK version and batch APIs but no code sample `snapshot-20261027`

WAS THIS USEFUL? 👍 👎

GOVERN

◆ AI/LLM Security

vArmor

Feedback ⓘ v0.10.6 NOTES ↗ RANK

The vArmor Kubernetes runtime security platform provides AppArmor, Seccomp, and eBPF policy enforcement for workloads.

> vArmor v0.10.6 introduces three automation features for EnhanceProtect policies—traffic-based network allowlisting, automated diagnosis, and automated policy optimization—alongside stricter validation and recreation requirements for enforcer configuration.

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

01 Stricter enforcer name validation on policy creation

BREAKING

65

Unknown `enforcer` components, duplicates, and separators such as `AppArmor-BPF` are now rejected on policy creation and update; only supported names or concatenated combinations such as `AppArmorBPF` are accepted.

— Gives exact rejected/accepted value formats v0.10.6

THINNER COVERAGE BELOW

02 NetworkProxy removal now requires policy recreation

BREAKING

50

Removing NetworkProxy from a combined enforcer policy now requires policy recreation; affected Pods must be cleaned up and recreated.

— States concrete before/after operational impact v0.10.6

03 Network Allowlist Advisor for NetworkProxy egress

NEW

45

Builds and validates NetworkProxy egress allowlists from live application traffic, with a user-review step before the allowlist is enforced.

— Describes mechanism but no CLI/API surface named v0.10.6

04 Automated Policy Optimization for EnhanceProtect

NEW

45

Refines `EnhanceProtect` policies through violation-log analysis combined with user-approved rule removal.

— Names config object and mechanism, lacks exact command v0.10.6

05 **Automated Diagnosis for policy troubleshooting** NEW

40

Adds read-only automated diagnosis covering policies, Pod injection, node enforcement, and NetworkProxy to help troubleshoot enforcement issues.

— Names covered areas but no usage steps given v0.10.6

└──▷ **BREAKING ON UPGRADE**

! Unknown `enforcer` components, duplicates, and separators such as `AppArmor-BPF` are now rejected on policy creation and update; only supported names or concatenated combinations such as `AppArmorBPF` are accepted.

! Removing NetworkProxy from a combined enforcer policy now requires policy recreation; affected Pods must be cleaned up and recreated.

WAS THIS USEFUL?  

ToolHive

Feedback ⓘ v0.57.0 NOTES ↗ RANK

ToolHive runs MCP servers in isolated containers and enforces identity and access policies for local and Kubernetes deployments.

> ToolHive's v0.57.0 release is almost entirely about authorization and identity plumbing for its Virtual MCP and auth server, adding backend-scoped vMCP authorization, richer OAuth token-exchange context, Dynamic Client Registration probing, and a cluster of new embedder/delegate controls over auth server identity and policy.

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

01 **OIDC transport flags in `thv run`** NEW 67

Wires OIDC transport flags into `thv run` auth middleware config, making OIDC transport options available directly on server launch.

– names the command and middleware but not flag names v0.57.0

02 **Dynamic Client Registration probing strategy** NEW 63

Adds `dcrConfig.registrationProbe` with an `authorization-endpoint` strategy for Dynamic Client Registration probing.

– names config key and strategy but not probing mechanics v0.57.0

THINNER COVERAGE BELOW

03 **`authorization_details` field on token exchange** NEW 52

Supports `authorization_details` on token exchange flows, enabling richer OAuth 2.0 authorization context during token exchange.

– names the field but not how context is consumed v0.57.0

04 **MCPRemoteProxy pod readiness classification** IMPROVED 50

Classifies `MCPRemoteProxy` pods by container readiness rather than phase alone, improving accuracy of pod status reporting.

– explains behavior change but no reporting interface named v0.57.0

05 **Backend-scoped authorization in Virtual MCP** NEW 48

Adds backend-scoped vMCP authorization, allowing per-backend authorization policies on the Virtual MCP server.

— describes capability without policy syntax or API v0.57.0

06 **Workload auth secret mounting in inline mode** NEW 42

Mounts workload auth secrets in inline mode, enabling credential injection for workload identity in inline deployments.

