THE AI TOOLCHAIN

NO. 094

Tail

THE DAILY RELEASE FIREHOSE

PUBLISHED NOVEMBER 18, 2026 · EVERY DAY

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

▸ // HOW THIS ISSUE IS MADE SHOW

VIEW

ISSUE VIEW

scroll



SCROLL | EXPAND IN PLACE

INFORMATION DENSITY

everything



BRIEFED | EVERYTHING

Every tool in the issue, in full.

all 757 changes

Do you prefer **scroll**? 👍 👎

▸ // CONTENTS SHOW 138 of 145 tools

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

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

Three unrelated things stand out today: pymilvus made nested and scalar retrieval schemas easier to evolve without reloading everything;

vim ? help

Microsoft Foundry turned custom agent code into a hosted workflow; lightrag now refuses to start when embedding-model drift would corrupt an existing workspace.

BUILD

Evolve retrieval schemas without rebuilding the whole store

Milvus added null-aware StructArray fields, path-based field changes, broader scalar indexing, and partial Array-field upserts. Teams can change nested metadata and update array fields in place instead of treating schema growth as a full reingestion project.

pymilvus

DEPLOY

Host custom agent code without rebuilding the serving stack yourself

Microsoft Foundry added a bring-your-own-code hosted agent workflow through the Azure Developer CLI, while Anyscale added KubeRay Connect for registering self-managed Kubernetes clusters as Anyscale clouds with RayJob submission and observability. Both reduce the glue work between agent code and production execution environments.

Microsoft Foundry · Anyscale

GOVERN

Stop bad retrieval state before it reaches users

LightRAG now detects embedding-model drift and refuses to start rather than mixing incompatible vectors in one workspace. knowhere blocks oversized or unreadable documents before billing or model calls, preventing ingestion jobs from wasting spend on inputs the system cannot safely process.

lightrag · knowhere

OPERATE

Keep live AI systems observable when the network or provider path gets messy

Kiro added real-time OpenTelemetry event streaming with private AWS PrivateLink/VPC Lattice connectivity for Enterprise observability. FailproofAI made telemetry delivery resilient to outages, credential refusals, and partial acknowledgements with retry budgets, so traces are less likely to disappear exactly when an incident needs them.

Kiro

Does **GPT-5.5** read better? 👍 👎

BUILD

◆ AI Coding Agents

Warp

Feedback ⓘ 2026.11.12 (v0.2026.11.11.08.17) NOTES ↗ RANK

Warp is an AI terminal that runs commands and helps developers build software.

> Warp's agent tooling gained visibility into MCP tool calls, more flexible Factory checkout timing, and richer failure context passed between parent and child agents.

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

01 **Richer parent-child agent failure context** Not checked yet

IMPROVED 60

Parent agents now receive child failure, blocked, and cancellation context instead of status-only lifecycle events, giving parents more detail to react to subtask problems.

– Clear before/after described but no concrete API or config surface.

2026.11.12 (v0.2026.11.11.08.17)

THINNER COVERAGE BELOW

02 **MCP tool calls shown in run transcripts** Not checked yet NEW

50

Warp now shows MCP tool calls from Codex, Claude Code, and Pi cloud agents in run transcripts alongside Warp's own agent calls, giving a unified view of tool activity across agent types.

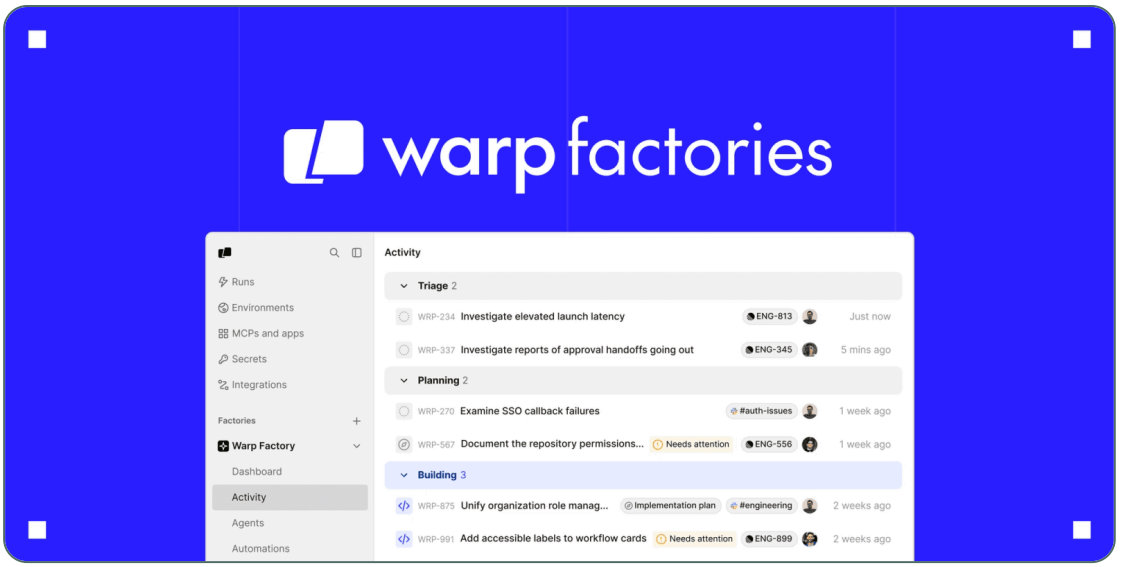
– Names the agents and feature but no UI path or mechanism.

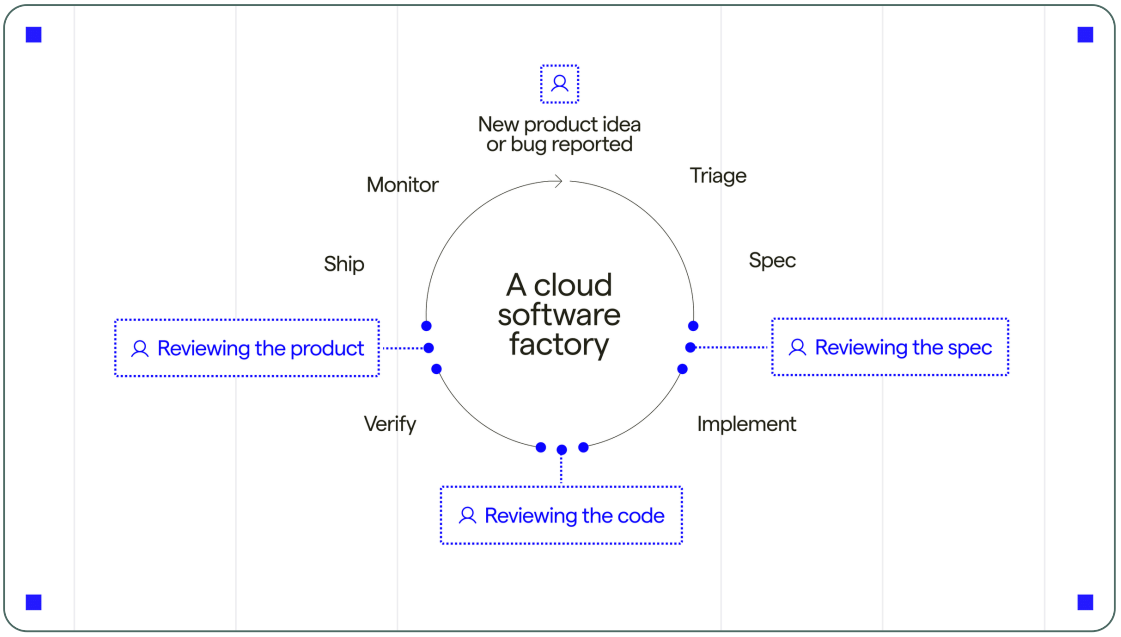
2026.11.12 (v0.2026.11.11.08.17)

03 **Deferred repository checkouts in Factories** Not checked yet

IMPROVED 45

Factories now support deferring repository checkouts in factory files until an agent actually needs them, rather than checking out repositories upfront.







— Describes the change but gives no config key or flag to use it.

2026.11.12 (v0.2026.11.11.08.17)

WAS THIS USEFUL? 👍 👎

SST OpenCode

Feedback ⓘ updated · Nov 18, 2026 NOTES ↗ RANK

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

> OpenCode introduces an in-app upgrade path to OpenCode 2.0, with background download and deferred install options, while warning that existing plugins will break under the new version.

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

01 **In-app upgrade to OpenCode 2.0** Not checked yet NEW 60

OpenCode now offers an in-app upgrade to OpenCode 2.0 that downloads in the background and installs only when the user chooses Restart to upgrade, Upgrade on quit, or Later, preserving existing sessions, projects, and settings.

– Names the three upgrade options but no commands or flags. snapshot-20261118

THINNER COVERAGE BELOW

02 **Plugin compatibility break in 2.0** Not checked yet BREAKING 25

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

– States the break but no migration guidance given. snapshot-20261118

└──▷ **BREAKING ON UPGRADE**

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

WAS THIS USEFUL? 👍 👎

Anthropic Claude Code

Feedback ⓘ v2.1.327 NOTES ↗

CODE ↗ RANK

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

> Claude Code's latest releases expand MCP server configuration and administration signals, add billing attribution and OpenTelemetry host tagging, broaden `/status` diagnostics, and ship smaller CLI, plugin and mod-authoring improvements.

└──▷ **WHAT SHIPPED · 13 FEATURES** most completely described first ? what's the number?

01 **OpenTelemetry host attributes** Not checked yet NEW 83

Setting the `OTEL_RESOURCE_INCLUDE_HOST` environment variable (off by default) includes `host.name` and `host.id` in OpenTelemetry resource attributes, useful when aggregating telemetry from multiple Claude Code instances.

Enrich OpenTelemetry traces with the host machine's name and ID — useful when aggregating telemetry from multiple Claude Code instances.

```
$ OTEL_RESOURCE_INCLUDE_HOST=true claude
```

— Env var, default, and runnable example given. v2.1.326

02 **Billing group attribution for sessions** Not checked yet NEW 81

The `attributionProject` setting in `.claude/settings.json` names a session's billing group, for Claude apps gateways that set `admin.billing_group_tag`.

Tag a session's API spend to a named billing group so gateway cost reports break down by project.

```json
{
  "attributionProject": "platform-team"
}
```

— Config key, file path, and example all given. v2.1.327

03 **Tool-rule startup warnings** Not checked yet NEW 81

Claude Code now warns at startup when an `--allowedTools`, `--disallowedTools`, or `--tools` value contains an entry that doesn't look like a complete rule, catching misconfigured tool allow/deny lists before a session begins.

Catch misconfigured tool allow/deny rules before a session begins by running Claude Code with explicit tool lists and observing startup warnings.

```
$ claude --allowedTools 'Bash(git *)' --disallowedTools 'Edit'
```

— Flags named with a runnable example demonstrating the warning. v2.1.326

---

04 **MCP server config via stdin piping** Not checked yet NEW

80

`claude mcp add-json <name> -` can now read a server's configuration from stdin, keeping secret tokens off the command line and out of shell history; it is not available inside an existing Claude Code session.

Add an MCP server whose config contains a secret token without exposing it in shell history or process listings.

```
$ cat mcp-server-config.json | claude mcp add-json my-server -
```

— Runnable command shown with its stated limitation. v2.1.327

---

05 **Expanded /status diagnostics** Not checked yet IMPROVED

73

`/status` now lists broken agent, skill, command and output style files, keybinding errors, an unavailable sandbox, ignored permission settings, and hook matchers that never match, plus a row indicating when your provider or an environment variable such as `DISABLE_TELEMETRY` turns features off.

— Command named with an exhaustive list of new diagnostics. v2.1.327 v2.1.326

---

06 **MCP administration and debugging signals** Not checked yet NEW

60

Gateway hint headers now include `x-claude-code-managed-settings-checksum` when managed settings set `CLAUDE_CODE_GATEWAY_HINT_HEADERS=1`; separately, Claude Code now explains why MCP servers stop loading when an administrator adds `managed-mcp.json` during a session.

— Two admin-facing MCP signals named, no runnable example. v2.1.327 v2.1.326

---

07 **New mod UI layout props** Not checked yet NEW

60

Mod authors can now use `overflowX="scroll"` on a mod's `Box`, and `maxLines` and `wrap="clip"` on its `Text`.

— Exact prop names given for mod developers. v2.1.326

THINNER COVERAGE BELOW

08 **Malformed plugin.json field notices** Not checked yet IMPROVED 50

`/plugin` and `claude plugin list` now show a note when a malformed `plugin.json` field such as `author` or `homepage` is ignored.

— Commands and field names given, no example provided. v2.1.326

09 **Confirmation before GitHub app install changes** Not checked yet IMPROVED 45

Claude Code now asks for confirmation before `/install-github-app` makes changes to a repository.

— Command named with a simple safety behavior. v2.1.327

10 **Code Mode compatible with agent teams** Not checked yet IMPROVED 45

Code Mode can now be enabled while agent teams are on, and sessions with teams now start in Code Mode automatically if it was already enabled.

— Behavior described but no config surface named. v2.1.326

11 **/tui classic renderer alias** Not checked yet NEW 40

`/tui classic` selects the non-fullscreen renderer, alongside the existing `/tui default`.

— Bare command naming with no further mechanism. v2.1.327

12 **Ctrl+O reveals hidden startup notices** Not checked yet NEW 35

Pressing Ctrl+O now shows hidden startup notices.

— Single keybinding, no further context given. v2.1.326

13 **/config override note** Not checked yet IMPROVED 30

`/config` now shows a note when another setting overrides a given row.

— Single-line behavior note, no mechanism given. v2.1.327

WAS THIS USEFUL?  

Block Goose

Feedback ⓘ v1.60.0 NOTES ↗ CODE ↗ RANK

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

> Goose v1.60.0 expands its oi-kit operation set with new skill, slash-command, bang-command and steering operations, adds reusable turn limits in the GDK, introduces Cortecs and Onysoft as declarative providers, and adds Dutch UI support, while removing the `goose review` subcommand and raising the desktop app's macOS requirement to 13.

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

01 **New oi-kit agent operations** Not checked yet NEW 75

Adds `SkillOperation` (replacing the previous Skills extension), `SlashCommandOperation` (with executable commands for running named slash commands), `BangCommandOperation` for invoking bang-prefixed shell commands, and `SteeringOperation` for redirecting or steering agent behavior mid-run, all to the oi-kit operation set.

— Names each new operation but gives no usage syntax v1.60.0

THINNER COVERAGE BELOW

02 **Removal of `goose review` subcommand** Not checked yet BREAKING 55

Removes the `goose review` subcommand; the local code-review workflow is no longer available, and any scripts or workflows calling it will fail after upgrading.

— Exact subcommand named with clear upgrade impact v1.60.0

03 **Cortecs and Onysoft declarative providers** Not checked yet NEW 50

Adds Cortecs and Onysoft as declarative providers, enabling Goose to route requests through the Cortecs and Onysoft inference backends.

— Named providers but no setup instructions given v1.60.0

04 **Reusable turn limits in GDK** Not checked yet NEW 45

Adds reusable turn limits in the GDK so agent developers can cap conversation length without reimplementing the logic per operation.

— No config key or default limit specified v1.60.0

05 **Raised macOS minimum version to 13** Not checked yet

BREAKING

45

Raises the desktop app's macOS minimum version requirement to 13; the app no longer runs on macOS 12 or earlier.

— Exact version threshold stated v1.60.0

06 **Dutch UI language support** Not checked yet NEW

25

Adds Dutch (nl) as a supported UI language.

— Bare addition with no further detail v1.60.0

└──▷ **BREAKING ON UPGRADE**

! The `goose review` subcommand has been removed; any scripts or workflows that call it will fail after upgrading.

! The macOS minimum version is raised to 13; the desktop app no longer runs on macOS 12 or earlier.

WAS THIS USEFUL? 👍 👎

yetone magpie

Feedback ⓘ v0.1.4402 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 addition this window is per-project instructions management via `magpie library project instructions`, alongside new SSH proxy support with host-key pinning, cross-machine Claude plugin sync, decision groups for gateway keys, skill sync by source, Codex web.run relay routing, a manual relay check with a provider speed breakdown, and a fix that stops `/api/state` from leaking proxy secrets.

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

01 Per-project instructions via `magpie library project instructions`

Not checked yet NEW 95

Set per-project instructions for specific agents with `magpie library project instructions <dir> text <text…>`, `--file <f>`, or `-` for stdin; magpie writes them between its own marker lines into the file each agent reads (e.g. `CLAUDE.md`, `AGENTS.md`, `GEMINI.md`, `QWEN.md`, `CODEBUDDY.md`) without touching the rest of the file. Control which agents receive a project's instructions with `magpie library project instructions <dir> agents <a,b…|none>`, which defaults to the agents the shared instructions are already on for, or every agent in the project that reads an instructions file. Inspect a project's current instructions and their per-agent file targets with `magpie library project instructions <dir> [--json]`.

Read project instructions from a file (or stdin with `-`) when the text is long or generated by another tool.

```
$ magpie library project instructions /path/to/myproject text --file project-rules.md
```

Review which agents are receiving the project's instructions and which files they were written into.

```
$ magpie library project instructions . --json
```

— Full command set with flags and runnable examples v0.1.4390

02 **Relay routing for Codex web.run search** Not checked yet NEW

90

Relay providers configured with `search=yes` now also receive Codex's `web.run` search requests at `/alpha/search`, in addition to Claude Code's `WebSearch` and Codex's `web_search`, so a relay can answer all three web-search surfaces itself without magpie substituting its own search.

Tell magpie that a relay handles all web-search traffic — including Codex's web.run `/alpha/search` — so requests pass through unmodified instead of falling back to magpie's own search.

```
$ magpie provider set my-relay search=yes
```

— Runnable CLI example plus exact endpoint named v0.1.4391

03 **SSH proxy support with host-key pinning** Not checked yet NEW

75

A saved proxy with an `ssh://` address can now tunnel traffic without a local SSH client. Trust-on-first-use host key pinning writes the server's host key into the saved proxy's `HostKey` field; a server presenting a different key is refused until the user re-trusts it from the proxy editor. Failed proxy connections now report the reason for failure in the settings UI.

— Names config field and UI flow, no CLI example v0.1.4395

04 **Decision groups for gateway keys** Not checked yet NEW

75

Gateway keys restricted to specific models can now be assigned a decision group, letting them reach that group at `/v1/systemone`. The key editor lists active decision groups (non-disabled, non-hidden) alongside chat groups so operators can select them when scoping a key's model access.

— Endpoint named but no example to try v0.1.4399

05 **Relay check and provider speed breakdown** Not checked yet NEW

75

Checking a single model on a single provider now triggers real `RelayPartCount`, `RelayPartMath`, and `RelayPartTool` calls through the gateway to surface reply identity, token counts, answer accuracy, tool-call success, and speed — activated only when the user presses Check. The usage speed ranking now also exposes `ProviderSpeed`, a per-provider breakdown of a model's speed alongside `Tries`, `Failed`, and `Swapped` counts.

— Names specific call types and fields, UI-only action v0.1.4388

06 **Cross-machine Claude plugin sync** Not checked yet NEW

65

Claude plugins installed on one computer now sync whole to other computers, with each computer's hooks using its own local folder path instead of the source machine's paths. Implemented via the new `CarriedPlugin` type in `internal/library/plugin_carry.go`.

— Names internal implementation, no user command given v0.1.4393

07 **Skill sync by source across computers** Not checked yet IMPROVED

65

Skills installed from GitHub or a git repository now sync as their source rather than their contents, so each computer fetches the skill itself and the skill's size no longer counts against the 64 MiB sync limit. Skills that can't be carried during a sync (too large or unfetchable) are now reported explicitly instead of silently omitted.

— Explains mechanism and limit, no command shown v0.1.4396

08 **Secrets removed from `/api/state` proxy payload** Not checked yet

IMPROVED

65

The `/api/state` endpoint (and `?last=1`) no longer sends a saved proxy's password, the LAN key (`LANKey`), Bark key (`BarkKey`), or GitHub token (`GitHubToken`) in its settings payload.

— Endpoint and exact fields named, no user action needed v0.1.4402

WAS THIS USEFUL? 👍 👎

Letta Code

Feedback ⓘ v0.34.65 NOTES ↗ RANK

Letta Code is a terminal coding agent with persistent memory and identity across development sessions.

> Letta Code gated a new managed cloud watch feature for PR monitoring behind a feature flag, and improved subagent orchestration and export safety.

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

01 **Managed cloud watches for PR monitoring** Not checked yet NEW

35

A feature-flagged `watch-pr` capability adds managed cloud watches for monitoring pull requests.

— Names the flag but no mechanism detail v0.34.65

02 **Subagent result reminders** Not checked yet NEW

35

Agents now support optional subagent result reminders, surfacing subagent outputs back to the orchestrating agent after task completion.

— Explains behavior but no config surface named v0.34.63

03 **XML text escaping in channel exports** Not checked yet IMPROVED

30

Channel exports now apply XML text escaping to exported content.

— Clear change but minimal mechanism detail v0.34.63

WAS THIS USEFUL? 👍 👎

mvschwarz OpenRig

Feedback ⓘ v0.7.18 NOTES ↗ RANK

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

> OpenRig's latest releases center on making seat handovers and runtime state recoverable — a `--recover` flag for failed handovers, in-place runtime-home recording with its own stale-switch recovery, and Pi-aware compaction state checks — alongside queue return-routing, Claude auto mode, Linux task limits, remote-seat TUI visibility, and clearer status messaging across restore, queue, and network commands.

└──▷ **WHAT SHIPPED · 18 FEATURES** most completely described first ? what's the number?

01 Seat runtime-home management and recovery

Not checked yet NEW

95

`rig seat runtime-home <seat>` shows a seat's next-launch runtime home, and `rig seat record-runtime-home <seat> --home <abs> --expect-occupant <gen> --expect-original <id> --expect-revision <n> --reason "<why>"` records it in-place without a relaunch. `rig seat set-runtime <seat> --recover` settles a stale `rig seat set-runtime` switch that was never committed, for use after a host reboot or tmux server restart.

Record the runtime home of an existing conversation in-place so the seat's next launch uses it, without triggering a relaunch.

```
$ rig seat record-runtime-home my-seat --home /Users/me/.claude --expect-occupant gen-42 --expect-original conv-abc123 --expect-revision 3 --reason "realign after manual move"
```

— Three exact commands with flags and recovery scenario. v0.7.16

02 Seat handover recovery

Not checked yet IMPROVED

85

`rig seat handover <seat> --recover` adopts the running successor of a failed handover by proving its process and identity against the failed attempt's record and recording its conversation as the seat's resume target; the same flag was later made to resume an interrupted Claude handover more accurately.

Recover from a failed seat handover by adopting the already-running successor process without manual intervention.

```
$ rig seat handover my-seat --recover
```

— Full mechanism and exact command given, accuracy improved later. v0.7.18 v0.7.16

---

03 **Seat compaction controls** Not checked yet IMPROVED 85

`rig compact <seat> --skip-map` force-compacts a seat at Claude Code's Context limit reached screen even when no context-usage reading is available; `rig compact <seat> --state` checks and reports the compaction attempt's status, including for Pi seats.

Force-compact a seat that has hit Claude Code's context limit when no usage reading is available.

```
$ rig compact my-seat --skip-map
```

— Exact flags and scenario given for both compaction states. v0.7.18 v0.7.16

---

04 **Queue return-route, tagging and successor tracking**

Not checked yet NEW 75

`rig queue create --return-route auto` or `--return-route self` hands work over and routes prompts back into your own conversation instead of leaving them in the terminal; `rig queue tag` adds tags to an existing queue row; and `rig queue list` now names a handed-off row's successors.

Receive your team's seat prompts back in your own conversation instead of leaving them in the terminal, so you can review and approve without switching context.

```
$ rig queue create --return-route auto
```

— Three named queue commands/flags grouped, each with a stated effect. v0.7.18

---

05 **New per-seat controls: input routing, auto mode, and task limits**

Not checked yet NEW 70

Adds new per-seat controls: native input routing via `rig seat set-input-route <seat> --route native --reason <text>` instead of terminal typing; Claude `autoMode` per seat via `claude_auto_mode` in seat config, or toggled at runtime with `rig seat set-permissions --auto-mode`; and a Linux systemd task limit per seat via `process_limits: { tasks_max: N }` in seat config, capping how many tasks a seat's systemd slice can spawn.

Switch a seat to native input routing when terminal-typing delivery is unreliable or you need programmatic control over what the agent receives.

```
$ rig seat set-input-route dev.impl --route native --reason
'programmatic input required'
```

Cap the number of tasks a seat's systemd slice can spawn on Linux, preventing runaway agent processes from starving the host.

```yaml
process_limits:
  tasks_max: 64
```

— Exact command given but mechanism of native routing unexplained. Named config key and flag, no behavioral detail on autoMode itself. Exact config key and purpose given.

v0.7.18

---

06 **rigs skill prompt routing and continuation** Not checked yet

IMPROVED

70

The `rigs` skill now routes a seat's permission, trust, or hook-review prompt to the person as a question by default instead of answering it autonomously; after the person answers a startup prompt, the skill tells their agent to run `rig seat continue <seat>` when `rig ps --nodes` shows Startup details for that seat ending in that command.

— Behavior change explained with trigger condition and follow-up command. v0.7.16

---

07 **Per-seat checkpoint file naming** Not checked yet BREAKING

65

Per-seat checkpoint files are now stored as `.rigged-checkpoint-<session>-<digest>.md` per session, replacing the shared `.rigged-checkpoint.md`; anything that reads the old shared path won't find restored checkpoints.

— Exact old and new filenames named, migration impact noted. v0.7.18

---

08 **CLI status message improvements** Not checked yet IMPROVED

65

An empty person registry now reads `undeliverable:no-registered-human` instead of silently failing; work for a hosted member with no wake trigger now reads `awaiting-pickup:` instead of `unroutable:`, waiting in its queue until the member reads it; `rig network status` text output now hides a hub URL's query string (`--json`

output is unchanged); and an external connection's `rig_queue_show` now labels who acted apart from the reader.

— Four named status strings/outputs grouped with exact before/after wording. v0.7.17

09 **Single seat launch into a rig** Not checked yet IMPROVED

65

Launch a single seat into a rig with `rig launch <rig> --seats <seat>` without being blocked by other paused or actively typed seats in the same rig.

— Exact command and flag with stated before/after behavior. v0.7.16

10 **Remote seat visibility in TUI** Not checked yet NEW

60

Enable reading a remote rig's seats in the TUI with the `ui.topology_remote_seats` config key, off by default.

Turn on remote-seat visibility in the TUI so you can inspect seats on a rig running on another host.

```yaml
ui.topology_remote_seats: true
```

— Exact config key and default state given. v0.7.17

THINNER COVERAGE BELOW

11 **Daemon HTTP routes for resume and maintenance pause**

Not checked yet NEW

55

Two new opt-in daemon HTTP routes let you resume one banked original conversation and pause automatic delivery for one rig during maintenance until an explicit end; documented in `docs/reference/daemon-http-clients.md`.

— Named doc path but no concrete endpoint paths given. v0.7.17

12 **MCP front-door mode restriction** Not checked yet IMPROVED

55

MCP front-door mode via `rig mcp serve --front-door` now restricts autonomous request pickup to running team seats only.

— Exact command named with scope restriction described. v0.7.16

13 **Private Slack channel for prompt delivery** Not checked yet NEW

50

Set up a private Slack channel for prompt delivery with `rig slack manifest --private-channel`.

— Exact command but no further mechanism described. v0.7.18

14 **Workflow command enhancements** Not checked yet IMPROVED 50

`rig workflow route` now adopts a second handoff to the same target, and `rig workflow revise` can now drop an obsolete required order from unstarted work.

— Named commands but behavior described only briefly. v0.7.17 v0.7.18

15 **Restore status improvements** Not checked yet IMPROVED 45

Restore now reads `partially_restored` when a seat was skipped or left attached, and names the old and new addresses when an adopted seat's session name can't be preserved.

— Named status value and scenario but no command to trigger. v0.7.17

16 **Roster label problem handling** Not checked yet IMPROVED 45

Roster labels over 200 characters or containing disallowed characters are now set aside with `label_problem` rather than failing the full member or roster load.

— Named status value and limit, but no command to inspect it. v0.7.16

17 **Bundle check schema validation** Not checked yet IMPROVED 35

`rig bundle check` now reports schema-invalid configuration names.

— One-line improvement with named command, no further detail. v0.7.17

18 **Queued sends to missing seat destination** Not checked yet IMPROVED 30

Sends to a missing seat destination are now queued and acknowledged rather than dropped.

— Behavior change stated but no command or config surface. v0.7.16

└──▷ BREAKING ON UPGRADE

! Each restored seat's checkpoint is now written to `.rigged-checkpoint-<session>-<digest>.md`; anything that reads `.rigged-checkpoint.md` won't find it.

! Four database migrations (two numbered 171, then 172 and 173) run before the daemon listens; rolling back to 0.7.17 after they run is not established — keep the pre-upgrade backup.

! openrig-core moves to 0.1.32 because its skills and refocus hook changed.

! Two database migrations (168 and 170) run before the daemon listens on startup; downgrading to 0.7.16 after applying them is not established — keep the backup created

during upgrade.

! Migration 167 adds the `seat_handover_attempts` table (for `rig seat handover --recover`) and migration 169 adds `seat_runtime_home_records` (for `rig seat record-runtime-home`); downgrading to 0.7.15 after these migrations runs is not established — keep the pre-upgrade database backup.

! `openrig-core` moves to 0.1.31 because its skills changed; the `openrig` plugin carries CLI version 0.7.16.

WAS THIS USEFUL? 👍 👎

agents-cli

Feedback ⓘ v1.15.0 NOTES ↗ RANK

The CLI and skills that turn any coding assistant into an expert at creating, evaluating, and deploying AI agents on Google Cloud.

> agents-cli v1.15.0 adds Cloud SQL-backed session storage for the LangChain extension and relocates first-party extensions to a new PyPI package, along with a Go scaffolding change for app-info serving.

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

01 **Go app-info serving moved to ADK Go package** Not checked yet

85

BREAKING

Go projects now serve `/apps/{app_name}/app-info` from ADK Go's `exp/appinfo` package instead of a local `appinfo/` folder. After running `agents-cli scaffold upgrade`, change the import `"<project>/appinfo"` to `"google.golang.org/adk/v2/exp/appinfo"` in any customized `main.go`.

— Names endpoint, old and new import paths, exact migration step v1.15.0

02 **Cloud SQL session storage for LangChain extension** Not checked yet

NEW

75

Store LangChain extension conversation checkpoints and long-term memory in Cloud SQL for PostgreSQL by passing `--session-type cloud_sql` to the template.

— Names exact flag and value, runnable from this alone v1.15.0

03 **First-party extensions moved to new PyPI package** Not checked yet

65

BREAKING

Install first-party extensions and templates from the new `google-agents-cli-extensions` package on PyPI, which no longer requires a trust prompt. Extensions installed from `google/agents-cli` now show a deprecation warning; install from `google/agents-cli-extensions` instead.

— Names old and new package sources, no migration command given v1.15.0

└──▷ **BREAKING ON UPGRADE**

! First-party extensions installed from `google/agents-cli` now show a deprecation warning; install from `google/agents-cli-extensions` instead.

! Go projects: change the import `"<project>/appinfo"` to `"google.golang.org/adk/v2/exp/appinfo"` in any customized `main.go` after running `agents-cli scaffold upgrade`.

WAS THIS USEFUL? 👍 👎

GitHub Copilot CLI

Feedback ⓘ v1.0.113 NOTES ↗ RANK

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

> GitHub Copilot CLI now lets administrators disable the built-in Browser Use feature via a managed setting, including mid-session.

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

01 **Admin control to disable Browser Use** Not checked yet NEW

60

Administrators can disable the built-in Browser Use feature by setting `features.browserUse` in managed settings, and this control applies even during active running sessions.

— Names exact config key but no file path or scope details v1.0.113

WAS THIS USEFUL? 👍 👎

Graphify Labs Graphify

Feedback ⓘ v0.9.112 NOTES ↗

RANK

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

> Graphify shipped a new section-reading command/MCP tool and a major memory optimization for Python resolution, with a breaking AST cache schema bump affecting several language extractors.

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

01 **Read markdown sections via CLI and MCP** Not checked yet NEW

79

The new `graphify section "<heading>"` command and `get_section` MCP tool let you read any markdown heading as its complete section body, derived from the source file at read time without touching the graph or cache; paths outside the project are refused.

— Names exact command and MCP tool with clear behaviour v0.9.110

02 **Lower-memory Python resolution** Not checked yet IMPROVED

70

Python resolution no longer memoizes parse trees across resolution passes, cutting peak memory by roughly 1.6–2 GB on mid-size repos and ~70% on large ones while producing byte-identical graph output. This bumps the AST cache schema to 47, so warm caches will re-run the affected JS/TS, SQL and nested-type (C#, Java, Scala, C++, TS) extractions.

— Concrete numbers and mechanism, but no user action beyond cache rebuild v0.9.110

NEW SUPPORTED PLATFORM Delphi form-to-unit graph linking

└──▷ BREAKING ON UPGRADE

! The AST cache schema is bumped to 47; warm caches re-run the affected JS/TS, SQL and nested-type (C#, Java, Scala, C++, TS) extractions.

> v0.9.112 NOTES ↗ LATEST STABLE
>
> Graphify adds Delphi form-to-unit linking, wiring `.dfm` / `.lfm` form classes and event handler stubs to sibling `.pas` units via INFERRED `references` edges.
>
> └──▷ GET THIS VERSION ▾

pipx

```
$ pipx install graphifyy==0.9.112
# already installed? upgrade to this version:
$ pipx install --force graphifyy==0.9.112
```

git

```
$ git clone --branch v0.9.112 https://github.com/Graphify-Labs/graphify.git
# already have the repo? check out this version:
$ git checkout v0.9.112
```

PyPI attestation · checked 18 Nov Wrong command?

Graphify now links Delphi forms to their implementing units: each `.dfm` / `.lfm` form class and `OnClick = Handler` stub gets an INFERRED `references` edge to the real class and method in the sibling `.pas` file, matched by base name and unique class/method lookup. Not checked yet

└──▷ **BREAKING ON UPGRADE**

! The AST cache schema is bumped to 48; warm caches will re-run extractions for Go, Java, JS/TS, C/C++, SQL, and Pascal.

WAS THIS USEFUL? 👍 👎

wrongstack

Feedback ⓘ v1.7.9 NOTES ↗ RANK

WrongStack is an AI coding agent that reads code, edits files, runs commands, and reasons through bugs across CLI, TUI, and browser interfaces.

> WrongStack's latest releases add a session inspector and desktop project management, automatic daemon replacement across its Chronicle, Kanban, Session Catalog, SAGE, and mailbox services, and broader dependency-manifest parsing alongside TUI input fixes for browser terminals.

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

01 **TUI icon rendering control** Not checked yet NEW 75

The `icons` config key accepts `auto`, `nerd`, or `ascii`, letting terminals without a Nerd Font render plain glyphs instead of broken icons.

Set TUI icons to plain ASCII on a server terminal that lacks a Nerd Font, so glyphs render correctly.

ini

```
icons = ascii
```

— Names exact config key, accepted values, and a runnable example v1.7.9

02 **Broader file and dependency manifest parsing** Not checked yet

IMPROVED 70

WrongStack now reads BOM-prefixed locks, registries, prompts, caches, manifests, and Kanban history files that previously failed to load, and reports malformed JSON roots rather than silently skipping them. It also parses Cargo virtual workspaces, wheel and sdist names, PEP 508 ranges with commas, and pubspec blocks at any indent as real dependency manifests.

— Names every file and manifest format affected v1.7.6

03 **Automatic daemon replacement across services** Not checked yet

NEW 65

A newer client now replaces Chronicle, Kanban, Session Catalog, SAGE, and mailbox daemons left by an older release, while daemons of the same or newer release are reused rather than replaced.

— Names all affected daemons and the replacement/reuse logic v1.7.3

THINNER COVERAGE BELOW

04 **TUI input fixes for browser terminals** Not checked yet IMPROVED

55

TUI input in browser terminals now handles Home/End keys correctly, survives split escape sequences, and treats pastes as data until Enter is pressed.

– Names specific input behaviors fixed but no command v1.7.6

05 **Session inspector for worker transcripts** Not checked yet NEW

45

A new session inspector surfaces each worker's transcript and plan alongside the leader transcript, without replacing the leader view.

– Explains what it shows but no path or command given v1.7.9

06 **Project reordering in Desktop menu** Not checked yet NEW

45

Projects can now be reordered or removed in the Desktop left menu without affecting their underlying files or history.

– Names the UI location but no further mechanism v1.7.9

07 **Stopped conversations resume via spare engine** Not checked yet

IMPROVED

40

Stopped conversations now resume using the spare engine already waiting in the conversation folder instead of booting a new one.

– Describes mechanism but no config or command to invoke v1.7.9

08 **BTW notes persist to the leader** Not checked yet IMPROVED

35

A BTW note now reaches the leader within the same session and is no longer discarded at the end of a turn.

– States before/after behavior but lacks usage detail v1.7.9

09 **Formatter checks report failure exit code** Not checked yet

IMPROVED

35

Formatter checks now report their exit code on failure instead of silently passing.

– Brief fix description with no further detail v1.7.6

WAS THIS USEFUL?  

OpenQodex

Feedback ⓘ openqodex@0.19.0 NOTES ↗ RANK

OpenQodex is an open source AI code review for coding agents. A separate reviewer checks every change your agent writes before you push, in Claude Code, Codex, Cursor and Cline.

> OpenQodex 0.19.0 adds cross-file evidence to findings, elevates convention-file violations to findings, and upgrades the bundled ruff linter to 0.17.3 with a much larger default rule set.

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

01 **Bundled ruff linter upgraded to 0.17.3** Not checked yet BREAKING

95

The bundled ruff linter moves to version 0.17.3, expanding the default rule set from 59 rules to 407 and adding security rules `S102`, `S110`, and `S112` ranked as high severity; repositories with no ruff `select` configured will see more findings. Add `select = ["E4", "E7", "E9", "F"]` to your repository's ruff config to keep the previous 59-rule default. Separately, when a `pyproject.toml` has no `[tool.ruff]` table, the linter now reads `requires-python` from that file as its target version, so changing that value triggers the ruff settings note.

Preserve the previous 59-rule ruff default set so your existing scans are not flooded with new findings after upgrading to 0.19.0.

```toml
select = ["E4", "E7", "E9", "F"]
```

— Exact version, rule counts, rule IDs and migration config given openqodex@0.19.0

02 **Cross-file evidence in findings** Not checked yet NEW 80

When a finding's evidence spans more than one file, OpenQodex now surfaces it in the `evidence` field of `report.json` and in a new Evidence in other files section of the report. Callers are drawn from the code graph (up to 8 per declaration), and verifier input is pulled from the cited lines in those other files.

— Names exact field, report section and caller limit openqodex@0.19.0

03 **Convention-file violations now reported as findings**

Not checked yet IMPROVED 75

Violations of rules in files listed under `review.instructions_files` are now reported as findings, carrying a `rule` field in `report.json`, instead of being downgraded to observations. Style, naming, docs, and test-coverage items still remain observations.

— Names config key and new report field precisely `openqodex@0.19.0`

└──▷ **BREAKING ON UPGRADE**

! The bundled ruff linter's default rule set expands from 59 rules to 407; repositories with no ruff `select` configured will receive more findings, including high-severity security rules `S102`, `S110`, and `S112`. Add `select = ["E4", "E7", "E9", "F"]` to your ruff config to preserve the previous behaviour.

WAS THIS USEFUL? 👍 👎

Alibaba Qwen Code

Feedback ⓘ desktop-v0.28.0 NOTES ↗

RANK

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

> Qwen Code added `qwen4-max` model support with new effort, cache-TTL and image-detail controls, extended managed-channel integrations to DingTalk and Aone Code, enabled external workspace execution with a `create_sub_session` capability, and shipped a round of Web Shell session-view and managed worktree improvements.

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

01 **Web Shell session-view enhancements** Not checked yet IMPROVED

75

Adds a sidebar workspace filter, per-turn hook statistics display for inspecting hook timing, active highlight styling for selected reasoning steps, dirty-tree checkout refusal resolution in the branch picker, a session-artifact header entry before replay on cold open, activity-date session grouping, a delivery-unit change overview for sessions, and visible Goal state with session navigation for at-a-glance orientation.

— Lists many named UI additions without exact navigation steps. v0.28.0

desktop-v0.28.0

02 **External workspace execution and create_sub_session**

Not checked yet NEW

65

Enables external workspace execution for managed-agent sessions so agents can run in workspaces outside the host process, and adds a `create_sub_session` capability to run a sub-session in a separate workspace.

— Names a concrete capability but gives no usage example. desktop-v0.28.0

03 **Managed worktree sessions in Web Shell** Not checked yet NEW

60

Adds managed worktree sessions in the Web Shell that are default-on and detached, with branch selection and handoff, matching Codex-style isolated execution.

— Describes defaults and behaviour but no exact UI path. v0.28.0

THINNER COVERAGE BELOW

04 **DingTalk and Aone Code channel integrations** Not checked yet NEW

55

Adds a DingTalk adapter on the managed-channel path so DingTalk conversations can drive agent sessions, and an Aone Code channel plugin with a polling event-stream inbound, extending the managed-channel roster for chat-driven agent sessions.

— Names both integrations and a mechanism, no setup steps given. desktop-v0.28.0 v0.28.0

05 **Loop skill controls** Not checked yet NEW 55

Publishes the managed `/loop` skill variant by profile, and lets loop switches be forced off as well as on.

— Names the `/loop` skill but gives no usage example. desktop-v0.28.0

06 **qwen4-max model support** Not checked yet NEW 55

Adds `qwen4-max` model support with a native effort ladder, `cache TTL`, and image detail controls.

— Names model and controls but no config key given. v0.28.0

07 **Shared agent loop as default execution path** Not checked yet IMPROVED 40

Runs headless sessions, the Ink TUI, and subagents on the shared agent loop by default.

— States the default change with no config exposed. desktop-v0.28.0

08 **Russian localisation and local-files bridge in Web Shell** Not checked yet IMPROVED 35

Completes the Russian localisation catalog for the Web Shell UI and makes the local-files bridge easier to discover and use.

— Thin description with no mechanism or UI path. desktop-v0.28.0

09 **Cross-session messages for ACP sessions** Not checked yet IMPROVED 35

Lets program-driven ACP sessions accept cross-session messages.

— Brief statement with no example or mechanism. desktop-v0.28.0

10 **Card view state stored in session transcript** Not checked yet NEW 35

Stores card view state in the session transcript for generated UI.

— Brief statement with no mechanism or example. v0.28.0

11 **Compaction retains latest complete operation** Not checked yet IMPROVED 30

Retains the latest complete operation during compaction.

— States the behaviour with no further detail. desktop-v0.28.0

12 **Mods can rewrite expanded thinking in terminals** Not checked yet

NEW

25

Lets Mods rewrite expanded thinking in the terminals.

— Single sentence with no mechanism or scope. desktop-v0.28.0

WAS THIS USEFUL? 👍 👎

Yuval Sarel cones

Feedback ⓘ v0.4.19 NOTES ↗ RANK

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

> cones added native Grok harness support with live state, token, and cost tracking, a `cones adopt` command to reparent sessions, and a reorganized session list that nests by folder and tags, alongside fixes to auto-update reliability and tag persistence.

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

01 **Native Grok harness support** Not checked yet NEW 90

cones now treats Grok as a native harness: each live dashboard row shows the conversation, Grok's state (working, waiting at a permission prompt, or idle), model, effort, title, tokens, context window, and session cost. `cones show`, `cones search`, history, `resume`, and `fork` read and drive Grok's session folders, and `cones wait` now settles on Grok rows. Launching Grok from the CLI or composer draws its screen in the pane and auto-answers the folder trust prompt with `--trust`, and `GROK_HOME` plus the `homes.grok` config key point cones at a custom Grok home directory.

Point cones at a non-default Grok home by setting `homes.grok` in your cones config, then verify active Grok sessions with their state, tokens, and cost.

yaml

```yaml
homes.grok: ~/grok-workspace
```

— Names commands, config keys, and a worked example. v0.4.16

02 **Session reparenting with cones adopt** Not checked yet NEW 85

`cones adopt ID` reparents a running session under another, using `--parent ID` to nest it under a specific session (mirroring `cones launch --parent`), and `--remove` to stand an adopted session alone again.

Move a runaway top-level agent session under a lead orchestrator session so it appears in the right hierarchy.

```
$ cones adopt <session-id> --parent <lead-session-id>
```

Detach a previously nested session so it stands alone at the top level again.

```
$ cones adopt <session-id> --remove
```

— Exact command and flags with worked examples. v0.4.19

03 **Folder-and-tag nesting in session list** Not checked yet IMPROVED

70

The session list (ctrl+l) now organises sessions by folder, tags, or both: a folder alone keeps a session under its folder heading, tags alone float it above folders, and combining folder with tags nests those tags inside the folder heading. A session under a worktree heading now shows that worktree as its folder in ctrl+l, with its tags nested under the worktree's heading, and tag headings are drawn in cyan to distinguish them from folder headings.

— Explains rules and UI path but no runnable command. v0.4.15

THINNER COVERAGE BELOW

04 **Auto-updater runs brew update before upgrade** Not checked yet

IMPROVED

55

The background auto-updater now runs `brew update` before `brew upgrade cones` so a newly announced release is actually installed.

— Names exact commands fixing a real update bug. v0.4.14

05 **Persistent session tagging via state/tags.json** Not checked yet

NEW

45

cones now persists session tags across restarts, saving ctrl+p highlights to `state/tags.json`.

— Names the storage file but no usage detail. v0.4.14

06 **Inline keyboard shortcut list on empty composer** Not checked yet

NEW

35

Pressing `?` on an empty dashboard composer now lists the current heading's keyboard shortcuts inline, below the input field.

— Describes behaviour but no command to run. v0.4.16

07 **Persistent ctrl+p highlight across resume/fork** Not checked yet

IMPROVED

35

A ctrl+p highlight now stays on a conversation when Claude resumes or forks it under a new id.

— Behaviour named but mechanism unexplained. v0.4.15

08 **Orange post-update restart banner** Not checked yet IMPROVED

30

After an update installs, the header now shows an orange banner prompting you to quit and reopen cones, and the footer notice turns orange instead of dim.

— Visual change only, no interaction described. v0.4.15

WAS THIS USEFUL? 👍 👎

zot

Feedback ⓘ v0.4.82 NOTES ↗ RANK

zot manages coding agent workflows from the command line.

> zot shipped two small usability updates: session ID visibility and compaction health tracking in the status bar.

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

01 **Session ID display and copy** Not checked yet NEW

30

zot now shows and copies the current session ID, making it easier to reference or share an active session.

— Describes what it does but no mechanism or UI path v0.4.82

02 **Compaction fallback and failure tracking** Not checked yet NEW

25

zot now tracks compaction fallbacks and failures in the status bar, giving visibility into store compaction health at a glance.

— Names the feature but no specifics on triggers or values shown v0.4.80

WAS THIS USEFUL? 👍 👎

Alibaba OpenCodeReview

Feedback ⓘ v1.13.11 NOTES ↗

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

OpenCodeReview adds branch filtering for session listing.

└──▷ GET THIS VERSION ▾

Homebrew

```
$ brew install open-code-review
# already installed? upgrade to this version:
$ brew upgrade open-code-review
```

npm

```
$ npm install -g @alibaba-group/open-code-review@1.13.11
```

git

```
$ git clone --branch v1.13.11 https://github.com/alibaba/open-code-review.git
# already have the repo? check out this version:
$ git checkout v1.13.11
```

Homebrew lists this version · signed build provenance · checked 18 Nov Wrong command?