— no named flag or path for the mounted secret v0.57.0

07 **Auth server identity and embedder policy controls** NEW 42

Decouples auth server identity from host auth types, giving embedders independent control over auth server identity; adds opt-in upstream subject identifiers so the auth server can surface upstream subject IDs to downstream clients; lets delegate clients change policy on restart without a full redeployment; and lets embedders refuse refresh grants against dead upstreams.

— four related embedder/delegate controls but no config keys or APIs named v0.57.0

WAS THIS USEFUL?  

Docker Sandboxes

Feedback ⓘ v0.50.0 NOTES ↗ RANK

Docker Sandboxes runs AI agents in isolated Docker environments with controlled filesystem, network, and tool access.

> Docker Sandboxes v0.50.0 adds machine-readable cloud sandbox inspection, configurable kit host mounts, cross-update persistent kit volumes, and a string of security hardening fixes spanning Git references, path escapes, and MCP OAuth handling.

└──▷ **WHAT SHIPPED · 12 FEATURES** most completely described first ? what's the number?

01 **Configurable kit host mounts** NEW 83

Adds `--mount HOST:SANDBOX_PATH` flag to `sbx create` and `sbx run` for binding a host directory to a kit-declared `host-mount@1` capability; environment files can declare the same bindings via `mounts`.

Mount a local host directory into a kit-declared mount point when creating a sandbox, so the agent can read/write project files from the host.

```
$ sbx create --mount /home/user/project:/workspace/project my-kit
```

— Flag, capability name, and env config key given with example v0.50.0

02 **Cloud sandbox inspection command** NEW 78

Adds `sbx --cloud inspect <id|name>` to show details for a single cloud sandbox, including published port URLs, with `--format json` or `--format yaml` output for scripting.

Inspect a specific cloud sandbox and get its published port URLs in machine-readable form for scripting or CI pipelines.

```
$ sbx --cloud inspect my-sandbox --format json
```

— Named command with flags and a runnable example v0.50.0

03 **sbx cp symlink escape protection** NEW 70

`sbx cp` rejects symbolic links that escape a sandbox workspace when the workspace is reached through an alternate host path, including a symlink, junction, macOS `/private/tmp`, or a Linux bind mount of a workspace subdirectory.

— Concrete mechanism and edge cases named for a specific command v0.50.0

THINNER COVERAGE BELOW

04 **Persistent kit volumes across updates and clones** IMPROVED

58

Persistent kit volumes now retain data across kit updates and sandbox removal/recreation with the same name, and sandboxes created with `--clone` preserve kit-provided workspace files and report clone initialization failures. This cross-update retention applies only to sandboxes created with v0.50.0; existing sandbox records keep their previous storage behavior.

— Behaviour and scope explained but no command example v0.50.0

05 **HTTP query-parameter network rules** NEW

55

Adds a repeatable `--query name=value` flag to HTTP network rules to require specific query-parameter restrictions on allowed traffic.

— Flag named but no usage example or mechanism detail v0.50.0

06 **TLS trust store override diagnostics** NEW

55

`sbx diagnose` and daemon logs on macOS and Windows now warn when `SSL_CERT_FILE` or `SSL_CERT_DIR` replaces the system trust store and can cause HTTPS certificate verification failures.

— Names exact env vars and diagnostic command, no example run shown v0.50.0

07 **MCP OAuth security hardening** IMPROVED

55

MCP OAuth requests now block cross-origin credential redirects, rejected silent refreshes preserve saved credentials without opening a browser, and browser-open logs omit URL query strings and fragments to keep OAuth parameters and tokens out of daemon logs and diagnostic bundles.

— Mechanism described for three related OAuth fixes, no config surface v0.50.0

08 **Windows daemon named pipe user verification** BREAKING

50

Windows clients now verify that the daemon's named pipes belong to the same Windows user; older elevated daemons may need to be stopped with the previous CLI or Task Manager before using the updated CLI.

— Clear migration note but limited mechanism detail v0.50.0

09 **Git identity and kit reference validation** IMPROVED

48

Host Git identity is now copied into sandboxes for repositories using the reftable ref format, and Git kit references reject ambiguous paths, non-ASCII hostnames, and SSH transport syntax.

– Describes behaviour change but no command surface v0.50.0

10 **sbx exec failure exit status** NEW 40

Adds exit status 125 to `sbx exec` when the command cannot run, such as when the sandbox does not exist.