Filter listed sessions by branch when managing review history. Not checked yet

NEW DETECTION Move language review rules

WAS THIS USEFUL? 👍 👎

jcode

Feedback ⓘ v1.0.0 NOTES ↗ RANK

Jcode is a high performance coding agent harness written in Rust, built for heavy multi-agent dev workflows

> jcode's v1.0.0 release adds cloud offload for sessions and swarms, phased parallel swarm execution, account/usage management controls, and notable terminal performance and cloud-history loading improvements.

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

01 **Cloud offload via /cloud and /local** Not checked yet NEW

85

Move a session or coordinated swarm to a cloud VM with `/cloud`, keep working after disconnecting, and return the conversation and repository to your local machine with `/local`.

Offload a long-running agent session to a cloud VM so it keeps running after you close your laptop.

```
$ /cloud
```

Pull a cloud session back to your local machine, restoring the conversation and repository state.

```
$ /local
```

— Exact commands given with clear before/after behavior v1.0.0

02 **Terminal CPU and memory optimizations** Not checked yet IMPROVED

75

Terminal now uses 25% less CPU, cuts idle wakeups from 2.68 to 1.51 per second, and reduces rendered-span memory from 20.4 MB to 14.6 MB, enabling larger parallel agent runs without added machine load.

— Concrete before/after metrics but nothing for reader to invoke v1.0.0

03 **Account priority, hot-swap, and usage controls** Not checked yet

NEW

70

Set account priority tiers, hot-swap accounts with `Ctrl+Alt+J/K`, inspect refreshed `/usage` cards, and configure context limits per session.

— Named keybind and command give a clear starting point v1.0.0

04 **Lazy-loaded cloud session history** Not checked yet IMPROVED

65

Cloud sessions now load recent history first and fetch older messages on demand, shrinking a 7,000-message session's initial history payload from 14.3 MB to 261 KB.

— Clear mechanism and numbers, but passive improvement v1.0.0

05 **Typo-tolerant /resume search picker** Not checked yet IMPROVED

65

`/resume` now offers a typo-tolerant search picker for finding a previous session by keyword across both local and cloud history.

Find a previous session by keyword with typo tolerance across both local and cloud history.

```
$ /resume
```

— Named command with a direct usage example v1.0.0

THINNER COVERAGE BELOW

06 **Phased parallel swarm execution** Not checked yet IMPROVED

50

Swarms now plan jobs in phases and parallel lanes, automatically start workers as dependencies finish, and show team progress in one pinned view.

— Describes mechanism but no named flags or commands v1.0.0

07 **Voice input waveform and streaming transcript** Not checked yet

NEW

30

Live voice input now includes a waveform and a streaming draft transcript.

— Thin description with no mechanism or usage detail v1.0.0

WAS THIS USEFUL? 👍 👎

flow-next

Feedback ⓘ flow-next-v9.13.0 NOTES ↗ RANK

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

> flow-next's biggest addition this window is a secret-store-backed credential command for Jira and Linear, alongside deeper unattended automation, a full Jira Product Discovery idea-to-parent workflow, and new tracker CLI commands for triaging and bookkeeping open issues.

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

01 **Per-tracker credential command for Jira and Linear**

Not checked yet NEW 90

Setting `tracker.perTracker.tokenCommand` to an argv list (e.g. `op read`, `security find-generic-password`, or a vault CLI) lets flowctl pull Jira or Linear credentials from any secret store at runtime instead of a manually exported token. The command runs without a shell, at most once per process, with a 10-second timeout, and only when the matching environment variable (`LINEAR_API_KEY`, `JIRA_API_TOKEN`, or `JIRA_PAT`) is unset; its output never reaches a log, receipt, or `.flow/` file.

Store per-tracker credentials so Jira and Linear calls pull tokens from your existing secret store rather than a manually exported value.

yaml

```yaml
tracker:
  perTracker:
    tokenCommand: "your-secret-store-command --name jira-token"
```

Store Jira credentials in 1Password and let flowctl fetch them automatically — no token in any developer's shell environment

yaml

```yaml
tracker:
  perTracker:
    jira-cloud:
      tokenCommand: ["op", "read", "op://Engineering/Jira/token"]
```

Use macOS Keychain to supply the Linear API key without exporting it into the shell

```yaml
tracker:
  perTracker:
    linear:
      tokenCommand: ["security", "find-generic-password", "-s",
"linear-api-key", "-w"]
```

— Full mechanism, exact config key, env vars and file path given `flow-next-v9.12.0`

---

02 **Jira Product Discovery idea-to-parent workflow** Not checked yet

NEW 85

Attended flow can take a Jira Product Discovery idea and generate specs as child issues under one parent that links back to the idea, on Jira Cloud. Declaring discovery projects in `tracker.perTracker.ideaProjectKeys` lets flow-next treat any issue key in those projects as a JPD idea; `flowctl tracker idea read <idea>` reads the idea's text, comments, and delivery-linked issues, `flowctl tracker idea parent <idea> --body-file F` (with `--use` to pick among multiple candidates) creates or reuses a delivery-linked parent issue, and `flowctl tracker idea link <idea> <parent>` writes a missing delivery link back to the parent.

Start attended flow from an approved Jira Product Discovery idea so specs and child issues appear under a single delivery parent linked back to the idea.

```
$ flowctl tracker idea parent JPD-42 --body-file .flow/specs/my-
spec.md
# review candidates, then:
flowctl tracker idea link JPD-42 PROJ-100
```

Enable Jira Product Discovery idea routing by declaring which Jira Cloud projects hold discovery ideas.

```yaml
tracker:
  perTracker:
    ideaProjectKeys:
      - JPD
      - DISCOVERY
```

— Three named commands and config key, workflow end to end `flow-next-v9.13.0`

---

03 **Tracker wire commands for issue and spec bookkeeping**

Not checked yet NEW 83

`flowctl tracker wire list-open --any-lane` lists every open issue across all four trackers, including when `tracker.readyState` is unset, with `--unlinked` filtering to issues no spec references; output includes compact rows of `id`, `identifier`, `title`, `url`, `labels`, and a 500-character `excerpt`. `flowctl tracker wire unready` moves an issue out of the ready state on Linear, Jira, GitHub, or GitLab, `flowctl tracker wire question --reply-settles` treats a later non-question comment as the answer to an open question round, and `flowctl specs --refs --tracker-id <id>` lists every spec linked to a tracker issue across the checkout and all local and remote-tracking branches.

Find all open issues that no spec in the current checkout links yet, so you can decide which to build next without making any tracker calls per issue.

```
$ flowctl tracker wire list-open --unlinked
```

See every spec branch that references a specific tracker issue before merging or closing it, to avoid leaving orphaned specs.

```
$ flowctl specs --refs --tracker-id WOR-17
```

— Four exact subcommands/flags named with runnable examples `flow-next-v9.11.0`

---

04 **Automated issue-to-merged-PR flow** Not checked yet NEW

82

`flow --auto --from-issue <key|url>` takes a ready-state issue through spec capture, review, build, and merge in one step, and `--until=merge` runs this unattended end to end, posting clarifying questions on the issue if it is unclear and merging once CI and review allow. A new `from:issue` capture mode writes the spec directly from tracker issue text, quoting the issue in Goal & Context with no invented acceptance criteria, and capture now records the named issue as the spec's tracker link in default flow-id mode so a later capture push updates that issue instead of creating a duplicate.

Automate a full issue-to-merged-PR cycle from a single ready-state board issue without any manual steps.

```
$ flow --auto --from-issue
```

Build a Linear or Jira issue all the way to a merged PR with no manual steps — flow posts clarifying questions on the issue if anything is unclear and waits for the issue to re-enter the ready state before continuing.

```
$ /flow-next:flow --auto --from-issue WOR-17 --until=merge
```

— Exact flags and flow given, slightly thin on question-posting mechanics

flow-next-v9.11.0

05 **Parent-issue creation across trackers** Not checked yet NEW

64

Setting a parent on a spec's tracker block makes `flowctl tracker create`, `tracker create-first`, `tracker create --parent`, or `flowctl spec create --tracker-parent` create the tracker issue under that parent in a single request, on Jira Cloud, Linear, or GitLab.

Create a tracker issue (story or task) nested under an existing epic in a single request, without a second parenting step in the tracker.

```
$ flowctl tracker create --parent PROJ-12
```

— Exact flags and three tracker targets named flow-next-v9.13.0

THINNER COVERAGE BELOW

06 **Attended-flow triage for unlinked issues** Not checked yet NEW

36

Attended flow can triage open tracker issues that no spec covers, producing a table with one proposed action per unlinked issue (new spec, fold, dependency, or leave) and a single confirming question before any work begins.

— No command or flag named, described only in prose flow-next-v9.11.0

WAS THIS USEFUL? 👍 👎

Factory

Feedback ⓘ

CLI v1.13.0, Desktop v1.13.0: Add to chat for pull requests and files, smoother streamed

BLOG ↗ RANK

Factory's Droids are AI software-development agents with a CLI, web app and API.

> Across nine releases, Factory's Droids gained a new Add to chat action for pulling pull requests and files into context, Windows command sandboxing, richer automation scheduling and creation flows, and a steadily expanding public API reference, alongside a long tail of smaller session, connector, and UI refinements across CLI, desktop, and web.

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

01 **/side command for Side Chat** Not checked yet NEW 70

Typing `/side` in the desktop app opens Side Chat alongside an active session, matching the existing `/btw` command.

— Exact runnable command given with comparison to existing one

CLI v1.6.0, Desktop v1.6.0: Side chat w…

02 **External browser opening for links and HTML files** Not checked yet

NEW 65

You can now right-click an HTML file link in chat and choose Open in external browser to open it outside the app, and right-click any web link in the desktop app and choose Open in Browser to open it in your default browser.

— Exact right-click menu labels given for two link types

CLI v1.9.0, Desktop v1.9.0: Open in Bro… CLI v1.11.0, Desktop v1.11.0: Windows c…

03 **Automation scheduling and creation menu** Not checked yet NEW

65

Automation schedules now support a chosen weekday of the month, such as the second Monday of each month, and a new automation menu on the Automations list offers Create from template, Create with Droid, or Create manually as distinct creation paths.

— Concrete example given but no config detail Exact menu labels and navigation path named

CLI v1.10.0, Desktop v1.10.0: Disk spac… CLI v1.5.0, Desktop v1.5.0: New automat…

THINNER COVERAGE BELOW

04 **Windows command sandboxing** Not checked yet NEW 55

Windows command sandboxing now covers commands, hooks, and connector servers on Windows 11 hosts that support it.

– Scope of what's sandboxed is specific, but no enablement steps

CLI v1.11.0, Desktop v1.11.0: Windows c…

05 **Public API reference documentation and endpoints** Not checked yet

IMPROVED

50

The public API reference grew across releases: it now lets callers list the service accounts available for running automations, documents the health check and organization billing token limit endpoints, documents reading your account and listing the organizations you belong to, documents six additional operations and the errors each can return, and now gives every operation a stable identifier with consistent error responses.

– Endpoints and operations named in prose, no exact method/path given

CLI v1.9.0, Desktop v1.9.0: Open in Bro… CLI v1.11.0, Desktop v1.11.0: Windows c… CLI v1.13.0, Desktop v1.13.0: Add to ch…

06 **Add to chat for pull requests and files** Not checked yet NEW

50

A new Add to chat action lets you pull pull requests and files from a session's panels directly into your next chat message in the app.

– Mechanism explained but no exact UI path given CLI v1.13.0, Desktop v1.13.0: Add to ch…

07 **Namespaced tool names in Droid** Not checked yet IMPROVED

50

Tool rows in Droid now show the namespace and capability for command, connector, and MCP calls, rather than just a bare tool name.

– Clear before/after but purely cosmetic, no action needed

CLI v1.12.0, Desktop v1.12.0: Namespace…

08 **Risk-level adjustment for pushes and pull requests**

Not checked yet IMPROVED

50

Plain pushes and pull requests now count as medium risk instead of high risk; force pushes remain high risk.

– Exact before/after risk levels stated CLI v1.12.0, Desktop v1.12.0: Namespace…

09 **Disk space banner** Not checked yet NEW

50

A new disk space banner appears on low-disk computers, names the biggest space consumers, and offers to reclaim space.

– Explains trigger and behaviour but not the reclaim mechanism

CLI v1.10.0, Desktop v1.10.0: Disk spac…

10 **Editable email draft cards** Not checked yet NEW 50

Email drafts now render as editable cards that you can copy or open directly in your mail app.

— Describes the card's two actions clearly CLI v1.3.0, Desktop v1.3.0: Email draft…

11 **Connector branding for Splunk, Elasticsearch, and BigQuery**

Not checked yet IMPROVED 45

Splunk, Elasticsearch, and Google BigQuery now show their logo, description, and category in Connectors settings.

— Names three connectors but change is cosmetic only

CLI v1.5.0, Desktop v1.5.0: New automat…

12 **Loops on Droid Computers** Not checked yet IMPROVED 45

Loops now run on Droid Computers and keep their schedule across idle pauses.

— States behaviour but not setup steps CLI v1.4.0, Desktop v1.4.0: Loops on Dr…

13 **Translucent sidebar appearance setting** Not checked yet NEW

40

Appearance settings now offer a translucent sidebar option on macOS and Windows 11.

— Names the settings location but little else CLI v1.11.0, Desktop v1.11.0: Windows c…

14 **Session performance and navigation improvements** Not checked yet

IMPROVED 35

Droid starts replying sooner in sessions where hundreds of tools are available, switching to a model from another provider is faster in sessions with long histories, and moving between the sessions sidebar and Settings, Wiki, or Analytics now renders at once without a slide animation.

— Performance claims with no numbers or benchmarks

CLI v1.3.0, Desktop v1.3.0: Email draft…

15 **Permission-skip inheritance for Droid-started tasks**

Not checked yet IMPROVED 35

Tasks that Droid starts now inherit the run's permission-skipping setting instead of stalling for approval.

— Behaviour change stated but no config surface named

CLI v1.3.0, Desktop v1.3.0: Email draft…

16 **Audit entry for service account key revocation** Not checked yet

NEW

30

Revoking a service account key now records an audit entry for the organization.

— States what happens but no detail on where to view it

CLI v1.12.0, Desktop v1.12.0: Namespace…

17 **Session sidebar UI refinements** Not checked yet IMPROVED

30

The session sidebar now hides automation sessions by default until you turn on that filter, and loading placeholders in the session sidebar and the new computer repository list are now visible instead of blank.

— Thin cosmetic sidebar changes with little mechanism

CLI v1.4.0, Desktop v1.4.0: Loops on Dr… CLI v1.12.0, Desktop v1.12.0: Namespace…

18 **Script image/PDF read logging** Not checked yet NEW

30

A Script's image or PDF read now records one line naming its type and size.

— Bare description of a logging change CLI v1.11.0, Desktop v1.11.0: Windows c…

19 **Droid context on question rationale** Not checked yet NEW

25

Droid now gives context on why a question is being asked and what the choice affects.

— Vague description with no concrete mechanism CLI v1.13.0, Desktop v1.13.0: Add to ch…

WAS THIS USEFUL? 👍 👎

All Hands AI OpenHands

Feedback ⓘ v1.41.0 NOTES ↗

RANK

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

> OpenHands' latest release focuses on usability polish: clearer hook-file visibility, ACP-driven reasoning-effort selection, tag-based automation organization, and expanded test coverage.

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

01 **Workspace and user hooks file naming in Available Hooks dialog**

Not checked yet IMPROVED 48

The Available Hooks dialog now names the workspace and user hooks files, letting practitioners see exactly which files are in scope.

– Names UI dialog and file scope but no mechanism detail v1.41.0

02 **Tag-based grouping in automations dashboard** Not checked yet NEW

42

The automations dashboard list can now be grouped by tag for easier navigation across large automation sets.

– Clear UI location but no detail on tagging mechanism v1.41.0

03 **Reasoning-effort selection via ACP** Not checked yet NEW

37

Reasoning effort can now be selected based on what the agent reports via the ACP integration.

– Names ACP integration but no concrete steps or values v1.41.0

04 **Expanded verification test scenarios** Not checked yet IMPROVED

30

Verification scenarios now cover paths beyond the happy path, including `map controls` and `browser timezone`.

– Names specific test scenarios but internal, not user-actionable v1.41.0

WAS THIS USEFUL? 👍 👎

Amazon Kiro

Feedback ⓘ Web: Terminal in Kiro Web (2026-11-16)

BLOG ↗ RANK

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

> Kiro shipped an integrated terminal in Kiro Web, CLI hooks that can auto-approve or rewrite tool calls, real-time OpenTelemetry event streaming with private AWS PrivateLink/VPC Lattice connectivity for Enterprise observability, persistent cross-session Memory in the IDE, and conversation forking, alongside a string of cloud-session and MCP refinements in the CLI.

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

01 Cloud session controls in CLI

Not checked yet IMPROVED

85

In `--no-interactive --cloud` runs, `--cloud --resume` now resumes the newest cloud session, exiting with code 1 when none exists; local `mcp.json` handling reports how many enabled servers in your local config are unused because the cloud session uses its own cloud configuration. Cloud sessions in Europe (Frankfurt) can now create and attach to cloud sessions for Kiro profiles in that region. `/knowledge add` and `/knowledge update` now report that the indexed path is inside the cloud sandbox. Cloud sessions now hide commands, subcommands, and tips for features they cannot run — such as `/permissions` and `/chat save` — from the `/` menu, Tab completion, and `/help`, explaining why if you type one anyway. Cloud sessions without a repository now attach files named in a prompt, reading them from your current directory, and Workflows now list cloud sessions in their own group, separate from the current directory's sessions.

Continue your most recent cloud session unattended in a CI job, failing fast if no session is available.

```
$ kiro --cloud --resume --no-interactive
```

— Every named flag/command listed plus a runnable example.

CLI: Hooks That Rewrite Tool Inputs and…

02 Real-time OpenTelemetry streaming and private connectivity

Not checked yet NEW

85

Kiro now streams each LLM request and completed tool call as a real-time OpenTelemetry event to an OTLP collector, tracking credits, models, conversations, response times, tool outcomes, client versions, and agent modes; each export — real-

time events, daily CSV reports, and OpenTelemetry metrics — can be turned on and tested in the Kiro console, with exports including usage data and user identifiers but excluding prompts, responses, code, and files, and OpenTelemetry data and daily CSV reports can now be sent over AWS PrivateLink or Amazon VPC Lattice instead of the public internet, with private connections never falling back to public delivery; real-time events and private connectivity are now available for Kiro Enterprise profiles in US East (N. Virginia) and Europe (Frankfurt).

Enable real-time OpenTelemetry event streaming and choose a private delivery channel so LLM request telemetry never traverses the public internet.

> 📍 In the Kiro console, go to the Enterprise observability settings, turn on `Real-time events`, select `AWS PrivateLink` or `Amazon VPC Lattice` as the delivery method, enter your OTLP collector endpoint, then click `Test export` to verify connectivity.

— Tracked fields and console steps given, exact endpoint not named. Named channels and no-fallback guarantee, no setup walkthrough attached.

`General: Real-time events and private c…`

---

03 **Hooks for approving and rewriting tool calls** `Not checked yet`

`NEW` 80

V3 CLI sessions gain two new hook types: a `PermissionRequest` hook that returns `allow` or `deny` to automate tool-call approvals without a human prompt, and a `PreToolUse` hook that can rewrite a tool call's inputs before it runs by returning `updatedInput` alongside an `allow`, `ask`, or `deny` decision.

— Named return values given but no runnable example.

`CLI: Hooks That Rewrite Tool Inputs and…`

---

04 **Four memory levels for CLI V3 sessions** `Not checked yet` `IMPROVED`

80

V3 session `config` now offers four memory levels — `Use and create automatically`, `Use and create when asked`, `Use but don't create`, and `Off` — replacing the earlier `read & write`, `read only`, and `off` levels and their separate automatic-updates checkbox, and shows the memory default set by a `KIRO_DEFAULT_SETTINGS` file instead of the built-in default.

— Exact level names and config file named, no example run.

`CLI: Hooks That Rewrite Tool Inputs and…`

---

05 **Persistent Memory across chat sessions** `Not checked yet` `NEW`

80

Kiro adds persistent Memory across chat sessions: open the Memory view via the Memories icon in the activity bar to review what Kiro remembers, and control behavior

via the `Kiro Agent: Memory` setting, including a Use only mode that applies memories without changing them. Memory levels now have clearer names describing when Kiro creates memories, and the Memory view and open memory tabs update automatically as memories change.

Prevent Kiro from writing new memories mid-session while still benefiting from previously saved context — useful when working on a one-off branch where you don't want conventions overwritten.

> 📍 In the IDE, open Settings and set Kiro Agent: Memory to Use only .

Audit what Kiro has stored about your project before starting a sensitive session — review or delete stale entries before the agent picks them up.

> 📍 In the activity bar, select the Memories icon to open the Memory view, then review and edit the stored entries.

— Exact setting name and UI path with two worked examples.

IDE 1.4.215: Memory Across Chat Session… | IDE: Fork Conversations, Workflow Credi…

---

06 **MCP OAuth scope fallback and connection timeout** Not checked yet

IMPROVED

70

V3 MCP servers now honor nested OAuth scopes in `oauth.oauthScopes` or `oauth.scopes` when top-level `oauthScopes` is not configured, and respect the `timeoutMs` connection timeout setting.

— Config keys named exactly, no usage example. CLI: Hooks That Rewrite Tool Inputs and…

THINNER COVERAGE BELOW

07 **Integrated terminal in Kiro Web** Not checked yet NEW

55

Kiro Web gains an integrated terminal that can be docked beside or below the conversation, letting you run builds, tests, and ad-hoc commands directly in the session workspace without leaving the browser.

Run a build and check its output without leaving Kiro Web — open the terminal docked beside the conversation so you can see agent responses and command output side by side.

> 📍 In Kiro Web, click the terminal icon and select Beside conversation to dock the terminal on the right, then run your build or test command.

— UI navigation given; no deeper mechanism described.

Web: Terminal in Kiro Web (2026-11-16)

---

08 **Daily OpenTelemetry metrics delivery time shifts** Not checked yet

55

BREAKING

Daily OpenTelemetry metrics delivery time moves to 00:30 UTC on November 26, 2026; until then, Dynatrace and Grafana Cloud will reject them as too old. The CSV report remains at 02:00 UTC.

— Exact times and affected vendors given, no mitigation step.

General: Real-time events and private c…

09 **Fork conversation at any turn** Not checked yet NEW 55

Select Fork conversation in a turn's footer to start a new chat that carries the conversation through that turn, leaving the original chat and files unchanged.

When you want to explore an alternative prompt direction mid-conversation without losing your original thread — fork from any past turn and experiment freely.

📍 In a chat, hover over any past turn's footer and select Fork conversation to open a new chat branch from that point.

— Exact UI label and navigation given. IDE: Fork Conversations, Workflow Credi…

10 **Workflow credit reporting** Not checked yet NEW 45

Workflow credit reporting now shows the credits an agent reports for finished steps on a workflow run's card, with step updates grouped in a Workflow report in the parent chat.

— Describes what appears but not where to enable it.

IDE: Fork Conversations, Workflow Credi…

11 **Concurrent chat turns for custom agents and Plan** Not checked yet

IMPROVED 35

Chats using a custom agent or Plan now run their turn at the same time as other chats, instead of waiting with The agent is currently busy until another chat's turn ends.

— Behavior change stated with no configuration or steps.

IDE: Fork Conversations, Workflow Credi…

└──▷ BREAKING ON UPGRADE

! Daily OpenTelemetry metrics delivery time moves to 00:30 UTC on November 26, 2026; until then, Dynatrace and Grafana Cloud will reject them as too old. The CSV report remains at 02:00 UTC.

WAS THIS USEFUL? 👍 👎

swttch

Feedback ⓘ v0.38.0 NOTES ↗ RANK

Swttch is a claude Code GUI plugin for JetBrains IDEs. (ex - Claude Code with GUI)

> Swttch's v0.38.0 release lets users retract a sent message before the CLI has read it.

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

01 **Retract unread message from waiting bubble** Not checked yet NEW

40

Users can now retract a sent message before the CLI has read it, directly from its waiting bubble.

— Describes the action and UI location but no mechanism detail. v0.38.0

WAS THIS USEFUL? 👍 👎

ogcode

Feedback ⓘ v0.59.3 NOTES ↗ RANK

ogcode executes code tasks via a local-first AI agent that manages tools, memory, permissions, and human oversight to turn intent into completed work.

> ogcode's v0.59.3 release adds an analyze_image tool so coding agents can read screenshots and diagrams without vision-capable chat models, alongside configurable provider request concurrency, a wave of new model catalog entries, a larger deep-search ranker budget, and a floating subagent activity dock.

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

01 **analyze_image tool for screenshots and diagrams** Not checked yet

NEW 95

A new `analyze_image` tool lets any coding agent read screenshots, scanned pages, and charts via a dedicated vision model, returning text so the requesting model never needs image support itself. Its `transcribe` mode returns visible characters in reading order, `describe` gives a factual account of diagram components and on-image text, and `answer` (the default) answers a pointed `question` from the image content; `effort` (`low`, `medium`, or `high`) is resolved against the vision model directly and silently dropped if unsupported. The vision model is configurable machine-wide via `ui_prefs.image_model` and `ui_prefs.image_provider` (migration 078) or via `GET` and `POST /models/image`, defaulting to the session model when unset if it supports vision.

Analyze a crash dialog screenshot from a CI artifact — the agent reads the error, stack trace, and next-step recommendation without needing a vision-capable chat model.

```
$ analyze_image(path='artifacts/crash-dialog.png', mode='answer',
question='What error is shown and what is the recommended fix?',
effort='high')
```

Pin a vision model for all image analysis on the machine so every session uses it, independent of whichever chat model is active.

```
$ curl -X POST http://localhost:<port>/models/image \
  -H 'Content-Type: application/json' \
  -d '{"model": "gemini-4-carbon", "provider": "google"}'
```

— Names modes, config keys, migration, and endpoint with examples. v0.59.3

02 **Configurable provider request concurrency ceiling** Not checked yet

NEW

95

Adds a configurable ceiling (floor: 2, default: 8) on process-wide provider request concurrency, set via `ui_prefs.provider_max_concurrent` (migration 079) or `GET` and `POST /providers/concurrency`, surfaced in Settings → Provider requests and taking effect immediately without preempting in-flight streaming requests. The `OGCODE_PROVIDER_MAX_CONCURRENT` environment variable overrides the stored value and makes the UI field read-only.

Cap provider concurrency at 4 to avoid rate-limit errors when running a heavy index alongside interactive sessions, overriding the default of 8.

```
$ export OGCODE_PROVIDER_MAX_CONCURRENT=4
```

— Names config key, endpoint, env var, floor/default values, and UI path. v0.59.3

THINNER COVERAGE BELOW

03 **Floating subagent activity dock** Not checked yet NEW

55

Subagent activity now shows in a floating dock above the composer with per-pill title, spinner, hover details for workload and elapsed time, `aria-live` announcements, and a `+N more` overflow for wide fan-outs.

— Describes UI elements in detail but no config key or command. v0.59.3

04 **New models added to catalog** Not checked yet NEW

50

Adds Claude Fable 5.2 (default, $8/$40 per million tokens, 1M window), Haiku 5.5, and Mythos 5.2 to the model catalog, alongside OpenAI GPT-6.1 Sol and GPT-6.1 Terra (Responses API only, 922K input cap), Google Gemini 4 Hydrogen and Gemini 4 Carbon, and Kimi K3.1 Pro and K3.1 Pro Highspeed.

— Lists many named models with pricing but no usage mechanism. v0.59.3

05 **Deep-search ranker output budget increased** Not checked yet

IMPROVED

40

Deep-search ranker output budget raised from 1,024 to 4,096 tokens so a reasoning model can emit rankings without exhausting its thinking budget first.

— Gives exact before/after numbers but no config surface to act on. v0.59.3

WAS THIS USEFUL? 👍 👎

agentty

Feedback ⓘ v0.9.57 NOTES ↗ RANK

Agentty is an AI pair programming in your terminal — one static binary, sub-ms startup, any model

> Agentty v0.9.57 adds support for Claude Code's hook protocol to intercept and control tool calls, extends sandboxing to 32-bit programs, and fills in missing keyboard editing keys across input fields.

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

01 Claude Code hook protocol support

Not checked yet NEW

88

Hooks defined in `hooks.json` can use `PreToolUse`, `PostToolUse`, and `PostToolUseFailure` events, delivered as JSON on stdin, to deny, allow, rewrite, or replace tool calls and their outputs. Hooks now also run in `agentty run` and inside subagents, not just the interactive session.

Intercept a tool call before it runs to inspect or deny it — useful in CI where you want to block destructive shell commands automatically.

json

```
# hooks.json
{
  "hooks": [
    {
      "event": "PreToolUse",
      "tool": "shell",
      "command": ["./scripts/check-shell-safety.sh"]
    }
  ]
}
```

— Names exact events, file, and command with a runnable config example. v0.9.57

THINNER COVERAGE BELOW

02 32-bit sandbox execution

Not checked yet NEW

55

The sandbox can now run 32-bit programs on x86_64 Linux with multilib installed; these processes get their own syscall filter matching the allowed set of the 64-bit filter.

— Explains mechanism and platform scope but no command to try it. v0.9.57

03 Editing keys in input fields

Not checked yet IMPROVED

32

Delete, Home, and End keys now work in the prompt, the login box, and the settings add field.

— Names exact keys and fields but is a simple fix with no further mechanism. v0.9.57

WAS THIS USEFUL?  

coffee-cli

Feedback ⓘ v4.0.6 NOTES ↗ RANK

coffee-cli is a clean AI terminal workspace with quick input, folder switching, copy & paste, and image pasting.

> coffee-cli added remote device control for operating another computer's sessions and files, plus shareable prompt links and inline next-prompt suggestions for Claude conversations.

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

01 **Remote device control window** Not checked yet NEW 85

Click the device status at the bottom left and select Control another device… to operate another computer's sessions and files in a dedicated window without re-pairing; the remote computer must be running v4.0.2 or later. A new device menu at the bottom left lists computers currently being controlled, recently controlled computers, and devices connected to this machine.

— Names exact UI path and version requirement v4.0.6

02 **Shareable prompt links** Not checked yet NEW 70

A new Save to Pre-input · Share · Clear button group above the input box lets you share prompts via a copyable link, letting recipients read, copy, or open the prompt directly in Coffee CLI.

— Names exact button group and behavior v4.0.5

THINNER COVERAGE BELOW

03 **Inline next-prompt suggestions for Claude** Not checked yet NEW 35

Claude conversations now show inline grey-text suggestions for what you might say next in the input box.

— Describes feature but no mechanism detail v4.0.5

WAS THIS USEFUL? 👍 👎

Cotool

Feedback ⓘ v0.80.0 NOTES ↗ RANK

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

> Cotool's v0.80.0 release adds three new AI models, reworks the Log Ingest interface with trend visibility and dedicated views, and improves agent performance for large tool catalogs.

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

01 **Reworked Log Ingest interface** Not checked yet IMPROVED

60

Log Ingest now prioritizes streams that need attention, surfaces 24-hour lag trends, provides clearer remediation guidance, and adds dedicated Streams, Runs, and Tables views.

— Describes behavior and named views but no exact navigation v0.80.0

THINNER COVERAGE BELOW

02 **New model support across plans** Not checked yet NEW 40

Cotool adds GPT-6 Astra and Gemini 4 Carbon as available models for token-based plans, and `Claude Fable 5.2` as an available model for organizations using their own Anthropic key.

— Names models and plan types but no setup steps v0.80.0

03 **Lazy-loaded integration actions for agents** Not checked yet

IMPROVED 25

Agents with large tool catalogs now load integration actions only when needed, reducing overhead.

— States benefit but no mechanism or scope detail v0.80.0

WAS THIS USEFUL? 👍 👎

juggler

Feedback ⓘ v0.8.1 NOTES ↗ RANK

The Juggler Code Agent

> Juggler v0.8.1 adds explicit reverse-proxy serving flags, multi-account provider support, and a worktree panel that can land commits by rebase or merge, alongside smaller improvements to per-window behavior, project file browsing, AGENTS.md discovery, and PDF attachments.

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

01 Port-strict and base-path serving flags

Not checked yet NEW

90

Adds `--port-strict` to make the server fail immediately on a busy port instead of silently moving to the next free one, and `--base-path` to serve Juggler under a reverse proxy path when the proxy does not send an X-Forwarded-Prefix header.

Prevent a Juggler server from silently occupying a different port when the configured one is busy — useful behind a reverse proxy that expects a fixed port.

```
$ juggler serve --port-strict
```

Serve Juggler under a subpath via a reverse proxy that strips the prefix but does not forward an X-Forwarded-Prefix header.

```
$ juggler serve --base-path /juggler --port-strict
```

— Named flags with runnable example commands. v0.8.1

THINNER COVERAGE BELOW

02 Worktree panel rebase/merge landing

Not checked yet IMPROVED

55

The worktree panel now shows unlanded commits and can land work by rebase or merge, including submodules.

— Describes behavior and options but no exact command or key. v0.8.1

03 AGENTS.md discovery in subfolders

Not checked yet NEW

55

Juggler now reads `AGENTS.md` from subfolders as the agent works in them, alongside `.local` and global ones.

— Named config files give a clear, if brief, mechanism. v0.8.1

04 **File browser for project files and pins** Not checked yet NEW

50

Project files and pinned folders now act as a file browser with click-to-preview, sizes, git status, and filtering.

— Names concrete UI capabilities but no exact navigation path. v0.8.1

05 **PDF attachment as extracted text** Not checked yet NEW

50

PDFs dropped into the composer or @`-mentioned` are now attached as their extracted text.

— Describes trigger and effect clearly, no further mechanism. v0.8.1

06 **Project folder rename and move** Not checked yet NEW 40

The agent can now rename or move its project folder, with windows following it automatically.

— Explains behavior; no mechanism or UI path given. v0.8.1

07 **Per-window theme/zoom and background task persistence**

Not checked yet IMPROVED 40

Theme and zoom are now set per window, so changing one window no longer restyles others, and background tasks now keep running after their project's window is closed.

— Clear before/after but no named setting or path. States behavior change only, no scope or limits.

v0.8.1

08 **Multiple accounts per provider** Not checked yet NEW 35

Add multiple accounts per provider and select which one to use per conversation.

— States capability but no config surface or UI path given. v0.8.1

09 **Mid-conversation date change notice** Not checked yet IMPROVED

25

Long-running conversations now tell the model when the date changes mid-conversation.

— Bare statement with no mechanism or trigger details. v0.8.1

WAS THIS USEFUL? 👍 👎

◆ AI App Builders

StackBlitz bolt.new

Feedback ⓘ updated · Nov 18, 2026

NOTES ↗ RANK

Bolt.new is an AI-powered web development platform that builds websites and applications from prompts.

> bolt.new tightened Free plan limits around collaborators and project transfers, and reorganized account settings into dedicated billing, member-management, and team pages.

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

01 Settings reorganized into Billing, Members, and Team pages

Not checked yet

78

BREAKING

The former Subscription & Tokens page is renamed Billing & usage and now includes billing details, payment method, billing history, plan changes, and subscription cancellation or resumption. Team admins now get three separate pages: Billing & usage (invoice, token refill date, and per-member token consumption table), Manage members (invitations, roles, upgrades), and Team settings (moved to the left sidebar, no longer behind the gear icon on the Team page). Non-admin team members now see a Usage page showing their own token balance and a Team page listing team members.

Navigate to your team's billing and member usage as an admin after the settings restructure.

> 📍 Open Settings › Billing & usage to see your team's next invoice, next token refill date, and each member's token consumption. Then open Settings › Manage members to invite new members or change roles.

— Names every renamed/relocated page and gives exact navigation steps. snapshot-20261118

THINNER COVERAGE BELOW

02 Collaborator cap on Free plan projects

Not checked yet NEW

42

Free plan projects now cap collaborators at `3`; existing invitees retain their access, while paid plans remain unlimited.

— Clear limit and scope, but no config or command to act on. snapshot-20261118

03 Free-to-Free project transfer restriction

Not checked yet

42

BREAKING

Project transfers now require at least one party (sender or recipient) to have a paid plan; Free-to-Free transfers are no longer allowed.

— Rule is explicit but there's no workaround or setting to configure. snapshot-20261118

└──▷ **BREAKING ON UPGRADE**

! Project transfers between two Free plan accounts are no longer allowed; at least one party must have a paid plan.

! The `Subscription & Tokens` settings page is renamed `Billing & usage`; any documentation or bookmarks pointing to the old name will need updating.

! The Team settings page is no longer opened via the gear icon on the Team page; it is now a standalone entry in the left sidebar.

WAS THIS USEFUL?  

Lovable

Feedback ⓘ Nov 15, 2026 (2026-11-16) BLOG ↗ RANK

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

> Lovable's biggest window yet: ten-plus new third-party connectors (Intercom, Attio, Deel, Retool, ActiveCampaign, Supermetrics, Mailchimp, Mailchimp Transactional, BambooHR, Mollie) landed alongside new AI models including Topaz video restoration, isolated draft-backend databases, a security-findings gate on publish, personal skills/instructions, background memory maintenance via Dreaming, and a string of enterprise admin, SSO, and audit controls.

└──▷ **WHAT SHIPPED · 32 FEATURES** most completely described first ? what's the number?

01 New AI models in Connectors → AI: OpenAI and Topaz

Not checked yet NEW 80

Three new OpenAI models: `openai/gpt-6.1-terra` for fast reasoning and image analysis at lower cost, `openai/gpt-image-2.6-flare` for fast image previews and in-app editors, and `openai/gpt-image-2.6-sunburst` for detailed visuals and precise edits — none available when EU inference is enabled. Both image models now support animated GIFs, guided edits via drawing on the image, and transparent-background stills as PNG or WebP. New `topazlabs/starlight-fast-2` model restores blurry, noisy, or low-resolution video clips at 1080p or 4K at the original frame rate, enabled by workspace admins or owners under Connectors → AI.

Use the low-cost reasoning model for a text-and-image analysis feature when you want GPT-6.1-class quality without the higher credit cost.

> 📍 In your Lovable project chat, type: 'Use openai/gpt-6.1-terra for the image analysis step in my product review feature.'

Enable the Topaz Starlight Fast 2 video restoration model for apps in your workspace so developers can restore low-quality clips to 1080p or 4K.

> 📍 In the workspace, go to Connectors → AI, find TopazLabs, and enable it. Then ask Lovable in your project: 'Use topazlabs/starlight-fast-2 to restore this clip.'

— Names exact model identifiers, formats, and enablement path Nov 6, 2026 (2026-11-08) Nov 5, 2026 (2026-11-07)

02 Security findings gate on publish with one-click fix

Not checked yet NEW 80

When a project uses the built-in backend and has unresolved critical security findings, Lovable now asks for review before publishing, with options to open the Security view,

publish anyway, or ask Lovable to `Fix all` in one request; the same `Fix all` prompt appears above the chat input whenever open critical findings or warnings exist, and sending them counts as one free fix.

Review and clear critical security findings before publishing so you do not accidentally expose a vulnerable backend app.

> 📍 In your project, select Publish. If critical security findings exist, Lovable surfaces the review prompt — choose `Open Security view` to inspect each finding, or select `Fix all` to send them to Lovable in one request, then publish once they are resolved.

Quickly remediate all open security warnings from the chat without leaving the editor.

> 📍 In the project chat, look above the input for the notice showing open critical findings or warnings. Select `Fix all` to send them to Lovable in one request (counts as one free fix).

— Names exact dialog options and the free-fix mechanic `Nov 12, 2026 (2026-11-13)`

---

03 **New third-party connectors: Intercom, Attio, Deel, Retool, ActiveCampaign, Supermetrics, Mailchimp, BambooHR, and Mollie**

`Not checked yet` **NEW** 75

Intercom: search and reply to support conversations, manage contacts and tickets, and search and publish help center articles. Attio: search, create, and update people, companies, and deals, add records to lists, and attach notes and tasks, with per-user access so each published-app user connects their own Attio workspace. Deel: read and update contracts, people, timesheets, time off, and payments. Retool: manage members, groups, and permissions of a Retool organization and embed existing Retool apps. ActiveCampaign: search and update contacts, manage mailing lists, schedule campaigns, build audience segments. Supermetrics (prebuilt chat connector, requires a subscription with API access): query campaign metrics, analytics, and store data, and read Supermetrics Studio dashboards. Mailchimp: add and update audience contacts, apply tags, read segments, and display campaign results (opens, clicks, unsubscribes, bounces). Mailchimp Transactional: send email from a verified sending domain using existing templates, schedule messages with attachments, view delivery activity. BambooHR: look up and update employee records, read the company directory, see who is out of office. Mollie: create payments, send customers to Mollie's hosted checkout, confirm payment status. All addable from Connectors or by asking Lovable in a project chat.

Add an Intercom integration to a Lovable project so users can search and reply to support conversations from inside the app.

📍 In your Lovable project, go to Connectors and select Intercom, then follow the prompts to authenticate and connect.

Add the Attio connector to a project so your app can create and update CRM records directly from a lead intake form.

📍 In your project, go to Connectors, find Attio, and click Add — or type `connect to Attio` in the Lovable project chat to have Lovable wire it up for you.

Connect a Deel or Retool data source to a new project so the app can read HR data or embed existing Retool tools.

📍 In your Lovable project, open Connectors and select Deel or Retool, or type a request such as `connect Deel` in the project chat to have Lovable configure the connector automatically.

Connect Mailchimp Transactional to a Lovable project to send order confirmations or receipts from your verified domain.

📍 In the Lovable project, go to Connectors, select Mailchimp Transactional, and follow the setup prompts — or ask Lovable: 'Connect Mailchimp Transactional so I can send order confirmation emails.'

When building an employee directory or out-of-office board, connect your app to BambooHR directly from the Lovable UI.

📍 In your Lovable project, go to Connectors, find BambooHR, and select Connect — or type a request in the project chat such as 'Connect this app to BambooHR to show who is out of office today'.

Add Mollie as a payment connector so your app can accept iDEAL, Bancontact, card, and SEPA payments through your own Mollie account.

📍 1. In your Lovable project, open `Connectors`. 2. Find Mollie and select it to connect. (Alternatively, in the project chat, ask Lovable: `Connect my app to Mollie`.) 3. Lovable will configure the integration to send customers to Mollie's hosted checkout and confirm payment status.

— Names ten connectors and their exact capabilities, thin per-connector

`Nov 16, 2026 (2026-11-16)` `Nov 15, 2026 (2026-11-16)` `Nov 12, 2026 (2026-11-13)` `Nov 6, 2026 (2026-11-08)` `Nov 5, 2026 (2026-11-07)` `Nov 3, 2026 (2026-11-04)` `Nov 2, 2026 (2026-11-04)`

---

04 **Isolated database for draft backend testing** `Not checked yet` NEW

75

Backend changes can now be tested in a draft environment with its own isolated database filled with example data, so schema changes such as new form fields can be verified before they reach the live database; selecting `Accept Draft` applies the schema change to the live database without carrying over the test data.

Safely test a new form field that writes to the database in a draft before the change goes live.

> 📍 Open a draft on a project with the built-in backend, add your new form field, submit test data to the draft's isolated database, then click Accept Draft to apply the schema change to your live database without the test data.

— Explains mechanism and exact before/after of applying a draft `Nov 16, 2026 (2026-11-16)`

---

05 **Custom SPF record entries for purchased domains** `Not checked yet`

NEW

75

Workspace admins and owners can add custom entries to a domain's SPF record to allow third-party inbox providers such as Google Workspace or Zoho to send email from a Lovable-purchased domain, via Domain Settings → SPF Record.

Allow Zoho to send transactional email from your Lovable-purchased domain by adding its SPF include to your domain's record.

> 📍 In the Lovable console, go to Domain Settings › SPF Record, enter the SPF include string provided by Zoho (e.g. include:zoho.com), and save.

— Exact settings path and example providers given `Nov 13, 2026 (2026-11-16)`

---

06 **Abandonment protection for Enterprise projects** `Not checked yet`

NEW

75

Enabling `Abandonment protection` in a project's settings, or selecting `Protect it` on the abandoned-project banner, prevents a project from ever being marked inactive or deleted; available on Enterprise plans and manageable by workspace admins and owners via `Security insights`.

Protect a sensitive Enterprise project from inactivity cleanup via the project settings UI.

> 📍 1. Open your project in Lovable. 2. Go to your project's Settings. 3. Find and enable `Abandonment protection`. (Alternatively, if the abandoned-project banner is showing, select `Protect it`.) 4. Workspace admins can also protect any project across the workspace via Security insights.

— Names exact setting, banner action, and admin surface `Nov 3, 2026 (2026-11-04)`

---

07 **Connector management and discovery improvements** Not checked yet

IMPROVED 70

Workspace admins and owners on Business and Enterprise plans can set the availability of several connectors in one step in the Connectors admin settings; the Connectors catalog now groups connectors by goal category (such as Sell and get paid or Add AI to your apps) and supports natural-language search; connector settings can be changed from More options → Manage connector on redesigned connector pages; a new Activity tab under Connectors shows a connector's request log and lets you send a failed request to the project chat with Fix with Lovable.

Enable or disable multiple connectors in one pass to keep only the integrations your team needs.

> 📍 In the workspace admin console, go to Connectors admin settings, select the checkboxes for each connector you want to enable or disable, then apply the change in one step.

Debug a failing ActiveCampaign connector call without leaving Lovable — view the request log and hand the error straight to the AI.

> 📍 In your project, go to Connectors › Activity tab, find the failed ActiveCampaign request, then select Fix with Lovable to send the error details to the project chat.

— Names exact UI paths and labels for each admin surface Nov 13, 2026 (2026-11-16)

Nov 10, 2026 (2026-11-12) Nov 6, 2026 (2026-11-08)

---

08 **Dreaming: automatic background memory maintenance** Not checked yet

NEW 70

On paid plans, Dreaming automatically keeps project memory current by checking and rewriting `AGENTS.md` rules against the code and chat history in the background without consuming credits; changes appear in version history as Lovable tidied up its notes.

— Explains mechanism, cost, and where to see changes Nov 4, 2026 (2026-11-04)

---

09 **New SSO options: session duration and Google sign-on**

Not checked yet NEW 65

Adds new SSO capabilities for Business and Enterprise plans: workspace admins and owners can set SSO session duration to 30 days or Indefinite for workspace members in Workspace Settings → SSO (defaulting to Indefinite when unset), and workspace members enforcing SSO through Google can log in once per session by selecting Continue with Google with a Google Workspace account on the workspace's verified domain.

Keep Business or Enterprise workspace members permanently logged in through SSO without repeated re-authentication prompts.

📍 In Workspace Settings, go to SSO, then set Session Duration to Indefinite.

— Exact setting, values, and navigation path given Clear action but no mechanism beyond the login flow

Nov 16, 2026 (2026-11-16) Nov 4, 2026 (2026-11-04)

10 **Two-factor authentication scoped to external collaborators**

Not checked yet NEW 65

Enterprise workspace admins and owners can require two-factor authentication from external collaborators only by choosing External collaborators, even when Enforce SSO is enabled.

Require two-factor authentication from external collaborators only — without blocking SSO-enforced internal members — in an Enterprise workspace.

📍 In the workspace admin console, go to Security settings, find the two-factor authentication section, and select External collaborators as the enforcement scope. This applies even when Enforce SSO is already enabled.

— Exact setting name and interaction with SSO enforcement Nov 10, 2026 (2026-11-12)

11 **Image resizing and video generation in Files and Chats**

Not checked yet NEW 65

From a project's Files tab image editing toolbar, Resize extends a scene to fill a new aspect ratio (such as square or widescreen) and Make video animates the image with Add prompt or Quick animate. The same toolbar is now available directly in Chats, with options to describe edits, erase, resize, or convert an image to a short video; each edit is saved as a new version.