– Specific exit code named but minor scope v0.50.0

11 **Disable Environment MCP host-execution commands** NEW 40

Adds `--skip-host-commands` flag to disable Environment MCP host-execution commands.

– Named flag but no further mechanism or example v0.50.0

12 **Local kit rebuild from digest-matching requirement** BREAKING 35

Each local source kit rebuilds once after upgrading due to a new digest-matching requirement for reuse and composition.

– States consequence but not the underlying mechanism v0.50.0

└──▷ **BREAKING ON UPGRADE**

! Windows clients now verify that the daemon's named pipes belong to the same Windows user — older elevated daemons may need to be stopped with the previous CLI or Task Manager before using the updated CLI.

! Each local source kit rebuilds once after upgrading due to the new digest-matching requirement for reuse and composition.

! Persistent kit volumes with cross-update retention apply only to sandboxes created with v0.50.0; existing sandbox records retain their previous storage behavior.

WAS THIS USEFUL?  

containarium

Feedback ⓘ v0.119.0 NOTES ↗ RANK

Open-source agent runtime — SSH-native isolation, eBPF egress policy, Kubernetes + LXC backends, GPU passthrough, MCP-native CLI

> Containarium shipped whole-box backup support, new MCP sprint-planning tools with env-var wiring for Claude Code/Cursor, Proof Pack audit hash-chain sealing, and tracker change-request submission from file sets, alongside smaller additions like SSH device approval, kernel-support reporting, and a guard against accidental version downgrades.

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

01 Whole-box backup support NEW

90

Adds a `backup create --box` subcommand for whole-box backup creation, a streamed instance export/import seam to support whole-box backup workflows, and whole-box scope to the `BackupService` contract, extending backup coverage beyond individual containers.

Create a whole-box backup of a running container, capturing the full instance state.

```
$ containarium backup create --box --server http://your-host:8080
```

— Names command, seam and contract with a runnable example v0.118.0 v0.119.0

02 Sprint planning tools and env wiring in MCP server NEW

85

Adds `sprint_board` and `sprint_list` tools to the MCP sprint client, exposing sprint planning surfaces via the platform MCP (`mcp-server`). Adds `CONTAINARIUM_SERVER_URL` and `CONTAINARIUM_JWT_TOKEN` environment variables for wiring `mcp-server` into Claude Code / Cursor via `~/.claude.json` .

Wire the platform MCP (including the new sprint_board and sprint_list tools) into Claude Code so an AI agent can manage containers and sprints.

```json
{
  "mcpServers": {
    "containarium": {
      "command": "/usr/local/bin/mcp-server",
      "env": {
        "CONTAINARIUM_SERVER_URL": "http://your-host:8080",
        "CONTAINARIUM_JWT_TOKEN": "<your-token>"
      }
    }
  }
}
```

— Config example shows exact env vars and file path `v0.118.0`

03 **Guard against downgrading via --github-tag** BREAKING

65

Adds a guard that refuses `--github-tag` values older than the running version, and an `--allow-downgrade` flag to override that guard.

— Flag and behavior named but no example command shown `v0.119.0`

04 **Tracker change-request submission from file sets** NEW

65

Adds `SubmitTrackerChangeFiles` to open a change request from a file set over the connection's broker credential, records the submitted change on the dispatch row, and exposes it via `ListTrackerDispatches`.

— API names given but no example invocation shown `v0.117.0`

05 **Proof Pack Workflow Enforcement** NEW

60

Proof Pack workflows now seal a Proof Pack into an audit hash chain and support verification, attaching sealed packs to tracker-dispatched PRs, and add a typed `CrewRole` per engineer-crew member that rejects a pack producer who is also the coder, enforcing separation of duties.

— Mechanism named but no verify command or endpoint shown Rule named but no config or command to configure roles

`v0.118.0`

THINNER COVERAGE BELOW

06 **KERNEL_UNSUPPORTED status reporting** NEW

45

Reports a `KERNEL_UNSUPPORTED` status when the host kernel is below the TCX kernel floor.

— Names status and threshold but no remediation steps v0.119.0

07 **GCS anchor retention lock warning** NEW 45

Warns when the GCS anchor retention lock is configured under 365 days, guarding against under-protected audit anchors.