Turn a static image into an animated short video directly inside a project's Files tab.

📍 In your project, open the Files tab, select an image, then choose Make video from the image editing toolbar. Select Add prompt to describe the animation in the project chat, or choose Quick animate to let Lovable decide. Lovable saves the result as a new file next to the original.

Convert a static hero image into a short video directly from a Chat conversation.

📍 Open the image in Chats, then use the image editing toolbar to select the video conversion option and describe the motion you want.

— Names exact toolbar options across both surfaces Nov 11, 2026 (2026-11-12)

Nov 10, 2026 (2026-11-12)

12 **Scheduled recurring automations in Chats** Not checked yet NEW

65

Recurring tasks — such as morning site checks, weekly signup summaries, or backlog reviews — can be automated by asking Lovable in Chats with a frequency, and results are posted there; available on Free, Pro, and Business plans.

Set up a weekly recurring task so Lovable summarises new signups every Monday morning without manual prompting.

> 📍 In Chats, type: 'Every Monday at 9am, give me a summary of new signups from this week.'

— Gives a runnable natural-language example and plan availability

Nov 11, 2026 (2026-11-12)

---

13 **Access-request flow for internal apps** Not checked yet NEW

65

Internally published apps now support visitors outside the audience requesting access with an optional message; the owner approves or declines from the Inbox, and Lovable emails the requester and adds approved users to the audience without republishing.

Gate an internal tool so external visitors can self-serve request access instead of emailing you directly.

> 📍 Publish your app internally, then monitor your Inbox for access requests — approve each one to add the person to the audience; they receive an email confirmation automatically.

— Explains the full request/approve/notify mechanism Nov 6, 2026 (2026-11-08)

---

14 **Subfolder-scoped AGENTS.md and CLAUDE.md instructions**

Not checked yet IMPROVED

65

Lovable now reads `AGENTS.md` and `CLAUDE.md` files in project subfolders: when it opens or edits a file, it reads instruction files in that file's folder and every folder above it.

— Names exact files and the folder-traversal mechanism Nov 3, 2026 (2026-11-04)

---

15 **Preview comments threading, repositioning, and guest access**

Not checked yet IMPROVED

60

Preview comments now number each thread (e.g. `#3`) for precise reference; a thread can be reattached to a different element by dragging its pin onto that element, or moved aside by dragging it to see what's underneath; guests on a shared preview link can now

read and reply in the comments panel, search threads, and see all comments in a list on a phone.

— Names concrete UI behaviours but no deeper mechanism Nov 15, 2026 (2026-11-16)

16 **Per-workspace personal memory isolation** Not checked yet IMPROVED

60

Lovable now keeps personal memory (what it learns about how a user likes to work) separate per workspace so preferences from one workspace do not carry over to others; personal memory has moved to account settings, where a workspace can be selected to view or disable its memory.

— Explains before/after and exact new settings location Nov 10, 2026 (2026-11-12)

17 **Admin approval required to enable backend services**

Not checked yet NEW

60

Enterprise plans can enable a workspace policy requiring an admin or owner to approve before backend services — database, user accounts, and file storage — can be turned on in any project; members submit a request with an optional message, and approvals or denials are handled in the Action center.

— Names the exact services gated and approval location Nov 7, 2026 (2026-11-08)

18 **Publish and Action center approval workflow improvements**

Not checked yet IMPROVED

60

Action center review dialogs now show each project's owner, list open security findings with severity, and let you ignore all open findings at once. The Publish dialog now lists blockers under Before you publish and, when workspace approval is required, shows the request status such as Awaiting review or Approved by with the approver's name.

— Names exact dialog labels and statuses shown Nov 2, 2026 (2026-11-04)

THINNER COVERAGE BELOW

19 **Searchable and filterable workspace API keys** Not checked yet

NEW

55

On Business and Enterprise plans, workspace API keys can now be searched, filtered, and sorted, including finding keys by IP restriction, expiry status, or creator.

— Names the filter fields but no exact UI steps Nov 13, 2026 (2026-11-16)

20 **Enterprise workspace claiming for verified domains**

Not checked yet NEW 55

Enterprise org admins can now claim workspaces whose primary owner uses a verified domain, with a 14-day notice period and the option for the owner to approve sooner.

— Clear mechanism and timeline, no UI steps given Nov 13, 2026 (2026-11-16)

21 **Communication personalization toggle** Not checked yet NEW

55

The Use my activity to personalize communication setting in account preferences controls whether Lovable personalizes its messages to a user; it is enabled by default outside the EU, EEA, UK, and Brazil.

— Names the exact setting and default-region behaviour Nov 2, 2026 (2026-11-04)

22 **Redesigned project and workspace domain settings** Not checked yet

IMPROVED 55

Project and workspace domain settings have a redesigned layout: buy a domain under Find and buy a domain, see connected domains with status (Active or Needs attention), and open any domain to view its alerts.

— Names exact labels in the redesigned UI Nov 2, 2026 (2026-11-04)

23 **Personal skills and personal instructions** Not checked yet NEW

50

Adds personal skills — private per-workspace playbooks only the user can use, with a separate set per workspace, configured in account settings or by asking Lovable in a project chat — and personal instructions, rules and preferences Lovable follows in every project the user opens in a workspace, invisible to teammates, configured separately for each workspace in account settings.

— Names config location, lacks concrete example Nov 10, 2026 (2026-11-12)

24 **Manual acceptance required for invitations** Not checked yet

50

BREAKING

Workspace and project invitations now require manual acceptance from email or the Inbox rather than being accepted automatically; Auto-accept invitations has been removed from account settings.

— States the removed setting and new required action Nov 11, 2026 (2026-11-12)

25 **Problem reporting from MCP clients** Not checked yet NEW

50

Problems can now be reported to Lovable directly from an MCP client by asking the client to send a report for issues such as misread requests, stuck projects, or failed builds.

— Names the surface (MCP client) but no example report shown Nov 5, 2026 (2026-11-07)

26 **Browser-side error monitoring on paid plans** Not checked yet NEW

50

Project monitoring on paid plans now also captures errors occurring in visitors' browsers, identifies the cause in your code, surfaces findings in the editor, and emails the project owner about important findings; active only while Visitor analytics is enabled.

— Explains the pipeline but no UI path given Nov 2, 2026 (2026-11-04)

27 **External invitee labelling in audience list** Not checked yet NEW

50

When sharing a published app, each external invitee's email address now appears beneath their name in the audience list, with an `EXT` label marking external invitees and collaborators.

— Names the exact label and where it appears Nov 4, 2026 (2026-11-04)

28 **Daily run usage spike alerts** Not checked yet NEW 45

Workspace admins and owners on paid plans now receive an email and in-app notification when daily run usage rises sharply compared with the previous 7 days.

— States the trigger but no threshold or config Nov 12, 2026 (2026-11-13)

29 **Published commit recorded in audit logs** Not checked yet IMPROVED

45

On Enterprise plans, each publishing event in audit logs now records the commit that was published, allowing a published app to be traced back to its code.

— Clear addition, no navigation detail given Nov 12, 2026 (2026-11-13)

30 **Inline audio playback in chat** Not checked yet NEW 45

Audio files attached in the project chat or included in Lovable's replies now play inline, with controls to jump to any point, loop, change volume, or set playback speed.

— Lists controls but no further mechanism Nov 12, 2026 (2026-11-13)

31 **Privacy & security settings grouped into five tabs**

Not checked yet IMPROVED 45

The workspace Privacy & security settings page is now grouped into five tabs: Access, Publishing, Security, Integrations, and Data protection.

— Names all five tabs but no behaviour change detail Nov 3, 2026 (2026-11-04)

32 **Email now sent from apex domain** Not checked yet IMPROVED 40

Email sent from a domain purchased through Lovable now comes from the apex domain with no subdomain, such as `noreply@yourdomain.com`.

— States the change with an example but no further detail Nov 13, 2026 (2026-11-16)

NEW INTEGRATION Jira connector

NEW INTEGRATION Mailjet connector

UPDATED INTEGRATION Sentry connector — in-app support

UPDATED INTEGRATION Resend connector — OAuth sign-in

NEW INTEGRATION PayPal integration

└──▷ ALSO FROM THESE RELEASES

.
.
.
.

└──▷ BREAKING ON UPGRADE

! The AI edits count has been removed from project settings; Credit usage in the same section now shows what the project has spent.

! Auto-accept invitations has been removed from account settings; workspace and project invitations must now be accepted manually from email or your Inbox.

> Nov 9, 2026 (2026-11-09) BLOG ↗ STABLE

> Nov 8, 2026 (2026-11-08) BLOG ↗ STABLE
>
> Lovable adds a PayPal integration supporting checkouts, payment capture, refunds, and order reads, accessible from Connectors or via in-project prompt.
>
> Add a PayPal integration to create checkouts, capture payments, issue refunds, and read order details — available from Connectors or by asking Lovable in a project.
>
> Not checked yet

WAS THIS USEFUL? 👍 👎

◆ Model Training & Fine-Tuning

Hugging Face Accelerate

Feedback ⓘ v1.16.0 NOTES ↗

RANK

Accelerate launches and runs PyTorch training across CPUs, GPUs, and TPUs with automatic mixed precision and configurable FSDP or DeepSpeed backends.

> Accelerate v1.16.0 broadens device backend support with torch.accelerator and initial single-chip TPU support, adds DeepSpeed sequence-parallel configuration and FSDP2/tensor-parallel workflow improvements, while raising minimum Python/PyTorch requirements and removing several deprecated integrations.

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

01 Removed FP8RecipeKwargs in favor of AORecipeKwargs/TERecipeKwargs

Not checked yet BREAKING 65

The deprecated `FP8RecipeKwargs` class has been removed — use `AORecipeKwargs` for TorchAO or `TERecipeKwargs` for Transformer Engine instead.

– Names exact classes and their replacements v1.16.0

02 FSDP2 model preparation without optimizer

Not checked yet NEW 60

Accelerate can now prepare a model without an optimizer under FSDP2, enabling workflows such as Trainer.evaluate() before training begins.

– Names concrete workflow and API call enabled v1.16.0

THINNER COVERAGE BELOW

03 Broader torch.accelerator backend support with TPU

Not checked yet NEW 57

Accelerate now uses `torch.accelerator` more broadly to support registered PyTorch backends without requiring a separate device-specific branch throughout the library, and adds initial single-chip TPU support via `torch-tpu`.

– Names mechanism and backend but no usage example v1.16.0

04 Raised minimum Python and PyTorch versions

Not checked yet 55

BREAKING

Accelerate now requires Python 3.11+ and PyTorch 2.6+; older environments will fail to install or run.

— Clear version thresholds but no migration mechanism beyond upgrade v1.16.0

05 **Clean shutdown for accelerate launch** Not checked yet IMPROVED

50

— placeholder v1.16.0

06 **DeepSpeed sequence-parallel configuration options** Not checked yet

NEW

45

New DeepSpeed configuration options let users distribute sequence-dimension workloads across ranks for sequence-parallel training.

— States purpose but no config keys or commands shown v1.16.0

07 **Tensor parallelism with replicated data parallelism**

Not checked yet NEW

35

Accelerate now allows tensor parallelism to be used alongside replicated data parallelism.

— Bare description of capability with no mechanism or example v1.16.0

08 **Removed PiPPy and MS-AMP integrations** Not checked yet DEPRECATED

30

The PiPPy and MS-AMP integrations have been removed from Accelerate.

— Names removed integrations but no replacement guidance v1.16.0

└──▷ **BREAKING ON UPGRADE**

! Accelerate now requires Python 3.11+ and PyTorch 2.6+; older environments will fail to install or run.

! The PiPPy and MS-AMP integrations have been removed.

! The deprecated `FP8RecipeKwargs` class has been removed — use `AORecipeKwargs` for TorchAO or `TERecipeKwargs` for Transformer Engine instead.

WAS THIS USEFUL?  

Preferred Networks Optuna

Feedback ⓘ v5.1.0 NOTES ↗

RANK

Optuna provides automated hyperparameter optimization with programmable search spaces, sampling algorithms, and trial pruning.

> Optuna v5.1.0 adds Python 3.15 support, drops Python 3.9, lets NSGAIIISampler accept plain Python lists for reference points, speeds up two visualization functions for large studies, and hardens `download_artifact` against race conditions.

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

01 Plain list support for NSGAIIISampler reference_points

Not checked yet IMPROVED 70

NSGAIIISampler now accepts a plain Python `reference_points` list directly, removing the need to convert it to a NumPy array before passing it in.

Run a multi-objective study with reference points supplied as a plain list, without converting to NumPy first.

python

```python
import optuna
from optuna.samplers import NSGAIIISampler

sampler = NSGAIIISampler(
    reference_points=[[0.0, 1.0], [0.5, 0.5], [1.0, 0.0]]
)
study = optuna.create_study(directions=['minimize', 'minimize'],
sampler=sampler)
study.optimize(objective, n_trials=200)
```

— Names the exact parameter and shows a runnable example v5.1.0

02 Faster plotting for large studies Not checked yet IMPROVED

65

`plot_contour` now runs faster by replacing O(N²) `list.index` lookups with dict lookups, and `plot_pareto_front` now runs faster by replacing O(N × |P|) list lookups in `_get_non_pareto_front_trials` with a set-based approach.

— Explains mechanism and complexity change but needs no user action v5.1.0

THINNER COVERAGE BELOW

03 **Python 3.9 support dropped** Not checked yet BREAKING 50

Python 3.9 is no longer supported; users must upgrade to Python 3.10 or newer before installing v5.1.0.

— Clear migration note but minimal further explanation v5.1.0

04 **Safer artifact downloads** Not checked yet IMPROVED 35

`download_artifact` now guards against overwriting a destination file that was created after the download began.

— Describes the fix but no concrete usage detail v5.1.0

05 **Python 3.15 support** Not checked yet NEW 25

Optuna now supports Python 3.15.

— Bare statement with no further detail v5.1.0

└──▷ **BREAKING ON UPGRADE**

! Python 3.9 is no longer supported; upgrade to Python 3.10 or newer before installing v5.1.0.

WAS THIS USEFUL? 👍 👎

◆ Speech & Voice AI

Parakeet ASR Server

Feedback ⓘ v0.12.0 NOTES ↗ RANK

Parakeet ASR Server transcribes audio on CPUs or GPUs using NVIDIA Parakeet ONNX models through a Whisper-compatible API.

> Parakeet ASR Server adds on-demand and idle-based model unloading to reclaim up to roughly 1 GB of RAM or VRAM when the server sits idle, with API endpoints to control and inspect model residency.

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

01 On-demand and idle-timeout model unloading

Not checked yet NEW

93

The ASR model, VAD, and diarization models can be unloaded automatically after a period of inactivity by setting `-model-unload-after-idle-duration` (e.g. `5m`); the default `0` keeps models resident. Unloading can also be triggered immediately via `POST /v1/models/unload` and reloaded via `POST /v1/models/load` without restarting the server, and `GET /v1/models` reports whether weights are currently in memory. Models reload automatically on the next incoming request after an unload, taking about 3.5s on CPU, and idle unloading has been shown to cut RSS from ~1GB to ~48MB on CPU.

Start the server so it automatically frees model memory after 5 minutes of inactivity, cutting RSS from ~1 GB to ~48 MB on CPU.

```
$ parakeet -model-unload-after-idle-duration 5m
```

Immediately reclaim VRAM or RAM from a running server without restarting it, then verify the weights are gone.

```
$ curl -X POST http://localhost:8080/v1/models/unload
curl http://localhost:8080/v1/models
```

Reload model weights back into memory on demand before a batch job so the first real request does not pay the 3.5 s reload cost.

```
$ curl -X POST http://localhost:8080/v1/models/load
```

— Names flags, endpoints, defaults, timing, and memory figures concretely. v0.12.0

WAS THIS USEFUL? 👍 👎

OpenWhispr

Feedback ⓘ v1.12.0 NOTES ↗ RANK

OpenWhispr is a desktop dictation and meeting-transcription app that supports local speech models and user-configured cloud providers.

> OpenWhispr's v1.12.0 release adds a floating, hotkey-activated assistant window, private per-workspace notes, Teams management with a People page, a one-screen local model setup backed by a signed model registry, background/parallel chats, organisation-level device and model policies, OpenRouter speech-to-text, Dragon vocabulary import, and configurable default summary templates.

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

01 Signed model registry with privacy control

Not checked yet NEW

70

A signed model registry pushes new and retired models between releases, and can be controlled in Settings → Privacy. Disabling it there prevents automatic model updates from arriving between releases.

Disable the signed model registry to prevent automatic model updates arriving between releases.

📍 Go to Settings → Privacy and turn off the signed model registry.

– Exact UI path to disable gives a reader something to act on v1.12.0

02 Floating hotkey-activated assistant window

Not checked yet NEW

62

A floating assistant window reachable by hotkey from any app lets you speak or type, edit selected text, fill form fields, draft emails and Slack replies, ask mid-task questions, and keep chats alive when minimised.

– Describes behaviour and use cases but no exact command or shortcut given v1.12.0

THINNER COVERAGE BELOW

03 People page for Teams management

Not checked yet NEW

55

A new People page adds bulk invites, bulk removals, and a teamspace overview for Teams and Guests.

– Names concrete admin actions but no exact navigation or API v1.12.0

04 Organisation device, update, and model policies

Not checked yet

NEW

55

Organisation controls cover managed device policy, update policy, and local model allow lists.

— Names three policy areas but no config keys or endpoints v1.12.0

05 **One-screen local model setup** Not checked yet NEW

45

A one-screen local setup recommends and downloads speech models suited to your machine.

— Explains behaviour but no specifics on model names or limits v1.12.0

06 **Background and parallel chats** Not checked yet NEW

38

Background chats keep generating after you leave them, and multiple chats can run in parallel.

— Describes behaviour only, no configuration or limit named v1.12.0

07 **Private notes in workspaces** Not checked yet NEW

35

A Private notes section is added to every workspace and syncs across devices.

— Bare description, no UI path or sync mechanism detailed v1.12.0

08 **OpenRouter speech-to-text option** Not checked yet NEW

35

OpenRouter speech-to-text is added as a transcription option.

— Names the provider but no setup detail given v1.12.0

09 **Dragon vocabulary import** Not checked yet NEW

35

You can now import your Dragon vocabulary into OpenWhispr.

— Names the source but no import steps or format given v1.12.0

10 **Configurable default summary template** Not checked yet IMPROVED

30

You can now choose your default summary template.

— Bare mention, no template options or location named v1.12.0

WAS THIS USEFUL? 👍 👎

◆ Coding Agent Add-ons

Token Optimizer

Feedback ⓘ v5.13.72 NOTES ↗ RANK

Token Optimizer is a coding-agent plugin that trims wasted context, checkpoints work across sessions and compactions, and reports token spend on a live dashboard.

> Token Optimizer shipped a self-contained HTML dashboard with configurable history windows, a new Cursor platform adapter with eleven wired hook events and read-cache interception, and an opt-in code map session hint for large codebases.

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

01 **Cursor platform adapter with hooks and read-cache** Not checked yet

NEW 90

Merges eleven Cursor hook events into `additional_context` — including `afterAgentResponse`, `postToolUseFailure`, `afterShellExecution`, `subagentStart`, and `subagentStop` — recording token fields and per-tool failure counts for offline correlation. Compresses Bash output and intercepts redundant file reads via a shadow-first read-cache configurable as `soft_block`, `block`, or `shadow` mode in `read-cache.json`; adds an opt-in status-line command via `cursor_statusline.py` showing context fill (`ctx ?` until verified); and pairs shell rewrites to executions by `tool_use_id`, with rewrites disabled (record-only) on native Windows due to unverified POSIX `shlex` quoting.

Switch read-cache interception from silent shadow recording to active blocking so Token Optimizer injects cached file content and denies redundant reads.

```json
{ "mode": "soft_block" }
```

— Detailed mechanism and named config/file surfaces, but no enable command product docs

02 **Self-contained HTML dashboard with configurable history**

Not checked yet NEW 86

Generates and opens a self-contained HTML dashboard covering Overview, Savings, Quick Wins, Medium, Deep, Habits, Trends, Health, Coach, Hooks, and Manage views, with `--user` to mark a person-initiated run distinct from hook invocations. History embedded at build time is controlled with `--days all` (default 30 days on new installs); the Trends view offers display ranges of 7, 14, 30, 90, 180 days or All that persist across regenerations and refreshes, and when served locally a range longer than the embedded window loads via a Load full history action without a full rebuild.

Kick off a person-initiated dashboard build that embeds all available history, so the Trends view can show any range without hitting the server.

```
$ token-optimizer dashboard --user --days all
```

— Names flags, views and ranges with a runnable command `product docs`

03 **Opt-in code map session hint** `Not checked yet` `IMPROVED`

78

The code map session hint is now opt-in via `TOKEN_OPTIMIZER_CODEMAP_HINT=1`; it indexes up to 6,000 files, falls back to a read-only index in sandboxes, and is cleaned up daily on macOS and Linux.

Opt into the code map session hint so the tool can index your project (up to 6,000 files) and pass context to the AI coding assistant without incurring token cost by default.

```
$ TOKEN_OPTIMIZER_CODEMAP_HINT=1 claude
```

— Clear env var and limits, exact command given `v5.13.72`

WAS THIS USEFUL? 👍 👎

Paradev

Feedback ⓘ cli-v2.0.9 NOTES ↗ RANK

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

> Paradev expanded platform reach with Android and web support and added per-environment cloud machine sizing, stacked-merge PR workflows, and concurrent background subagent chat, alongside a string of smaller UI and plugin additions.

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

01 **Per-environment machine sizing and management** Not checked yet

NEW 75

Adds environment management: set up environments from chat, configure per-environment machine size, and manage environments from a new environment page in Settings. The database schema now stores per-environment machine sizing via `environments.machine_size`, and users can view box pressure for every cloud workspace and change machine size directly from the UI.

Resize a cloud workspace's machine directly from the UI when a workspace is under heavy load.

> 📍 In the cloud workspace tab bar, open the box pressure indicator, then select a new machine size from the change machine size control.

— Names config key and gives exact UI path cli-v2.0.9 desktop-v2.0.9

THINNER COVERAGE BELOW

03 **Pages view enhancements** Not checked yet NEW 50

Users can duplicate a page directly from the pages view, e.g. via right-click and Duplicate, and pressing ⌘F while a Pages pane is focused opens a find bar to search within the page, useful for long wiki or documentation pages.

When iterating on a draft page, duplicate it to preserve the original while experimenting with edits.

> 📍 In the Pages view, right-click (or open the context menu on) the page you want to copy and select Duplicate.

Search for a term inside a Pages pane without leaving the keyboard — useful when reviewing long wiki or documentation pages.

📍 Press ⌘F while a Pages pane is focused to open the find bar and search within the page.

— Simple action with clear UI step Exact keystroke given, runnable immediately cli-v2.0.9 desktop-v2.0.9

04 **Ports button in cloud workspace tab bar** Not checked yet NEW

50

A new Ports button beside Share in the cloud workspace tab bar gives quick access to forwarded ports, letting users view and manage forwarded ports without leaving the editor.

Inspect and manage forwarded ports for a running cloud workspace without leaving the editor.

📍 In the cloud workspace tab bar, click the Ports button (beside Share) to view and manage forwarded ports.

— Exact UI location with runnable steps given cli-v2.0.9 desktop-v2.0.9

05 **Workspace moves carry terminal sessions** Not checked yet IMPROVED

45

Moving a workspace to the cloud from any of your computers now carries over Claude terminal sessions, and the Move submenu no longer requires a confirmation dialog.

— Describes before/after behavior without full mechanism cli-v2.0.9 desktop-v2.0.9

08 **Ahrefs MCP plugin added** Not checked yet NEW

35

Adds an Ahrefs MCP plugin to the plugins catalogue.