— Threshold named but no config key or command given v0.118.0

08 **AI turn journaling for Claude and Codex** NEW 35

Journals how Claude and Codex turns ended, capturing AI agent run-completion state.

— Describes capture but no storage location or query surface v0.117.0

09 **SSH device approval on same connection** NEW 25

Adds device approval on the same SSH connection in the enforcing plugin.

— Single sentence, no mechanism or command given v0.119.0

└──▷ ALSO FROM THESE RELEASES

Verify a release binary has not been tampered with using its keyless cosign signature before deploying to a host.

```
$ cosign verify-blob \
  --bundle containariumd-linux-amd64.sigstore.json \
  --certificate-identity
https://github.com/FootprintAI/Containarium/.github/workflows/release.
yml@refs/tags/v0.118.0 \
  --certificate-oidc-issuer
https://token.actions.githubusercontent.com \
  containariumd-linux-amd64
```

└──▷ BREAKING ON UPGRADE

! Upgrade attempts using `--github-tag` with a version older than the running daemon are now refused unless `--allow-downgrade` is also passed.

WAS THIS USEFUL? 👍 👎

Backslash

Feedback

AI Plugin Inventory: Uncover Hidden Risks | Backslash - Backslash (2026-10-26)

BLOG ↗

Backslash detects and prioritizes reachable vulnerabilities in cloud-native application code and dependencies.

Backslash introduces a Plugin Inventory with child-component risk scoring, digital signature validation, inline host-level guardrails, and a unified inventory covering plugins, hooks, local models, rules files, and MCP servers.

Adds a dedicated Plugin Inventory surfacing publisher, source repository, unpacked contents, risk severity, associated agents, and last sync timestamp via the Plugins tab — fully filterable and sortable.

NEW DETECTION Introduces child-component risk scoring: a plugin's overall severity is set by the highest-risk element among its nested scripts, tools, hooks, and rule files.

Adds digital signature validation for hook files to verify provenance and detect tampering before execution.

Adds a Hooks Inventory that tracks whether a hook was added by Backslash, a developer, or a third-party plugin.

Enforces host-level inline guardrails that block unapproved command executions, credential access, and data egress before actions execute.

NEW DETECTION AI Rules file scanning for instruction poisoning and exposed credentials

NEW SUPPORTED PLATFORM Local AI model discovery on macOS and Windows for Ollama, LM Studio, and llama.cpp

NEW INTEGRATION MCP server and agent skill inventory mapping

NEW COMPLIANCE FRAMEWORK Automated audit records for EU AI Act, NIS2, DORA, and SOC

WAS THIS USEFUL? 👍 👎

NoLabs nono

Feedback ⓘ CVE-2026-47128 CVE ↗

The nono runtime runs AI agents in isolated environments with policy-controlled filesystem, network, and credential access.

`CVE-2026-47128`: nono: Sandbox escape on Linux via D-Bus: `systemd-run --user` (CVSS 6.1, affects `< 0.55.0`)

WAS THIS USEFUL? 👍 👎

◆ AI Runtime Defense & Guardrails

PipeLab Pipelock

Feedback ⓘ CVE-2026-91179 CVE ↗ STABLE

Pipelock is an open-source AI agent firewall that scans MCP and HTTP egress for exfiltration, SSRF, and prompt injection and emits signed action receipts.

`CVE-2026-91179`: hostile-model preset: shell-profile destination-write bypass via MCP move_file (equivalent-operation coverage gap) (affects `<= 3.5.0`)

▾ 3 more releases in this issue · 2026-08-20 → 2026-10-03

> CVE-2026-91180 CVE ↗ STABLE
>
> `CVE-2026-91180`: Pipelock media policy bypass through mismatched response content types (CVSS 5.3, affects `>= 2.2.0, < 3.5.0`)

> CVE-2026-91178 CVE ↗ STABLE
>
> `CVE-2026-91178`: URL DLP Scanner Bypass via Path-Query Secret Splitting (CVSS 5.8, affects `< 3.4.0`)

> CVE-2026-92145 CVE ↗ STABLE
>
> `CVE-2026-92145`: Hot-reload can silently retain stale security config under burst edits (affects `< 2.6.0`)

WAS THIS USEFUL? 👍 👎

◆ AI Red Teaming & Testing

Giskard

Feedback ⓘ CVE-2026-40319 CVE ↗ STABLE

Giskard tests and evaluates LLM applications for performance, bias, and security issues.