— Names the plugin but not its capabilities desktop-v2.0.9

WAS THIS USEFUL? 👍 👎

AGENT

◆ AI Agent Frameworks

Arcade

Feedback ⓘ updated · Nov 18, 2026 NOTES ↗ RANK

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

> Arcade's most concrete changes this window are deployment and governance controls — a new `apt_packages` config key for baking system packages into deployed servers, a per-project pending-deployment cap, and project API keys that can be scoped and deactivated — alongside breaking removals of the `porter` deploy provider, several runner-image dependencies, and required Engine database migrations.

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

01 apt_packages for deployed servers

Not checked yet NEW

85

Declare system-level Debian packages a deployed server needs with the `apt_packages` field in `.arcade/deploy.json`; the packages are installed at build time and baked into the image rather than installed on every pod start.

Add a system Debian package (e.g. `libpq-dev`) that your deployed server needs at build time, so it is baked into the image rather than installed on every pod start.

```json
{
  "apt_packages": ["libpq-dev", "libxml2"]
}
```

— Names exact config key, file, and example packages with a runnable snippet.

snapshot-20261118

02 Pending deployment cap per project

Not checked yet NEW

80

Cap pending deployments per project with the `deployments.max_pending_per_project` setting in `engine.extraConfig`, off by default.

Prevent a runaway CI pipeline from queuing unlimited deployments by capping pending deployments per project to 3.

```yaml
deployments:
  enabled: true
  max_pending_per_project: 3
```

— Named config key and file with a runnable config example. snapshot-20261118

---

03 **Tool integration improvements across Gmail, Docs, Drive, GitHub**

Not checked yet IMPROVED 80

Several tool integrations improved: Gmail's `GetThread` now accepts a flag to omit each message's HTML body; Google Docs' `EditDocument` now operates on only named tabs instead of the whole document; Google Drive's `SearchFiles` now searches every shared drive by default, flags incomplete results, reports the last editor, and makes its query argument optional; GitHub's `DownloadFile` on a directory now explains the situation and lists its contents instead of failing silently; and GitHub GraphQL tools now work on GitHub Enterprise Server.

— Names every tool and its exact changed behavior, though no flag names given.

snapshot-20261118

---

04 **Server-side search, filter, and sort in admin list views**

Not checked yet IMPROVED 65

Several admin list views now search, filter, and sort server-side: the secrets list (which also reports whether each secret is the currently active one for its key), the OAuth providers list, the connected-users picker, the gateway-requirement form, and the duplicate-name check.

— Names every affected view; UI path is a clear but not runnable starting point.

snapshot-20261118

---

05 **Required Engine database migrations** Not checked yet BREAKING 60

The Engine database migration drops the old account-based gateway binding and blocks while any gateway still uses it (required migration), and Engine tool-execution-history API-key/gateway columns are now nullable (required migration).

— Names the required migrations and blocking condition clearly. snapshot-20261118

THINNER COVERAGE BELOW

06 **Client ID Metadata Document authorization** Not checked yet NEW 50

Support Client ID Metadata Document authorization for servers that require it, with authorization requests and responses now recording an expected issuer.

— Names the auth mechanism and new issuer field but no config surface. snapshot-20261118

07 **Runner image dependencies removed** Not checked yet BREAKING

50

The runner image's preinstalled `unixODBC`, `libaio`, GSSAPI, and Oracle Instant Client compatibility link have been removed; see Upgrade notes.

— Names all removed dependencies but no replacement guidance given. snapshot-20261118

08 **Scoped and deactivatable project API keys** Not checked yet NEW

45

Admins can now deactivate a project API key immediately without deleting it, leaving its secret intact, and when creating a key they can choose its scope — call tools, manage the project, or both — instead of always granting full access.

— Explains behavior and scope options but no exact UI/API path given. snapshot-20261118

09 **Registry catalog browsing via gateway tools** Not checked yet NEW

45

A gateway can now expose two Registry tools letting a signed-in agent browse the full project MCP catalog, controlled by a per-gateway setting that is off by default.

— Describes the capability but the per-gateway setting is unnamed. snapshot-20261118

10 **Distinct error kind for tool refusals** Not checked yet IMPROVED

45

Tool execution results now report a new error kind for deliberate tool refusals (bad input, business rules), distinct from the fatal-error kind reserved for unhandled crashes.

— Clear behavior distinction but no field name or API shape given. snapshot-20261118

11 **Dashboard list pagination defaults** Not checked yet IMPROVED

45

Dashboard tables and lists now open at 20 rows, offer up to 200 rows (100 for audit logs), and remember your last page size per table.

— Concrete numbers but no actionable config or navigation. snapshot-20261118

12 **Porter deployment provider removed** Not checked yet BREAKING

35

The `porter` deployment provider has been removed; see Upgrade notes.

— Names the removed provider but gives no migration detail. snapshot-20261118

└──▷ BREAKING ON UPGRADE

! The `porter` deployment provider has been removed; see Upgrade notes.

! The runner image's preinstalled `unixODBC`, `libaio`, GSSAPI, and Oracle Instant Client compatibility link have been removed; see Upgrade notes.

! The Engine database migration drops the old account-based gateway binding and blocks while any gateway still uses it (required migration).

! Engine tool-execution-history API-key/gateway columns are now nullable (required migration).

WAS THIS USEFUL?  

OpenAI Agents SDK

Feedback ⓘ v0.25.0 NOTES ↗ RANK

OpenAI Agents SDK is an open-source framework that provides tools, handoffs, guardrails, and tracing for agentic applications.

> OpenAI Agents SDK v0.25.0 exposes new utility APIs for building third-party session adapters while removing the Runloop sandbox provider.

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

01 **Runloop sandbox provider removed** Not checked yet BREAKING

55

The Runloop sandbox provider has been removed from the `sandbox` module; any code referencing it will break on upgrade.

— Names exact module and breaking impact, no migration path given v0.25.0

02 **Utility APIs for third-party session adapters** Not checked yet

NEW

30

The SDK now exposes utility APIs that let developers build third-party session adapters.

— Names capability but no API names or usage shown v0.25.0

└──▷ **BREAKING ON UPGRADE**

! The Runloop sandbox provider has been removed from the `sandbox` module; any code referencing it will break on upgrade.

WAS THIS USEFUL? 👍 👎

HeyGen HyperFrames

Feedback ⓘ v0.9.65 NOTES ↗ RANK

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

> HyperFrames' biggest changes this window were security-focused: a sandboxed Studio preview origin (`sandboxOrigin` / `VITE_STUDIO_OPAQUE_PREVIEW_ENABLED`) and a new restriction requiring `--include-outside-files` for trusted renders that reference outside files. Alongside these, the CLI gained new diagnostic and layer-scoring commands (`check`, `measure`, `diff --hide`), the rendering pipeline got performance overlap improvements, and Studio's API, exports, and sealed-preview sharing all picked up incremental additions.

└──▷ **WHAT SHIPPED · 13 FEATURES** most completely described first ? what's the number?

01 Studio Preview Sandboxing

Not checked yet NEW 90

Studio preview can now run on its own sealed origin via `sandboxOrigin` or `VITE_STUDIO_OPAQUE_PREVIEW_ENABLED` (off by default), continuing to support images, media, fetched files, the console errors panel, GSAP keyframe previews, and 3D previews inside the sealed context; Studio hosts can also share the preview's editing-port client with other users and seal the preview frames that NLEPreview draws, with the sealed preview exporting its checked element reads and stand-ins and forwarding Studio's playback keys.

– Exact config keys, default, and preserved behaviors listed. Names surfaces like NLEPreview but no setup steps shown.

v0.9.54 v0.9.58

02 Trusted renders restricted to project-linked files

Not checked yet BREAKING 80

Trusted renders now refuse outside-linked compositions unless the `--include-outside-files` flag is passed, e.g. `hyperframes render --include-outside-files`.

Render a trusted composition that references files outside the project folder — previously this would be refused outright.

```
$ hyperframes render --include-outside-files
```

– Exact flag and runnable command given. v0.9.61

03 Layer scoring via diff --hide

Not checked yet NEW 80

Score a project's layers by what each one contributes using the new `diff --hide` CLI flag, e.g. `npx hyperframes diff --hide`.

Compare project layers by their contribution — useful for identifying which layers add the most to a composition before a final render.

```
$ npx hyperframes diff --hide
```

— Exact flag and runnable command given. v0.9.59

04 **CLI check and measure command enhancements** Not checked yet NEW

70

`hyperframes check` now detects text overlaid on a face in a composition, and `hyperframes measure` reports hand presence in addition to trimming the names of resumed takes in its output.

— Named commands and behaviors, but no sample output shown. v0.9.64

05 **Upload and config file manifest tracking** Not checked yet IMPROVED

60

Uploads now list every file they carry from outside the project, giving a full manifest of external assets included in an upload; the `add` command also now lists files from a received config's registry as received.

— Names the `add` command but not the manifest's exact format. v0.9.53

THINNER COVERAGE BELOW

06 **Rendering pipeline overlap improvements** Not checked yet IMPROVED

55

The engine now checks still frames across several pages at once, so text films start rendering sooner, and the producer now starts drawing frames while the audio mix is still finishing — overlapping work that was previously sequential.

— Clear before/after, but internal change with no exposed control. v0.9.63

07 **Output-dimension enforcement before capture** Not checked yet NEW

50

HyperFrames' Producer now enforces expected output dimensions before capture begins, catching mismatches before any frames are rendered.

— What it does is clear; no config to adjust tolerance. v0.9.61

08 **Source-parsed animations in Studio API** Not checked yet NEW

50

HyperFrames' Studio API now exposes an element's source-parsed animations, making animation metadata available to integrations and automations.

— Names the API surface but no endpoint or schema given. v0.9.61

09 **Live geometry streaming from preview to editor** Not checked yet

NEW

45

The preview can now stream an element's geometry to the editor on every frame, enabling live positional feedback during playback.

— Mechanism described but no UI entry point named. v0.9.63

10 **Describe label and standInHandle export field** Not checked yet

IMPROVED

45

Describe replies now carry the picker's element label, and Studio exports now include a `standInHandle` field.

— Field named but no schema or use case detailed. v0.9.59

11 **Render telemetry for still-frame reuse failures** Not checked yet

IMPROVED

40

HyperFrames' render telemetry now reports why a page cannot reuse still frames, surfacing the reason in Engine output.

— No list of possible reasons or log format shown. v0.9.61

12 **CLI command for paid plan management** Not checked yet NEW

40

Adds a CLI command to show and open the page to manage or cancel a paid plan.

— Describes purpose but not the exact command invocation. v0.9.58

13 **403 rejection metric in Studio Server** Not checked yet NEW

35

Studio Server now counts the API requests it refuses with `403`, giving a host visibility into rejected traffic.

— No metric name, dashboard, or threshold given. v0.9.65

WAS THIS USEFUL? 👍 👎

PydanticAI

Feedback ⓘ v2.70.0 NOTES ↗ RANK

PydanticAI is a framework that builds type-safe Python agents with dependency injection, model integrations, tools, and structured outputs.

> PydanticAI's v2.70.0 release adds four new agent capabilities for streaming guardrails, tool-result pruning, model retries, and tool search, alongside sub-agent task control, Temporal durability improvements, and Skills file access.

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

01 **StreamGuardrail capability for streamed output** Not checked yet

NEW 70

`StreamGuardrail` screens each streamed text, thinking, and tool-call part before the caller sees it, letting you flag or block model output before it reaches users.

Guard a streamed agent response so your code never sees a flagged output part — useful when model output reaches users directly.

python

```python
from pydantic_ai_harness import StreamGuardrail

from pydantic_ai import Agent

agent = Agent(
    'openai:gpt-6.1-sol',
    capabilities=[StreamGuardrail(...)],
)
```

— Runnable example and clear mechanism, but no detail on flagging rules. v2.70.0

02 **PruneToolResults capability** Not checked yet NEW 70

`PruneToolResults` lets the model replace a tool result it just received with a summary, used to keep the context window small.

Replace a bulky tool result with a model-written summary mid-conversation to keep the context window small.

```python
from pydantic_ai_harness import PruneToolResults

from pydantic_ai import Agent

agent = Agent(
    'openai:gpt-6.1-sol',
    capabilities=[PruneToolResults(...)],
)
```

— Named class with runnable example and stated purpose. v2.70.0

03 **Sub-agent delegation task controls** Not checked yet NEW

65

SubAgent(on_model_error=) lets you choose retry or return for a failed sub-agent independently of `on_failure`, and a finished sub-agent task can be resumed via `message_task`. A parent can also check a running background delegation using `check_task` and DelegationTasks.snapshot().

— Several named APIs but no code example of use. v2.70.0

04 **RetryOnModelErrors capability** Not checked yet NEW

60

`RetryOnModelErrors` retries the same model automatically on transient model API errors, avoiding custom retry logic.

Automatically retry the same model after transient API errors without writing custom retry logic.

```python
from pydantic_ai.capabilities import RetryOnModelErrors

from pydantic_ai import Agent

agent = Agent(
    'openai:gpt-6.1-sol',
    capabilities=[RetryOnModelErrors()],
)
```

— Clear runnable example but no retry limits or backoff behavior given. v2.70.0

05 **Durable execution controls for Temporal** Not checked yet NEW

60

`register_workflows` was added to Temporal's `AgentPlugin` so activity-only workers can skip child workflow registration, and `ToolDefinition.durable` lets a

tool run outside its durable unit on every durable execution engine.

— Named config/flag surfaces but no example or before/after shown. v2.70.0

THINNER COVERAGE BELOW

06 **Skills bundled file access** Not checked yet NEW 55

`Skills` can now read a loaded skill's bundled files with `read_files=True`, and `${SKILL_DIR}` placeholders are substituted.

— Named flag and placeholder but no usage example. v2.70.0

07 **ToolSearch tools selector** Not checked yet NEW 40

`ToolSearch` is a tools selector added for pre-search deferral.

— Named but no usage example or mechanism detail provided. v2.70.0

08 **CapabilityPart request part** Not checked yet NEW 35

`CapabilityPart` is a new request part representing what a capability tells the model, rendered by the framework before any model sees it.

— Bare description of a new part type with no example. v2.70.0

NEW SUPPORTED PLATFORM Gemini 4 Hydrogen model support

NEW SUPPORTED PLATFORM MiniMax model profile for Bedrock, OpenAICompatibleProvider, LiteLLMProvider, TogetherProvider

NEW SUPPORTED PLATFORM Windows host support via WindowsWorkspace

NEW INTEGRATION Kubernetes sandbox integration

NEW SUPPORTED PLATFORM Mistral Large 4 reasoning_effort mapping

v2.69.0 NOTES ↗ STABLE

PydanticAI adds peer messaging, loop detection, Windows/Kubernetes sandboxes, tool budgets, sub-agent partial output, and new model settings for Groq, OpenAI, and Anthropic data residency.

└──▷ GET THIS VERSION ▾

pip

```
$ pip install clai==2.69.0
# already installed? upgrade to this version:
$ pip install --upgrade clai==2.69.0
```

```git
$ git clone --branch v2.69.0 https://github.com/pydantic/pydantic-ai.git
# already have the repo? check out this version:
$ git checkout v2.69.0
```

PyPI attestation · checked 18 Nov Wrong command?

Add `anthropic_inference_geo` setting to `AnthropicModel` to control Anthropic data-residency region. Not checked yet

Add `service_tier` to `GroqModel` for tiered inference priority. Not checked yet

Add `openai_safety_identifier` model setting for OpenAI requests. Not checked yet

Add `tool_call_ids_are_response_scoped` profile flag to generalize response-scoped tool call ID handling. Not checked yet

Add `is_retry` to message history to identify retried messages. Not checked yet

+24 more

└─▷ **BREAKING ON UPGRADE**

- ! `RequestInput` is folded into `HandleDeferredToolCalls`; its handler can now take a resume token — code using `RequestInput` directly must migrate.
- ! `tool_call_ids_are_response_scoped` replaces the previous response-scoped tool call ID mechanism — any code relying on the old behaviour must update to use the new profile flag.
- ! Advisor's `mode` parameter is deprecated in favour of `native=` and `local=`; the old `mode` kwarg will be removed in a future release.
- ! Harness `subagents`, `filesystem`, and `spend` packages are renamed to their capability names, and `ACP*` and `*Error` class names are aligned — imports from the old package paths will break.
- ! `DockerSandbox` is renamed to `DockerContainer`; `DockerSandbox` remains as a deprecated alias.
- ! `BubblewrapSandbox` is renamed to `BubblewrappedWorkspace`.

WAS THIS USEFUL? 👍 👎

LangChain

Feedback ⓘ langchain-core==1.8.2 NOTES ↗ RANK

LangChain is an open-source framework that orchestrates applications powered by language models.

> LangChain expanded LangSmith gateway configuration and routing across langchain-core, langchain-mistralai and langchain-groq, while langchain-groq gained a performance service tier and explicit context-overflow handling; langchain-mistralai now drops Python 3.10 support.

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

01 Groq performance tier and context-overflow handling

Not checked yet NEW 85

`ChatGroq` adds a `performance` service tier (set via `service_tier="performance"`) for Groq's performance-optimised, low-latency inference path, and now maps context-length errors to a new `ContextOverflowError` exception so callers can catch that specific type instead of a generic error.

Catch context-length overflows explicitly in a Groq-backed chain so your app can truncate or retry rather than crash on a generic exception.

python

```python
from langchain_groq import ChatGroq
from langchain_core.exceptions import ContextOverflowError

llm = ChatGroq(model="llama3-8b-8192")

try:
    result = llm.invoke(very_long_prompt)
except ContextOverflowError:
    # truncate input and retry
    result = llm.invoke(very_long_prompt[:4000])
```

Use the `performance` service tier when you need Groq's lowest-latency inference path for time-sensitive workloads.

```python
from langchain_groq import ChatGroq

llm = ChatGroq(model="llama3-8b-8192",
service_tier="performance")
result = llm.invoke("Summarise this incident report in one
sentence.")
```

— Runnable code examples show both new capabilities directly `langchain-groq==1.2.0`

02 **LangSmith gateway configuration and routing** Not checked yet

IMPROVED

67

The `langsmith_gateway` configuration in langchain-core now supports a `headers` field to pass custom HTTP headers (with proxy traffic ignored on that path), and can be overridden per context via the `langsmith_gateway` setting to route tracing traffic to different endpoints. `ChatMistralAI` and `ChatGroq` calls are now routed through the LangSmith gateway to enable tracing and observability.

— Names config key and headers field but no full example shown `langchain-groq==1.2.0` `langchain-mistralai==1.2.0` `langchain-core==1.8.2` `langchain-core==1.8.1`

THINNER COVERAGE BELOW

03 **Python 3.10 dropped in langchain-mistralai** Not checked yet

40

BREAKING

langchain-mistralai==1.2.0 drops support for Python 3.10; Python 3.11 or later is now required to use the package.

— States the requirement but gives no migration guidance `langchain-mistralai==1.2.0`

└──▷ **BREAKING ON UPGRADE**

! Python 3.10 is no longer supported in langchain-mistralai; Python 3.11 or later is required.

WAS THIS USEFUL? 👍 👎

CopilotKit

Feedback ⓘ v1.89.0 NOTES ↗ RANK

CopilotKit connects AI agents to application interfaces with shared state, tools, and generative UI components.

> CopilotKit's v1.89.0 release improves performance for long Intelligence threads and adds agent display names plus model-context sharing for MCP Apps.

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

01 **Distinct display names for agents** Not checked yet NEW

35

CopilotKit now supports distinct display names for agents, letting multi-agent setups surface each agent's identity separately in the UI.

– Describes behaviour but no config key or API v1.89.0

02 **Model-context sharing for MCP Apps** Not checked yet NEW

35

MCP Apps now support model-context sharing, allowing agents to share context with connected MCP applications.

– General description without named surface or example v1.89.0

03 **Improved performance for long Intelligence threads**

Not checked yet IMPROVED

20

CopilotKit Intelligence threads now handle long conversations with improved performance and reliability.

– No mechanism or metrics given, just a claim v1.89.0

WAS THIS USEFUL? 👍 👎

cotal

Feedback ⓘ v0.166.1 NOTES ↗ RANK

Cotal is the open standard for agent coordination

> Cotal's biggest window change was a short-lived cross-session Claude permission-approval API that shipped in v0.165.0 and was fully removed two releases later, alongside lasting additions like a multi-channel history API, in-app live terminal access, a connector-less `cotal mcp` gateway, persistent broker configuration, and Windows tray auto-updates.

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

01 Dashboard session and posting controls

Not checked yet NEW

85

Open an authenticated local dashboard browser session with `cotal service start web`; pass `--no-open` to preserve manual access while still reusing a running dashboard. Dashboard service posting can now be enabled with an explicitly supplied separate credential, while read-only mode remains the default.

Open an authenticated dashboard in your local browser, reusing any already-running dashboard instance; skip the auto-open to preserve manual access.

```
$ cotal service start web --no-open
```

— Exact command and flag with behavior fully detailed. v0.166.1

02 Cross-session Claude permission approval, added then removed

Not checked yet BREAKING

85

v0.165.0 added `CotalEndpoint.openPermissionHolder`, `openPermissionReader`, `watchPermissions` and `answerPermission` so a Claude agent's permission prompts could be approved or denied from a separate session, with answers riding the `cotal.<space>.perm.<agentOwner>.<agentActor>.<runtimeId>.<owner>.<actor>` rail and `permissionRequests` reported on every Manager `status.connectors` row. v0.166.0 then removed all of it: the `perm` answer subject, the `cotal_permissions_<space>` bucket and its readers, the `@cotal-ai/core/permissions` export, the `request` field on a presence condition, and `permissionRequests` on `status.connectors` are gone, and a Claude Code connector no longer holds a permission prompt answerable from another session.

Watch pending permission prompts for a Claude agent and answer them from a separate operator session without interrupting the agent's own dialog.

```javascript
const reader = await endpoint.openPermissionReader();
for await (const prompt of endpoint.watchPermissions(reader)) {
  await endpoint.answerPermission(prompt, { approved: true });
}
```

— Full lifecycle documented, but feature no longer exists. v0.166.0 v0.165.0

03 **Persistent broker settings via `cotal up --open`** Not checked yet

NEW 80

Configure the broker's message size limit and retain it across restarts with `cotal up --open --max-payload <bytes>`. Native broker log and rotation settings can now also be changed or enabled on a stopped stack and are kept for subsequent starts.

Cap the broker's message size to 10 MB when opening a space, so that large payloads are rejected and the limit survives daemon restarts.

```
$ cotal up --open --max-payload 10485760
```

— Runnable command with a concrete byte-size example. v0.165.4

04 **Multi-channel history API** Not checked yet NEW 80

Read the newest messages across a channel family in one call using CotalEndpoint.multiChannelHistory(['events.>'], { limit }), which uses one broker filter and returns subject-tagged, stream-ordered rows.

Fetch the latest messages across an entire channel family in one call, avoiding per-channel enumeration.

```javascript
const rows = await endpoint.multiChannelHistory(['events.>'], {
limit: 50 });
```

— Code example and filtering mechanism both given. v0.165.0

05 **Connector-less gateway via `cotal mcp`** Not checked yet NEW 75

Serve gateway tools from apps without a native Cotal connector using `cotal mcp`; the stdio host keeps one local mesh identity until disconnect and retires it on shutdown.

Expose Cotal gateway tools from an app that has no native connector, so it participates in the mesh without a full connector install.

```
$ cotal mcp
```

— Command and mesh-identity mechanism both named clearly. v0.166.1

06 **Persona command validation and draft safety** Not checked yet

IMPROVED 70

`cotal personas new --role` now refuses a value that is not a single NATS-safe token (`[A-Za-z0-9_-]`), matching the validation already applied by `cotal persona create` . `cotal persona edit` now keeps your edited draft and names its path when saving fails for any reason.

— Named commands and validation rule, no full example. v0.166.0

07 **Agent files reject cwd/worktree frontmatter** Not checked yet

65

BREAKING

Agent files now reject `cwd` and `worktree` frontmatter fields explicitly; set the directory on the spawn request with `--cwd` instead.

— Named fields and the exact migration flag. v0.165.0

08 **Named reconnect errors for missing DM durable** Not checked yet

IMPROVED 65

Reconnect errors for a missing DM durable after a broker outage now name the durable and suggest recovery via `cotal_reconnect` , with `connectionIssue` in `cotal_connection_status` surfacing the reason `dm-durable-missing` .

— Named recovery command and status reason code. v0.165.0

09 **Retained-generation provenance on self-update** Not checked yet

IMPROVED 65

Stage complete execution generations before the installed CLI is swapped during updates, keeping the original execution graph intact for running commands; inspect this provenance with `cotal status --verbose` or `cotal status --json` .

Check that a staged generation was retained correctly and inspect its provenance after a self-update.

```
$ cotal status --verbose
```

Parse retained-generation provenance programmatically in a CI script after running `cotal update --self` .

```
$ cotal status --json
```

— Named flags are runnable but mechanism stays brief. v0.164.16

10 **Daemon install reliability** Not checked yet IMPROVED 60

Linux daemon installation now accepts home and config directories containing spaces, and desktop service setup now waits up to 30 seconds to read the account's application PATH, reporting a timeout and advising `cotal daemon install` instead of checking a login shell on failure.

Re-register the Cotal daemon after a failed PATH read on Windows to recover without debugging your login shell.

```
$ cotal daemon install
```

— Thin one-line fix with no command detail. Named timeout value and exact recovery command.

v0.166.0 v0.164.17

THINNER COVERAGE BELOW

11 **Broker websocket port fallback in `cotal up`** Not checked yet

IMPROVED 55

`cotal up` now starts the broker on a new websocket port when another process took the picked port before the broker could bind it, instead of failing with `Unable to listen for websocket connections`.

— Describes before/after behavior; fix is automatic. v0.166.0

12 **Accurate outcome reporting for detached spawn** Not checked yet

IMPROVED 55

`cotal spawn --detach --on <instance>` now reports its real outcome when the manager briefly serves nothing right after accepting it, instead of ending at once with `OPERATION_UNKNOWN`.

— Named flag and error code, but automatic fix. v0.166.0

13 **Live terminal access to managed agents** Not checked yet NEW 55

Open a managed agent's live terminal in the app's middle column and interact with it via openTerminalSession() from `@cotal-ai/workspace/attach`.

— Named API and UI location, no runnable snippet. v0.165.0

14 **Codex message redelivery pause after repeated failures**

Not checked yet IMPROVED 55

Automatic Codex message redelivery now pauses after three failed turns; exhausted messages stay unacknowledged and remain readable with `cotal_inbox`.

— Named inbox command, no runnable example given. v0.165.0

15 **Windows tray-hidden auto-updates** Not checked yet IMPROVED 50

Windows desktop apps can now download and apply verified updates while hidden in the system tray or when Restart to update is chosen, restoring the pre-update window state after relaunch.

— Describes behavior; no user action required. v0.164.17

16 **Agent-profile credential refusal in console and watch**

Not checked yet IMPROVED 45

Agent-profile credentials are now refused in console and watch before announcing a feed, with instructions to mint a separate observer credential.

— Describes restriction and remedy, no named command. v0.166.0

17 **Browser contract modules require regeneration** Not checked yet 40

BREAKING

Generated browser contract modules must be regenerated this release because they now import the shared interpreter from json-schema-library 11.6.16.

— Names the dependency version but no regen command. v0.165.0

└──▷ **BREAKING ON UPGRADE**

! The `perm` answer subject, the `cotal_permissions_<space>` bucket and its readers, the `@cotal-ai/core/permissions` export, the `request` field on a presence condition, and `permissionRequests` on a Manager `status.connectors` row are removed. The Manager cluster advances to revision 51. A Claude Code connector no longer holds a permission prompt answerable from another session.

! Agent files that set `cwd` or `worktree` in frontmatter will now be refused instead of silently ignoring those fields; move the directory to the spawn request using `--cwd`.

! Generated browser contract modules must be regenerated with this release because they now import the shared interpreter from json-schema-library 11.6.16.

WAS THIS USEFUL?  

Vercel AI SDK

Feedback ⓘ ai@7.0.174 NOTES ↗ RANK

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

> Vercel AI SDK added structured output support to its batch generation APIs and introduced extraction model support for provider registries and custom providers.

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

01 **Structured output in batch generation APIs** Not checked yet NEW

70

`startBatch` now accepts a structured `output` parameter, and `getBatchResults` parses that output so callers get typed batch results instead of raw text.

– Names exact functions and behaviour, no code example given ai@7.0.174

THINNER COVERAGE BELOW

02 **Extraction model support in provider registries** Not checked yet

NEW

25

Provider registries and custom providers gained support for extraction models.

– Only mentioned in passing, no mechanism or usage detail ai@7.0.174

WAS THIS USEFUL? 👍 👎

Agno (formerly Phidata)

Feedback ⓘ v3.2.1 NOTES ↗

RANK

Agno is an agent framework and runtime that orchestrates and runs multi-agent systems.

> Agno's v3.2.1 release adds API visibility into session compaction and company task resources, alongside parameter expansions across several tool integrations.

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

01 Session compaction history endpoint

Not checked yet NEW

88

The new `GET /sessions/{session_id}/compactions` endpoint on `SqliteDb` and `PostgresDb` lists a session's compaction history, showing message counts replaced and token counts before and after, with results ordered newest first.

Inspect how many messages and tokens a session compacted away — useful for auditing context-window management on long-running sessions.

python

```
import httpx

response = httpx.get(
    'https://<your-agentos-
host>/sessions/<session_id>/compactions',
    headers={'Authorization': 'Bearer <token>'}
)
for compaction in response.json():
    print(compaction)
```

— Named endpoint with exact behaviour and a runnable example v3.2.1

02 Company task files and notes APIs

Not checked yet NEW

83

Read a task's files and shared notes over the API with `GET /company/tasks/{task_id}/files` and `GET /company/notes`, each supporting a `/{path}` sub-route to fetch a single file.

Fetch the file tree for a company task to let an agent inspect task-scoped attachments programmatically.

```python
import httpx

files = httpx.get(
    'https://<your-agentos-host>/company/tasks/<task_id>/files',
    headers={'Authorization': 'Bearer <token>'}
).json()

one_file = httpx.get(
    'https://<your-agentos-
host>/company/tasks/<task_id>/files/report.pdf',
    headers={'Authorization': 'Bearer <token>'}
).content
```

— Two named endpoints with routes and a working example v3.2.1

03 **Configurable request timeouts for LinearTools, DiscordTools, ModelsLabTools**

Not checked yet NEW 66

`LinearTools`, `DiscordTools`, and `ModelsLabTools` now accept a `timeout` parameter applied to every request, defaulting to 30 seconds (120 seconds for `ModelsLabTools`).

— Exact defaults named across three tools, no example given v3.2.1

04 **Default role for unlisted Slack members** Not checked yet NEW 65

The new `default_role` setting on the Slack integration gives members not explicitly listed in `users=` a managed role's permissions.

— Named config key but no usage example v3.2.1

THINNER COVERAGE BELOW

05 **New ExaTools search parameters** Not checked yet IMPROVED 58

`ExaTools` now accepts `include_text`, `exclude_text`, `user_location`, `moderation`, `highlights`, and `subpages` parameters.

— Parameters named but no mechanism or example shown v3.2.1

06 **Batch size for Qdrant async insert** Not checked yet IMPROVED 50

Qdrant 's `async_insert` now accepts a `batch_size` parameter, matching the behaviour of `insert`.

— Named parameter, minimal mechanism described v3.2.1

07 **Boolean return from MCPTools.connect()** Not checked yet IMPROVED

40

MCPTools.connect() now returns True when the connection succeeds and False when it does not.

— Clear behaviour change but no example or context v3.2.1

WAS THIS USEFUL? 👍 👎

Langflow

Feedback ⓘ CVE-2026-7646 CVE ↗

Langflow is a visual builder for agents and RAG workflows that exports each flow as an API or Python object for deployment.

CVE-2026-7646 : MCP Path Traversal Allows Full-Filesystem Read and JWT Secret Exfiltration Leading to Account Takeover (CVSS 9.6, affects < 1.9.1) Not checked yet

WAS THIS USEFUL? 👍 👎

lexmount Moli

Feedback ⓘ v1.1.26 NOTES ↗ RANK

Moli is a best headless browser for AI agents. Lite, Fast, High-Compatibility. Built in Rust

> Across two releases Moli added CDP session support over Chromium's remote debugging pipe for CLI-driven automation, native clipboard event payloads, and accessibility improvements, alongside canvas transfer, service worker fetch, and XHR interception changes.

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

01 **OffscreenCanvas.transferToImageBitmap support** Not checked yet

NEW

70

Implements `OffscreenCanvas.transferToImageBitmap` to transfer canvas bitmap data off the main thread.

— Names exact API and explains the off-thread behavior v1.1.25

THINNER COVERAGE BELOW

02 **CDP session over remote debugging pipe** Not checked yet NEW

55

Moli now serves a single CDP session over Chromium's remote debugging pipe, enabling CLI-driven browser automation.

— Names the pipe mechanism but no flags or example given v1.1.26

03 **Native FetchEvent construction in service workers** Not checked yet

NEW

50

Implements native `FetchEvent` construction in service workers.

— Names API and context but lacks usage detail v1.1.25

04 **Semantic link URLs via computed roles** Not checked yet IMPROVED

40

Shows semantic link URLs using computed roles in accessibility and fetch contexts.

— Names contexts but no concrete API or flag v1.1.26

05 **XHR interception during synchronous requests** Not checked yet

IMPROVED

40

Keeps XHR interception live during synchronous requests.

— Describes behavior change without a named surface v1.1.25

06 **Renderer output transport sender** Not checked yet NEW

35

Exposes a renderer output transport sender, per the commit `feat(core): expose renderer output transport sender`.

— Only a commit message describes this addition v1.1.25

07 **Native clipboard event payloads** Not checked yet NEW 25

Implements native clipboard event payloads for richer clipboard interaction support.

— Bare description, no mechanism or API named v1.1.26

WAS THIS USEFUL? 👍 👎

openexecutive

Feedback ⓘ v0.7.6 NOTES ↗ RANK

Openexecutive is an ai-powered virtual executive team — a single coherent executive persona backed by 8 specialist agents (FastAPI + Next.js).

> Openexecutive's orchestrator now gates all workflows through Approvals alone, while the Home screen gets a handover-focused redesign, new Schedule/People/Approvals pages, a Google Chat integration for announcements and live workflow updates, and browser delegation plus per-person memory for deeper context tracking.

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

01 Home screen redesign with handover briefing

Not checked yet

IMPROVED

75

Home now opens with a headline and a list of items to hand over, adds a briefing count for everything waiting on the user, renames the Work — Documents section to Made for you, and decides each Needs you card's words and buttons on the server. The Done for you rows now appear in the What it does view.

— Names exact section rename and server-side behaviour change v0.7.5 v0.7.4

THINNER COVERAGE BELOW

02 Google Chat integration for announcements and workflows

Not checked yet NEW

55

Adds a Google Chat integration that posts announcements and department messages to Google Chat spaces, lets the orchestrator know how people reach the Executive in Google Chat, and shows the live steps of a chat-started workflow as they happen in Chat.

— Names the integration and three behaviours, no setup steps v0.7.6

03 Approvals as sole orchestrator gate

Not checked yet BREAKING

50

The orchestrator now uses Approvals as the only gate, replacing previous multi-gate logic.

— Names the before/after but no migration detail given v0.7.4

04 Schedule, People, and Approvals pages

Not checked yet NEW

50

Adds a Schedule page that lists everything on its schedule in one view, a People page with Needs a look surfaced at the top, and an Approvals settings page that consolidates

all Approvals configuration in one place.

— Three distinct pages named, clear UI starting points v0.7.5

05 **Per-person chat memory and search** Not checked yet NEW

48

Keeps cards and decisions from a person's own chat in per-person memory and shows the speaker's recent chats in chat, so past context and past work are retained and included in search.

— Explains mechanism but no retrieval detail or limits v0.7.5

06 **MCP turn attribution to the requesting person** Not checked yet

IMPROVED

43

MCP now attributes an AI app's turn to the person themselves, and adds card answers and workflow runs to that turn's context.

— Names mechanism but no configuration surface given v0.7.5

07 **Edit card plan via reply** Not checked yet NEW

35

Allows changing a card's plan directly from a reply on the card.

— Clear action but no example of a reply or constraints v0.7.5

08 **Browser delegation using the user's session** Not checked yet NEW

28

Lets the Executive use the person's own browser for tasks that require their logged-in session.

— States capability only, no mechanism or scope v0.7.5

WAS THIS USEFUL?  

octop-harness

Feedback ⓘ v1.0.11 NOTES ↗ RANK

Octop-harness is a production-grade agent runtime engineered from the Harness Engineering philosophy.

> octop-harness's two releases added richer control over managed ACP turns — MCP server/session attachment with graceful fallback and a YOLO auto-approval mode — alongside per-agent context compaction settings and ACP runner forwarding via a new config field.

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

01 ACP runner forwarding via managed_runner config

Not checked yet

NEW

80

An agent's turns can be routed to a designated ACP runner instead of the model loop by setting `managed_runner` on that agent's `HarnessAgentConfig`, as in the example forwarding an `acp-forwarded-agent` to `<your-acp-runner-id>`.

Forward all turns from a specific agent to an ACP runner instead of running them through the model loop.

python

```
config = HarnessAgentConfig(
    name="acp-forwarded-agent",
    managed_runner="<your-acp-runner-id>",
    workspace_dir=root,
    backend={"type": "filesystem", "root_dir": str(root),
"virtual_mode": True},
)
```

— Names exact config key with a runnable example configuring it v1.0.10

02 Per-agent context compaction configuration

Not checked yet NEW

70

Each agent can now independently configure automatic context compaction — trigger, retention, and truncation — including disabling the auto-trigger entirely; the `/compact` command and context-overflow summarisation continue to work regardless of these settings.

— Names the `/compact` command and exact settings governed, but no config key shown

v1.0.10

THINNER COVERAGE BELOW

03 **MCP server and session mode attachment on ACP turns**

Not checked yet NEW 55

A managed ACP turn can now have MCP servers, session mode, and additional prompts attached; if the agent refuses the MCP set, the session continues without those tools instead of failing outright.

— Describes mechanism and fallback behaviour but no exact flag or field name v1.0.11

04 **YOLO auto-approval mode for managed ACP turns** Not checked yet

NEW 45

Enabling YOLO mode on a managed ACP turn automatically approves permission requests that are not hard-blocked, switching the session into the agent's declared no-review mode.

— Clear behaviour described, no concrete command or setting name given v1.0.11

WAS THIS USEFUL?  

atomic-agent

Feedback ⓘ v0.8.0 NOTES ↗ RANK

Atomic Agent is a local-first AI agent. Runs open-weight models on your own machine via llama.cpp.

> atomic-agent shipped browser-based MCP server authentication and a server catalog across both the CLI and desktop app, alongside offline/incognito privacy modes, six new chat platform integrations, a wider set of cloud and local model providers, and desktop-only additions like projects, multi-format chat export, and cross-platform voice.

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

01 MCP server authentication and catalog

Not checked yet NEW

95

Authenticate to MCP servers from the CLI with `/mcp auth <server>` and `/mcp signout <server>`, with per-MCP-server permission control to choose whether a server is asked every call, only on writes, or never. A catalog of ready-to-add MCP servers includes Exa, Firecrawl, Apify, Bright Data, Tavily, Brave Search, and Monid. The desktop app adds one-click sign-in to MCP servers in Settings › Connections › MCP servers, finishing authentication in the browser without copying keys, plus a searchable card catalog of ready MCP servers to browse when adding a new connection.

Authenticate to a Monid MCP server so the agent can access live X/Reddit data and leads without manually managing API keys.

```
$ /mcp auth monid
```

— Names exact CLI commands, UI path, and catalog providers. v0.8.0 desktop-v0.0.9

02 **Claude Code, Codex, and Cursor integration** Not checked yet NEW

75

Share agent memory and skills read-only with external tools such as Claude Code, Codex, and Cursor using the `atag mcp-server` command. The desktop app lets you install Claude Code, Codex, and Cursor plugins directly from inside the app.

Expose your agent's memory and skills read-only to Claude Code or Cursor so those tools can draw on the same context.

```
$ atag mcp-server
```

— Names the command and plugin path but not deeper mechanism. v0.8.0 desktop-v0.0.9

03 **New model providers and local model reuse** Not checked yet NEW

75

Connect new cloud and network model providers: Azure OpenAI, Amazon Bedrock, Google Vertex AI, W&B Inference, Lemonade, Foundry Local, and a network model server running on another computer. The desktop app also lets you use models already downloaded by LM Studio, Jan, or llama.cpp directly without a second download.

— Names every provider but no connection steps given. v0.8.0 desktop-v0.0.9

04 **Offline mode, privacy controls, and incognito sessions**

Not checked yet NEW

70

Enable offline mode so no data leaves the machine except to a local model: in the CLI press `o` in `/privacy`, or in the desktop app toggle it in Settings › Privacy, where you can now also edit your stored memory. Start incognito sessions with `/incognito` so nothing is saved and the chat disappears on exit.

— Names both CLI keypress and desktop settings path. Clear command but no further mechanism given.

v0.8.0 desktop-v0.0.9

05 **New chat commands and keybindings** Not checked yet NEW

70

New chat commands `/rewind`, `/btw`, `/why`, and `/resume` let you jump back to an earlier message, ask a side question, inspect the last step, or finish a cut-off reply; add files inline with `@`, open your editor from chat with `/edit`, and turn on vim keybindings with `/vim`.

— Four exact commands with stated purpose each. Three named shortcuts, minimal explanation of behaviour.

v0.8.0

06 **Chat import from Claude and other instances** Not checked yet NEW

70

Import chats from Claude app exports and other Atomic Agent instances via `atomic-agent import`, preserving original project or folder structure and skipping duplicates.

— Names command and describes preserved structure and dedup. v0.8.0

07 **Phone push notifications** Not checked yet NEW 65

Enable phone push notifications without a bot using `atomic-agent push on` from the CLI. The desktop app sets up notifications through the free ntfy app in three steps.

Enable phone push notifications so you receive alerts from the agent without setting up a bot account.

```
$ atomic-agent push on
```

— Names exact command and the notification mechanism. v0.8.0 desktop-v0.0.9

THINNER COVERAGE BELOW

08 **Global quick-entry shortcut** Not checked yet NEW 55

Open a quick-entry box from anywhere with ⌘⇧Space (Ctrl+Shift+Space on Windows and Linux).

— Exact keyboard shortcut given across platforms. desktop-v0.0.9

09 **Six new chat platform integrations** Not checked yet NEW

50

Adds six new chat integrations — Slack, Signal, Matrix, iMessage, Microsoft Teams, and Google Chat — each off by default.

— Names all six integrations but no setup detail. v0.8.0

10 **Projects for grouping chats** Not checked yet NEW 50

Group chats into projects with per-project instructions that apply to every chat in the folder.

— Describes the feature but no UI path given. desktop-v0.0.9

11 **Chat archiving and export formats** Not checked yet NEW

50

Archive and export chats to Markdown, PDF, Word, or HTML.

— Names four export formats, no further mechanism. desktop-v0.0.9

12 **Voice conversations on Windows and Linux** Not checked yet IMPROVED

Hold voice conversations with the agent; voice input now works on Windows and Linux in addition to macOS.

— States platform expansion, no setup detail. desktop-v0.0.9

WAS THIS USEFUL? 👍 👎

EXXETA exxperts

Feedback ⓘ v0.16.0 NOTES ↗ RANK

The exxperts framework runs local-first AI agents with approval-gated memory, model-provider integrations, MCP tools, and web search.

> exxperts v0.16.0 locks down room execution with OS-level sandbox protection and three command-permission modes, adds confirmation prompts for connector actions and GitHub Copilot enterprise sign-in, and merges the two room file-access modes into one.

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

01 Confirmation and allow-listing for connector actions

Not checked yet NEW 75

Connector actions that may change data now ask for confirmation and show what the room would send before running; choosing Allow for this room permits that specific action without future prompts, and allowed actions are visible and manageable in Room settings, Connectors.

Allow a specific connector action to run without asking each time, including in scheduled tasks.

📍 When the connector-action confirmation prompt appears, choose Allow for this room. To review or remove allowed actions later, go to Room settings, Connectors.

— Exact UI path and allow-list mechanism with example v0.16.0

02 OS-level sandbox protection for room commands Not checked yet

NEW 70

Room commands can now be sandboxed at the OS level on macOS and Windows 11 (24H2+), keeping them inside the room's folders; protection can be turned off per-room with the Protection switch in Room settings, Folder.

Turn off sandbox protection for one room whose commands need access to files outside its folder, such as a git push over SSH.

📍 In Room settings, go to Folder, then flip the Protection switch off for that room.

— Names platforms and exact toggle location v0.16.0

03 Command-permission modes and confirmation panel Not checked yet

NEW 70

Rooms gain a command-permission mode chosen via the chip under the message box — Ask, Semi, or Auto: Ask confirms each command and new website, Semi runs commands silently but asks before new websites, and Auto reaches public websites without asking. A confirmation panel above the message box lets users approve running a command, switching folders, reaching a new website, or turning on commands.

— Modes and panel described but no runnable example `v0.16.0`

---

04 **Unified room file-access mode** `Not checked yet` **BREAKING**

70

Room file-access modes Full access and Bounded are now merged into one; Room settings, Folder lists allowed folders and permitted sites under `Also allowed`. After upgrading, a room previously set to Bounded has its commands turned off and must be re-enabled in Room settings, Folder.

— Names setting, breaking effect, and fix path `v0.16.0`

---

05 **Desktop screen history navigation shortcuts** `Not checked yet` **NEW**

65

Users can navigate back and forward through previously opened screens in the Mac and Windows desktop apps using `⌘[` / `⌘]` on Mac or `Alt+Left` / `Alt+Right` on Windows.

— Exact runnable keyboard shortcuts named `v0.16.0`

---

06 **GitHub Copilot enterprise sign-in** `Not checked yet` **NEW**

60

Users can sign in to GitHub Copilot using a company GitHub Enterprise instance (entering an address like `yourcompany.ghe.com` in Settings, AI setup) to access only the models that account can use.

Sign in to GitHub Copilot using a company GitHub Enterprise instance to access only the models that account can use.

> 📍 In Settings, AI setup, open the GitHub Copilot sign-in and enter your company's GitHub address (e.g. `yourcompany.ghe.com`) instead of leaving it blank.

— Exact settings path and example domain given `v0.16.0`

THINNER COVERAGE BELOW

---

07 **Chat-selected model becomes room's Conversation model**

`Not checked yet` **IMPROVED**

50

A model chosen in chat now also becomes the room's Conversation model shown in Room settings, so new conversations and scheduled tasks start on it.

— States behaviour but only names settings page v0.16.0

08 **Room mention via @ in message box** Not checked yet NEW

45

Users can pick a room to consult after typing `@` and a space, or add it as a label on the message box.

— Clear trigger syntax but thin mechanism v0.16.0

09 **Claude Fable 5 removed from sign-in model list** Not checked yet

45

BREAKING

Claude Fable 5 is removed from the Claude sign-in model list; a room still set to it must choose another model before starting a new conversation or scheduled task.

— Names removed model and required action v0.16.0

└──▷ **BREAKING ON UPGRADE**

! A room previously set to Bounded has its commands turned off after upgrading; re-enable them in Room settings, Folder.

! A room on a Claude sign-in still set to Fable 5 must choose another model before starting a new conversation or scheduled task (Claude Fable 5 is removed from the sign-in model list).

WAS THIS USEFUL? 👍 👎

CopilotKit OpenBot

Feedback ⓘ v1.1.0 NOTES ↗ RANK

OpenBot runs self-hosted AI coworkers that each get their own browser, files, and gated tools, with every action decided and recorded.