`CVE-2026-40319`: Regular Expression Denial of Service (ReDoS) in RegexMatching Check (CVSS 5.5, affects `<= 1.0.1b1`)

▾ 3 more releases in this issue · 2024-11-14 → 2026-04-13

> CVE-2026-40320 CVE ↗ STABLE
>
> `CVE-2026-40320`: Unsandboxed Jinja2 Template Rendering in ConformityCheck (CVSS 7.8, affects `<= 1.0.1b1`)

> CVE-2026-34172 CVE ↗ STABLE
>
> `CVE-2026-34172`: Server-side template injection via ChatWorkflow.chat() using non-sandboxed Jinja2 Environment (CVSS 8.8, affects `<= 0.3.3`)

> CVE-2024-52524 CVE ↗ STABLE
>
> `CVE-2024-52524`: ReDoS in Giskard Scan text perturbation (affects `<= 2.15.4`)

WAS THIS USEFUL? 👍 👎

AI MODELS

◆ Frontier Models

Anthropic

Feedback ⓘ snapshot-20261027 NOTES ↗ RANK

The Claude Developer Platform is the API for building applications on Anthropic's Claude models that generate, analyze, and automate work.

> Anthropic added Admin API key scopes that let administrators read and write Claude Tag settings for Slack channels and workspaces, including visibility into setting origins and Access bundle attachment.

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

01 **Admin API scopes for Claude Tag in Slack** NEW 70

New `read:claude_tag` and `write:claude_tag` Admin API key scopes allow reading and writing Claude Tag settings on Slack channels and workspaces. The API exposes which Claude Tag settings apply in a given Slack channel, traces where each setting originates, and supports attaching Access bundles via the API.

— Names exact scopes and behavior but no endpoint paths or examples. snapshot-20261027

WAS THIS USEFUL? 👍 👎

Mistral AI

Feedback ⓘ snapshot-20261027 NOTES ↗ RANK

Learn how to deploy and use Mistral AI's Large Language Models with our comprehensive documentation, guides, and tutorials.

> Mistral AI shipped a wave of new models headlined by Mistral Large 4, the Voxtral 3 audio family and Z.ai's GLM 5.3 Flash (several with 1M-token context windows), alongside OCR API enhancements, Custom Guardrails for Agents and Conversations, several model deprecations, and multiple breaking changes to the Completions and Chat APIs.

└──▷ **WHAT SHIPPED · 17 FEATURES** most completely described first ? what's the number?

01 **OCR API parameters for layout and confidence** IMPROVED

90

Adds `confidence_scores_granularity` (supporting `'block'` granularity for page-, block-, or word-level confidence scores), `include_blocks` (when `true`, returns a `blocks` array per page with paragraph-level bounding boxes and structural labels in reading order), support for comma-separated digits and ranges (e.g. `'0,1,2'`, `'0-5'`, `'0,2-4'`) in the `pages` parameter alongside lists of integers, `table_format` (choice between `markdown` and `html`), and `extract_footer` / `extract_header` parameters plus hyperlinks output.

— Every OCR parameter named with exact behavior and values snapshot-20261027

02 **Model deprecations and renames** DEPRECATED

85

`mistral-ocr-4-0` is deprecated and retires on September 30, 2026, with migration to `mistral-ocr-4-1` or `mistral-ocr-latest`; `labs-leanstral-1-5` is deprecated and retires on September 30, 2026; `zai-glm-5-2` is deprecated and retires on October 31, 2026, with migration to `zai-glm-5-3`; and the `mistral-tiny` endpoint is renamed to `mistral-tiny-2312` and will be deprecated in three months.

— Exact model names, retirement dates, and migration targets given snapshot-20261027

03 **Z.ai GLM 5.3 Flash model** NEW

75

Adds `zai-glm-5-3-flash`, a multimodal 320B-parameter (18B active) MoE model from Z.ai with a 1M-token context window, priced at $0.15/M input, $0.03/M cached input, and $0.50/M output tokens.

— Detailed specs and pricing but no call example snapshot-20261027

04 **Voxtral 3 audio model family** NEW

73

Public Preview launch of the Voxtral 3 family: `voxtral-mini-tts-3-0` (text-to-speech with zero-shot voice cloning, streaming, and contextual delivery), `voxtral-`

`mini-transcribe-3-0` (offline transcription and diarization with timestamps), `voxtral-mini-transcribe-realtime-3-0` (real-time transcription), and `voxtral-small-transcribe-3-0`.

— Names four model IDs and their capabilities but no API call shown `snapshot-20261027`

---

05 **Custom Guardrails for Agents and Conversations** NEW

70

Adds Custom Guardrails support for Agents and Conversations via the `guardrails` parameter on an Agent and per-request on `POST /v1/conversations`.

— Named parameter and exact endpoint given `snapshot-20261027`

---

06 **Completion API streaming chunk format change** BREAKING

70

A new parameter `p` added to chunks streamed by the Completion API may break applications relying on strict parsing of those chunks; users relying on the `mistral-common` package must update to version 1.8.4 or higher.

— Exact field and required package version given `snapshot-20261027`

---

07 **Mistral Large 4 flagship model** NEW

65

Adds `mistral-large-4`, an open-weight multimodal Mixture-of-Experts model with a 1M-token context window, released in Public Preview with launch pricing at 50% off for 2 weeks.

— Model ID and specs given, no usage example `snapshot-20261027`

---

08 **Mistral Medium 3.5 frontier model** NEW

65

Adds `mistral-medium-3-5`, a frontier-class multimodal model with adjustable reasoning via the `reasoning_effort` parameter, released as open weights under a Modified MIT license.

— Names model, parameter, and license but no usage example `snapshot-20261027`

THINNER COVERAGE BELOW

---

09 **Mandatory tool_call_id in tool-role messages** BREAKING

55

`tool_call_id` is now mandatory in chat messages with the `tool` role; requests omitting it will fail.

— Named field and failure condition, no migration example `snapshot-20261027`

---

10 **Mistral Small 2603 hybrid model** NEW

50

Adds `mistral-small-2603` (`mistral-small-4`), a hybrid model unifying instruct, reasoning, and coding in a single multimodal model with a 256k context window.

– Model named with context window spec, no usage example `snapshot-20261027`

11 **Strict API parameter validation** BREAKING 50

The API now strictly enforces parameter validity and rejects unsupported parameters with `Extra inputs are not permitted`, whereas previously they were silently ignored.

– Describes behavior change and exact error message `snapshot-20261027`

12 **Pixtral-12b temperature scaling change** BREAKING 50

The temperature parameter for `pixtral-12b` was downscaled by a multiplier of 0.43, changing output behavior for existing integrations.

– Exact multiplier given but no migration guidance `snapshot-20261027`

13 **Voxtral TTS 2603 model** NEW 45

Adds `voxtral-tts-2603`, a text-to-speech model with zero-shot voice cloning, multilingual support, and real-time streaming.

– Model named with capabilities but no API detail `snapshot-20261027`

14 **Leanstral open-source Lean 4 code agent** NEW 40

Adds `labs-leanstral-2603` (`Leanstral`), the first open-source code agent designed for Lean 4 formal proof engineering.

– Names the model and its purpose, little else `snapshot-20261027`

15 **Devstral 2.0 paid API access** BREAKING 35

`Devstral 2.0` moves to paid API access.

– States the change but no pricing or date detail `snapshot-20261027`

16 **Documentation path redirects to /vibe/*** IMPROVED 35

Documentation paths `/le-chat/*` and `/mistral-vibe/*` now redirect to `/vibe/*`.

– Exact old and new paths given, minor documentation change `snapshot-20261027`

17 **Mistral Moderation 2603 model** NEW 20

Adds `mistral-moderation-2603` (Mistral Moderation 2603).

– Only a name is given, no capability description `snapshot-20261027`

└──▷ BREAKING ON UPGRADE

- ! A new parameter `p` added to chunks streamed by the Completion API may break applications relying on strict parsing of those chunks; users relying on the `mistral-common` package must update to version 1.8.4 or higher.
- ! `tool_call_id` is now mandatory in chat messages with the `tool` role; requests omitting it will fail.
- ! The API now strictly enforces parameter validity and rejects unsupported parameters with `Extra inputs are not permitted`, whereas previously they were silently ignored.
- ! `mistral-ocr-4-0` is deprecated and retires on September 30, 2026; migrate to `mistral-ocr-4-1` or `mistral-ocr-latest`.
- ! `labs-leanstral-1-5` is deprecated and retires on September 30, 2026.
- ! `zai-glm-5-2` is deprecated and retires on October 31, 2026; migrate to `zai-glm-5-3`.
- ! The `mistral-tiny` endpoint is renamed to `mistral-tiny-2312` and will be deprecated in three months.
- ! The temperature parameter for `pixtral-12b` was downscaled by a multiplier of 0.43, changing output behavior for existing integrations.
- ! `Devstral 2.0` moves to paid API access.
- ! Documentation paths `/le-chat/*` and `/mistral-vibe/*` now redirect to `/vibe/*`.

WAS THIS USEFUL? 👍 👎

OpenAI GPT-6.1 Terra

Feedback ⓘ released · Oct 27, 2026

MODEL ↗ BLOG ↗

OpenAI: "Use GPT-6.1 Terra to match product photos to catalog entries, interpret technical drawings, act on application screens, or draft and revise text."

GPT-6.1 Terra from OpenAI was released on 26 Oct 2026. 1050K-token context; $1.00 in / $5.00 out per million tokens (OpenRouter).

Succeeds GPT-5.6 Terra.

WAS THIS USEFUL? 👍 👎

OpenAI GPT-6.1 Terra Pro

Feedback ⓘ

released · Oct 27, 2026 MODEL ↗

GPT-6.1 Terra Pro from OpenAI was released on 26 Oct 2026. 1050K-token context; $1.00 in / $5.00 out per million tokens (OpenRouter).

Succeeds GPT-5.6 GPT-6.1 Terra Pro.

WAS THIS USEFUL? 👍 👎

◆ Open-Weight Models

vLLM Project d3-ultra

Feedback ⓘ released · Oct 27, 2026

MODEL ↗

d3-ultra from vLLM Project was released on 24 Oct 2026. Open weights on Hugging Face.

WAS THIS USEFUL? 👍 👎

Xiaomi

Feedback ⓘ v0.1.16 NOTES ↗ RANK

MiMo Code: Where Models and Agents Co-Evolve

> MiMo Code v0.1.16 retires the legacy Compose agent and workflow in favor of a new `/compose-next` build-agent flow, adds lazy MCP tool loading and durable turn-queue scheduling, and moves several previously-automatic behaviors (agent-CLI skills, Claude Code session import) behind opt-in environment variables.

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

01 Compose agent retired for /compose-next workflow

BREAKING

85

The `compose` agent, its built-in workflow, and the entire `compose:*` skill set are removed and replaced by `/compose-next` on the build agent, which runs a clarify → spec → implement → verify → review → finish sequence for complex tasks.

— Names exact command and replaced workflow stages v0.1.16

02 Agent-CLI skills and Claude Code import now opt-in

BREAKING

78

Agent-CLI skills (`claude-code` , `codex` , `grok-build`) are gated behind the `MIMOCODE_ENABLE_AGENT_CLI_SKILLS` environment variable instead of being active by default, and Claude Code session import on startup is likewise made opt-in via an environment variable rather than running automatically.

— Names exact env var and all gated skills v0.1.16

03 Lazy MCP tool loading and background connection

IMPROVED

65

MCP tool handling now uses conversation-scoped lazy loading via `mcp_tool_use` , discovering tools on demand and calling them once ready instead of dumping the full catalog up front; MCP servers connect in the background with a bounded first-turn wait and concise readiness hints for tools still deferred.

— Explains mechanism and names the trigger, no user-facing command v0.1.16

THINNER COVERAGE BELOW

04 Installer and environment variable changes

BREAKING

55

`npm i @mimo-ai/cli` now installs only a thin launcher, with the native installer required for the full experience, and leftover `OPENCODE_*` environment variables have been renamed in this release.

— Names exact install command and affected env var prefix v0.1.16

05 **Plugin SDK served from host** IMPROVED 35

The plugin SDK is now served directly from the host instead of requiring separate installation.

— Clear before/after but no further detail v0.1.16

06 **Durable turn-queue scheduling and subagent waiting** NEW 30

A turn-queue subsystem now provides durable scheduling and lets subagents wait within a turn.

— Names the subsystem but no behavioral detail given v0.1.16

07 **Host-defined API error classification and retry policies** NEW 30

Hosts can now define classification of API errors and corresponding retry policies.

— States the capability without mechanism or config surface v0.1.16

08 **Default voice control model updated** IMPROVED 30

The default voice control model is switched to `mimo-v2.6-flash`.

— Names the model with no further behavioral detail v0.1.16

09 **Experimental Orchestrator mode removed** BREAKING 25

The experimental Orchestrator mode has been removed from MiMo Code.

— Bare removal notice, no replacement or migration detail v0.1.16

└──▷ **BREAKING ON UPGRADE**

! The `compose` agent, all built-in `compose:*` skills, and its workflow are removed; switch to `/compose-next` on the build agent for complex tasks.

! The experimental Orchestrator mode is removed.

! Claude Code session import no longer runs automatically on startup — it must be explicitly enabled via an environment variable.

! The agent-CLI skills (`claude-code` , `codex` , `grok-build`) are no longer active by default and must be enabled via `MIMOCODE_ENABLE_AGENT_CLI_SKILLS` .

! `npm i @mimo-ai/cli` now installs only a thin launcher; the native installer is required for the full experience.

! `OPENCODE_*` environment variables have been renamed (leftover `OPENCODE_*` env vars were renamed in this release).

WAS THIS USEFUL? 👍 👎

DeepSeek V4 Pro 0813 Feedback ⓘ

price increase · Oct 26, 2026 MODEL ↗

DeepSeek V4 Pro 0813 is a chat model from DeepSeek.

Price increase on OpenRouter (deepseek/deepseek-v4-pro): $0.2088 → $0.9553 in, $0.4176 → $1.9105 out per million tokens (+358%), first seen 26 Oct 2026.

Price increase on OpenRouter (deepseek/deepseek-v4-pro-0813): $0.66 → $1.32 in, $1.98 → $3.96 out per million tokens (+100%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

DeepSeek V4.1 Flash Feedback ⓘ

price increase · Oct 26, 2026 MODEL ↗

DeepSeek V4.1 Flash is a multimodal model from DeepSeek.

Price increase on OpenRouter (deepseek/deepseek-v4.1-flash-1017): $0.15 → $0.30 in, $0.60 → $1.20 out per million tokens (+100%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

Z.ai GLM 5.3

Feedback ⓘ price cut · Oct 26, 2026 MODEL ↗

GLM 5.3 is a reasoning model from Z.ai.

Price cut on OpenRouter (z-ai/glm-5.3): $0.079 → $0.0699 in, $1.40 → $1.40 out per million tokens (-12%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

Google Gemma 4 26B A4B

Feedback ⓘ

price increase · Oct 26, 2026 MODEL ↗

Gemma 4 26B A4B is a multimodal model from Google.

Price increase on OpenRouter (google/gemma-4-26b-a4b-it): $0.0675 → $0.09 in, $0.225 → $0.30 out per million tokens (+33%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

Alibaba Qwen3 30B A3B

Feedback ⓘ

price increase · Oct 26, 2026 MODEL ↗

Qwen3 30B A3B is a chat model from Alibaba.

Price increase on OpenRouter (qwen/qwen3-30b-a3b-instruct-2507): $0.09 → $0.10 in, $0.30 → $0.30 out per million tokens (+11%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

Alibaba Qwen3.5-122B-A10B

Feedback ⓘ

price cut · Oct 26, 2026 MODEL ↗

Qwen3.5-122B-A10B is a multimodal model from Alibaba.

Price cut on OpenRouter (qwen/qwen3.5-122b-a10b): $0.40 → $0.26 in, $3.20 → $2.08 out per million tokens (-35%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

Alibaba Qwen3.6 27B

Feedback ⓘ

price increase · Oct 26, 2026 MODEL ↗

Qwen3.6 27B is a multimodal model from Alibaba.

Price increase on OpenRouter (qwen/qwen3.6-27b): $0.25 → $0.32 in, $2.00 → $3.20 out per million tokens (+60%), first seen 26 Oct 2026.

WAS THIS USEFUL? 👍 👎

// END OF TRANSMISSION

The AI Toolchain · curated daily · October 27, 2026
Releases & features from the open-source and commercial AI tools that practitioners actually run.

▷ Subscribe Feedback on this issue