> OpenBot's v1.1.0 release introduces multi-organization support with per-team management, moves its entire API onto a documented, versioned `/api/v2/` surface, and locks down access with install-token auth and PostgreSQL row-level-security roles, alongside a string of breaking deployment and packaging changes.

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

01 Install token authentication for personal edition

Not checked yet

86

BREAKING

Personal-edition installs now require all API requests to carry an install token via the `X-OpenBot-Install-Token` header or an API key from Settings > API keys; a single-user container won't start without `OPENBOT_INSTALL_TOKEN` set.

Authenticate a script against a personal-edition OpenBot that now requires a token on every request.

```
$ curl -H 'X-OpenBot-Install-Token: <your-install-token>'
https://<openbot-host>/api/v2/organizations/default/bots
```

— Runnable curl command demonstrates the exact header v1.1.0

02 Unified REST and WebSocket API v2

Not checked yet BREAKING

85

OpenBot's REST API is now documented in `api/openapi.yaml` (OpenAPI 3.1) and its WebSocket channels in `api/asyncapi.yaml` (AsyncAPI 3.1), with every resource under `/api/v2/organizations/{organizationId}/`. All earlier `/api/...` addresses are retired, so clients must move to `/api/v2/` and reverse proxies must allow `/api/v2/` with WebSocket upgrades; the API now also returns 400 for a query parameter sent more than once where the spec takes one value, for a body that isn't valid JSON on operations where the body is optional, and for an empty `days=` on check-in usage (previously read as 7).

— Endpoint pattern and spec files named with exact failure conditions v1.1.0

03 **Database schema migrations for goals, routines and MCP servers**

Not checked yet BREAKING 85

Migration `0089_goal_ends_when_met` adds an `ends_when_met` field to `responsibilities` (defaults false for existing goals); migration `0090_routine_goal_linked_by_run` adds `goal_linked_by_run` to `routines`; and migration `0091_remote_mcp_oauth` adds `oauth` to `mcp_servers`. These migrations are irreversible at runtime (restore the pre-upgrade backup to roll back), and existing routines linked to a goal via a 1.0.x firing now have `goal_linked_by_run` set false, so completing that goal switches the routine off and deleting the goal deletes it.

— Every migration ID and field named with exact rollback consequence v1.1.0

---

04 **SMTP-based invitation emails** Not checked yet NEW 81

Invitation emails are sent when `SMTP_HOST` and `MAIL_FROM` are configured; `SMTP_PORT` defaults to 587, or 465 when `SMTP_SECURE=true`. With `SMTP_HOST` set, the server won't start without `MAIL_FROM`, with an invalid `SMTP_PORT` or `SMTP_SECURE`, or with only one of `SMTP_USER` and `SMTP_PASSWORD`.

Require SMTP-based invitation emails by supplying all required mail settings so the server will start and send invitations.

env

```
SMTP_HOST=smtp.example.com
SMTP_PORT=587
SMTP_USER=alerts@example.com
SMTP_PASSWORD=s3cr3t
MAIL_FROM=alerts@example.com
```

— Runnable env config with all required variables and defaults v1.1.0

05 **Deployment and packaging changes** Not checked yet BREAKING

80

The `openbot` Kubernetes chart no longer installs Intelligence, and a release on the 0.2 chart that runs the chart-installed Intelligence is refused at upgrade. Compose removes the `spire-init`, `spire-server`, `spire-agent`, `agent-bot` and TypeScript `agent-langgraph` services and the `BOT_PORT`, `BOT_IMAGE`, `AGENT_BOT_MODEL` and `WORKER_SHARED_SECRET` settings; the `migrate` and `worker` services are no longer started. `bash scripts/start.sh` no longer runs the server from the checkout; it pulls the published images instead. OpenBot images are now signed with the key in `docs/cosign.pub` and no longer carry a GitHub build attestation or a `sha-<commit>` tag; deploy using a version tag or the digest from `container-images.json`.

— Names every removed service and setting, no migration script given v1.1.0

06 **PostgreSQL row-level security roles** Not checked yet BREAKING

76

PostgreSQL now enforces row-level security using two new non-login roles, `openbot_app` and `openbot_system`, to keep organizations isolated; a connection pooler must pass on the `role` start-up parameter, and login must be allowed to create roles or an administrator must create the two roles before upgrade.

— Role names and pooler requirement given, no setup command shown v1.1.0

07 **API error format, pagination and idempotency** Not checked yet

IMPROVED

73

Errors from the API now return problem details (`application/problem+json`) with a stable `code`, and all list endpoints page using `cursor` and `limit`; every write operation accepts an Idempotency-Key header.

— Exact header and field names given, no example call shown v1.1.0

08 **Multi-organization support** Not checked yet NEW 72

Setting `OPENBOT_MULTIPLE_ORGANIZATIONS=true` on the server enables multi-organization mode; each user then gets their own organization on account creation, can create additional organizations, and joins others by invitation.

Turn on multi-organization mode so each new user automatically gets their own organization and can create or join others.

```env
OPENBOT_MULTIPLE_ORGANIZATIONS=true
```

— Config flag and behaviour given, invitation flow left vague v1.1.0

09 **Team management within organizations** Not checked yet NEW 70

Owners and admins can create, nest, and delete teams within an organization, control Bot publishing and pinning per team, assign apps and skills to teams, and set permissions and network rules per team; SCIM directory groups also now surface as teams inside OpenBot, kept in sync with the company directory and marked as directory-managed.

— Capabilities described but no exact UI navigation path v1.1.0

THINNER COVERAGE BELOW

10 **Per-organization provider usage counting** Not checked yet IMPROVED 45

Admin > Providers > Usage now counts every run that reached a model — including background turns — per organization.

— UI path given but counting mechanism unexplained v1.1.0

11 **Model allowlist for every run type** Not checked yet IMPROVED 45

The model allowlist now applies to every run type: built-in Bots, harness Bots, agent hand-offs, routines, check-ins, Slack/Teams/SMS messages, and group conversations.

— Lists covered run types, no config surface named v1.1.0

12 **Theme reset to System default** Not checked yet BREAKING

35

Every user's theme is reset to System; anyone who had explicitly chosen light or dark must select their preference again.

— Minimal description of a UI-only reset v1.1.0

└──▷ **BREAKING ON UPGRADE**

! All earlier `/api/...` addresses are retired; every client must move to `/api/v2/` as described in `api/openapi.yaml` or `api/asyncapi.yaml`, and reverse proxies must allow `/api/v2/` with WebSocket upgrades.

! The API now returns 400 for a query parameter sent more than once where the spec takes one value, for a body that isn't valid JSON on operations where the body is optional, and for an empty `days=` on check-in usage (previously read as 7).

! Personal-edition OpenBot refuses requests with no credentials; scripts must supply an API key or the `X-OpenBot-Install-Token` header, and a single-user container won't start without `OPENBOT_INSTALL_TOKEN`.

! PostgreSQL login must be allowed to create roles, or the two roles `openbot_app` and `openbot_system` must be created by an administrator before upgrade; a connection pooler must pass on the `role` start-up parameter.

! Database migrations from `0088_approval_reviewer_guidance` onward are irreversible at runtime; restore the pre-upgrade backup to roll back.

! Existing routines that were linked to a goal via a 1.0.x firing now have `goal_linked_by_run` set false, so completing that goal will switch the routine off and deleting the goal will delete it.

! The `openbot` Kubernetes chart no longer installs Intelligence; a release on the 0.2 chart that runs the chart-installed Intelligence is refused at upgrade.

! Compose removes the `spire-init`, `spire-server`, `spire-agent`, `agent-bot` and TypeScript `agent-langgraph` services and the `BOT_PORT`, `BOT_IMAGE`, `AGENT_BOT_MODEL` and `WORKER_SHARED_SECRET` settings; the `migrate` and `worker` services are no longer started.

! `bash scripts/start.sh` no longer runs the server from the checkout; it pulls the published images instead.

! With `SMTP_HOST` set, the server won't start without `MAIL_FROM`, with an invalid `SMTP_PORT` or `SMTP_SECURE`, or with only one of `SMTP_USER` and `SMTP_PASSWORD`.

! OpenBot images are now signed with the key in `docs/cosign.pub` and no longer carry a GitHub build attestation or a `sha-<commit>` tag; deploy using a version tag or the digest from `container-images.json`.

! Every user's theme is reset to System; anyone who had explicitly chosen light or dark must select their preference again.

WAS THIS USEFUL?  

agentlas-os

Feedback ⓘ os-v1.2.99 NOTES ↗ RANK

Agentlas-os is the Open-Source AI Agent OS inspired by OpenAI Dots, Meta Muse, OpenClaw, Hermes & Mem0. Autonomous 3-tier headhunting, Receipt & Verifier loops, scoped memory, and BYOM for Claude Code, Codex & Cursor.

> Agentlas OS reworked its memory system with scoped time boundaries, a redefined fact/preference split, a new four-scope memory model with a unified `memory.record` write API, and a `hephaestus register-agent` command for local agent registration, alongside storage reliability fixes and a breaking memory-store schema change.

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

01 Local agent registration via hephaestus register-agent

Not checked yet NEW 95

`hephaestus register-agent <ref>` registers a Hub or Cloud agent locally, provisioning a folder under `~/.agentlas/agents` and reporting availability status (`not_registered` , `no_folder` , etc.); the base path can be overridden with the `AGENTLAS_AGENT_HOME` environment variable.

Register a remote Hub agent on the local machine so the Hephaestus engine can recruit and invoke it in task graphs.

```
$ hephaestus register-agent agentlas-hub://my-org/data-analyst-agent
```

Override the default agent folder root when your home directory is on a non-standard path or a shared volume.

```
$ export AGENTLAS_AGENT_HOME=/mnt/shared/agents
hephaestus register-agent agentlas-hub://my-org/finance-agent
```

— Exact command, env var, and runnable examples given os-v1.2.98

02 Time-bounded memory records

Not checked yet NEW 80

A `boundary` object with `type` , `from` , and `to` fields can now be passed to `memory.record` or the memory ticket schema to scope a person's memory to a specific time window; the engine accepts the field without storing it, while the Agentlas desktop app keeps it.

Record a memory fact with a time boundary — for example, to scope a person's availability preference to a specific window without permanently storing it.

```python
memory.record({
  "content": "Prefers no meetings before 10 AM",
  "kind": "fact",
  "boundary": {
    "type": "time",
    "from": "2025-01-01",
    "to": "2025-06-30"
  }
})
```

— Names exact fields and includes a runnable example `os-v1.2.99`

03 **Four-scope memory model and memory.record write API**

Not checked yet NEW 72

Agentlas OS now structures memory across four scopes — personal, agent, project, and session — with a single recall rule applied across all scopes the owner has access to. A single `memory.record` call can write, rewrite, condense, or remove memory lines, with the model managing memory edits directly rather than storing inferences as facts.

— Explains scopes and API mechanism but no example call shown `os-v1.2.98`

THINNER COVERAGE BELOW

04 **Memory store schema version 2 incompatibility** Not checked yet

58

BREAKING

After the Agentlas desktop app converts a memory store to schema version 2, engines 1.2.96 and 1.2.97 cannot open that store and will run without memory until updated; update the engine before opening the 2.0 desktop app.

— Clear migration warning but no fixed engine version named `os-v1.2.98`

05 **Redefined fact vs preference memory categories** Not checked yet

IMPROVED 54

Memory rules 1.17.0 now classifies what a person says about their own life, tastes, and wishes as a `fact` available to every agent, while `preference` is scoped to how one specific agent should respond to them; lines already stored are not relabelled.

— Clear before/after but no command to act on it `os-v1.2.99`

06 **Memory storage reliability fixes** Not checked yet IMPROVED

51

A newly created project marker now carries a unique id, so two projects sharing the same folder name no longer share a key, letting a project's code index live alongside its memory. On Windows, memory reads and writes now work for paths longer than 259 characters by opening the memory tree with the extended-path prefix.

— Two concrete fixes but neither exposes a command or key `os-v1.2.98`

└──▷ **BREAKING ON UPGRADE**

! After the Agentlas desktop app converts a memory store to store schema version 2, engines 1.2.96 and 1.2.97 cannot open that store and will run without memory until updated — update the engine before opening the 2.0 desktop app.

WAS THIS USEFUL? 👍 👎

AutoGPT

Feedback ⓘ CVE-2026-129984 CVE ↗ STABLE RANK

AutoGPT automates multi-step workflows with continuous AI agents that users can build and deploy.

`CVE-2026-129984`: Billing-manager-only members are treated as ordinary members and can read other team members' runs (CVSS 4.3, affects `>= 0.6.67, < 0.11.3`)

Not checked yet

▾ **1 more release in this issue** · 2026-11-16

> CVE-2026-129317 CVE ↗ STABLE
>
> `CVE-2026-129317`: Anyone with an active OAuth app's client ID can stall the REST API by sending wrong client secrets (CVSS 5.9, affects `>= 0.6.41, < 0.11.2`)
>
> Not checked yet

WAS THIS USEFUL? 👍 👎

Graphicle AI Graphicle

Feedback ⓘ v0.1.18 NOTES ↗

RANK

Graphicle is a code graph and memory for your coding agents, in one Rust program. For Claude Code, Codex, Cursor and any MCP client.

> Graphicle's v0.1.18 release adds OpenCode as a supported MCP client, introduces a profiling environment variable for update timing, and surfaces a new marker for incomplete dependency invalidation on very large repositories.

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

01 **OpenCode integration via connect and install commands**

Not checked yet NEW 95

Run `graphicle connect opencode` to add the project's MCP server to OpenCode's global `opencode.json` or `opencode.jsonc` (comments preserved); `graphicle disconnect opencode` removes it, handling both the OpenCode 1.x (`mcp`) and OpenCode 2.x (`mcp.servers`) config shapes. The graph engine's OpenCode skill and plugin can also be installed separately with `graphicle graph install --platform opencode` .

– Exact commands and config file paths given, fully actionable. v0.1.18

02 **Incomplete dependency invalidation marker** Not checked yet NEW

75

Graphicle now surfaces incomplete dependency invalidation via an `invalidation_incomplete` marker on sources that exceed the 100,000-artifact walk limit, counted by `freshness` and each repository's freshness count until the walk completes or the source is removed.

– Names the marker and limit but no remediation steps given. v0.1.18

03 **GRAPHICLE_PROFILE env var for update timing** Not checked yet NEW

65

Setting `GRAPHICLE_PROFILE=1` enables timing of an update's graph-read, manifest load, and changed-file split phases.

– Names the exact env var but lacks output detail. v0.1.18

WAS THIS USEFUL? 👍 👎

Mads Lorentzen AI Job Search

Feedback ⓘ v2.0.0

NOTES ↗ RANK

AI Job Search is the job search that runs on your machine. AI job application framework built on Claude Code: evaluate postings, tailor CVs, write cover letters, prep interviews.

> AI Job Search v2.0.0 adds fork-local overrides for security guards and employment-type tracking on scraped postings, while moving portal job-board skills out of the core repository into an external community package.

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

01 **Portal job-board skills moved to community-portals repo**

Not checked yet

77

BREAKING

Portal job-board skills have been moved out of the main repository to the external `community-portals` repository; existing setups that relied on bundled portal skills will lose `/scrape` portal support until the relevant skill folders are manually copied from `community-portals` into `.agents/skills/`.

– Gives exact migration path and affected command v2.0.0

02 **Fork-local security guard negations** Not checked yet NEW

75

Forks can override built-in security guards by adding negation rules to `.github/security-guards-local.txt`, keeping customisations out of upstream files so they survive updates.

– Names exact config file path and mechanism v2.0.0

THINNER COVERAGE BELOW

03 **Employment type field on scraped postings** Not checked yet

IMPROVED

35

Scrape and rank postings now record and surface each listing's employment type alongside existing posting fields.

– Describes field addition but no schema or example shown v2.0.0

└──▷ **BREAKING ON UPGRADE**

! Portal job-board skills have been moved out of the main repository to the external `community-portals` repository; existing setups that relied on bundled portal skills will lose `/scrape` portal support until the relevant skill folders are manually copied from `community-portals` into `.agents/skills/`.

WAS THIS USEFUL?  

ag2

Feedback ⓘ v1.3.2 NOTES ↗ RANK

AG2 (formerly AutoGen): The Open-Source AgentOS.Join us at: https://discord.gg/sNGSwQME3x

> ag2 v1.3.2 adds batched multi-question decision calls, overhauls token usage accounting across providers, upgrades to ACP 1.0, and adds Python 3.15 support.

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

01 **Multi-question decision requests via response_schema**

Not checked yet NEW 93

A single ask() call can now gather several answers at once by setting `response_schema` to a dataclass, `TypedDict`, or Pydantic model whose fields carry `Question` annotations; the shared input is sent and billed once. Works with `OpenAIDecisionsConfig` and `TypeSafeConfig`.

Triage a support ticket with multiple questions in one model call, reducing cost and latency when inputs are shared.

```python
from dataclasses import dataclass
from typing import Annotated
from ag2 import Agent
from ag2.config import OpenAIDecisionsConfig
from ag2.decisions import Question

@dataclass
class TicketTriage:
    is_spam: Annotated[bool, Question("Is this message spam?")]
    department: Annotated[Department, Question("Which team should
handle this ticket?")]
    urgency: Annotated[Urgency, Question("How urgent is this
ticket?")]

triage = Agent(
    "triage",
    prompt="You triage support tickets.",
    config=OpenAIDecisionsConfig(),
    response_schema=TicketTriage,
)

reply = await triage.ask("I was charged twice this month.")
result = await reply.content()  # TicketTriage(is_spam=False,
...)
confidence = reply.response.metadata["department"]["confidence"]
```

— Full mechanism plus a runnable code example `v1.3.2`

02 **Normalised token usage accounting across providers**

`Not checked yet` 81

`BREAKING`

`Usage.prompt_tokens` now always covers the full input including cache reads and writes, with `cache_read_input_tokens` and `cache_creation_input_tokens` reported as parts of that total, and `total_tokens >= prompt_tokens + completion_tokens`. This is a breaking change: on Anthropic and Bedrock, `prompt_tokens`, `total_tokens`, telemetry `gen_ai.usage.input_tokens`, and the input metric now grow by previously excluded cached tokens, and on Gemini Live thinking models `completion_tokens` and `total_tokens` now include thinking tokens — so token budgets and `TokenMonitor` thresholds are reached sooner on these workloads.

— Named fields and before/after but no migration command `v1.3.2`

03 **ACP 1.0 support in the acp extra** `Not checked yet`

`BREAKING`

76

The `acp` extra now targets ACP 1.0, requiring `agent-client-protocol>=1.0.0,<2` (0.12.x no longer supported). Each message chunk is now delivered as its own `ModelMessage` event, `turn_messages` replaces `turn_text` on `BridgeState`, and `PermissionPolicy` is exported from `ag2.acp`.

— Names exact dependency version and renamed fields v1.3.2

THINNER COVERAGE BELOW

04 **Python 3.15 compatibility** Not checked yet NEW

25

ag2 now supports and type-checks on Python 3.15.

— Bare compatibility statement with no further detail v1.3.2

└──▷ **BREAKING ON UPGRADE**

! The `acp` extra now requires `agent-client-protocol>=1.0.0,<2`; version 0.12.x is no longer supported — update any pinned `agent-client-protocol` dependency before upgrading.

! On Anthropic and Bedrock, `prompt_tokens`, `total_tokens`, telemetry `gen_ai.usage.input_tokens`, and the input metric now grow by the previously excluded cached tokens; token budgets and `TokenMonitor` thresholds are reached sooner on cached workloads.

! On Gemini Live thinking models, `completion_tokens` and `total_tokens` now include thinking tokens; token budgets and `TokenMonitor` thresholds are reached sooner on those workloads.

! `BridgeState.turn_text` is replaced by `turn_messages` in the ACP client.

WAS THIS USEFUL? 👍 👎

◆ MCP Servers & Tooling

godot-mcp

Feedback ⓘ v0.40.0 NOTES ↗ RANK

Godot-MCP — Model Context Protocol (MCP) integration for the Godot Engine. dev.

> Godot-MCP shipped a new command-line interface covering project scaffolding, plugin install, agent setup, headless CI checks and RQL scene queries, and added official .NET 10 SDK support alongside .NET 8.

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

01 **New godot-cli command-line interface** Not checked yet NEW

90

A new CLI adds commands for the full workflow: `create-project` to scaffold a project (e.g. `--dotnet --name MyGame`), `install-plugin` to install the Godot plugin, `login` to sign in, `setup-mcp` to configure an AI agent (e.g. `claude-code`), `rql` for headless Relational Query Language queries against the scene tree (e.g. `nodes | where name starts-with "Tree" | count`), `check` to run a CI gate (import, build, editor boot, GDScript load check), and `export` to produce a release build.

Bootstrap a new project, install the plugin, sign in, and configure an AI agent in one go.

```
$ godot-cli create-project ./MyGodotProject --dotnet --name
MyGame
godot-cli install-plugin ./MyGodotProject
godot-cli login
godot-cli setup-mcp claude-code ./MyGodotProject
```

Query the scene tree headlessly in CI — counts nodes whose name starts with `Tree` without an open editor.

```
$ godot-cli rql 'nodes | where name starts-with "Tree" | count'
```

Run the full CI gate on a project: import, build, editor boot, and GDScript load check, then export the release preset.

```
$ godot-cli check ./MyGodotProject
godot-cli export ./MyGodotProject
```

— Names every CLI subcommand with runnable examples product docs

THINNER COVERAGE BELOW

02 **.NET 10 SDK support** Not checked yet NEW

50

godot-mcp now officially supports the .NET 10 SDK alongside the existing .NET 8 requirement, with live extension tests running on .NET 10 and documentation recommending .NET 10 as the preferred runtime.

— States what changed but no migration steps v0.40.0

WAS THIS USEFUL? 👍 👎

whodb

Feedback ⓘ 0.139.0 NOTES ↗ RANK

Whodb is a where data access meets operational intelligence

> WhoDB's latest release adds CLI sync support, exposes computed source fields, and improves MCP support, though details remain sparse.

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

01 **CLI sync support** Not checked yet NEW 20

WhoDB adds CLI sync support, per the release summary.

– Named only in summary, no mechanism or command given 0.139.0

02 **Computed source fields exposed** Not checked yet NEW 20

WhoDB now exposes computed source fields, per the release summary.

– No explanation of which fields or where they appear 0.139.0

03 **Improved MCP support** Not checked yet IMPROVED 15

WhoDB improves MCP support.

– Bare statement with no mechanism or scope detail 0.139.0

WAS THIS USEFUL? 👍 👎

gograph

Feedback ⓘ v1.7.21 NOTES ↗ RANK

Gograph is a local-only Go static analysis engine with a built-in MCP server. Gives AI coding agents deterministic structural awareness: call graphs, impact analysis, symbol search, and more.

> gograph adds a new `uncovered` command that uses real `go test -coverprofile` data to find untested changed lines, architecture-enforcement policy predicates for `validate`, and configurable taint-flow sources/sinks, alongside a `whatsnew` release-notes command, `tests` and `complexity` filtering flags, and a dropped Windows build.

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

01 **Custom taint-flow sources, sinks and sanitizers** Not checked yet

NEW 96

Repository-specific taint sources and sinks can be defined in `gograph/flow.json` with `"sources"`, `"sinks"`, and `"sanitizers"` entries, each carrying `"kind"`, `"return"` or `"argument"`, and optional `"severity"`. Custom `kind` values are usable via new `--source http_request|decoded_json|environment|<custom>` and `--sink sql_query|process_execution|filesystem|outbound_http|<custom>` flags on the `flow` command to scope taint-flow analysis, plus a new `--parameter-only` mode and a budget limit.

Enforce that only your auth package is allowed to call os/exec by adding a callers_only_from policy, and define a custom MCP-response sink with high severity in gograph/flow.json.

```json
{
  "sinks": [
    {
      "symbol": "mcpio.WriteMCPResponse",
      "kind": "mcp_response",
      "argument": 1,
      "severity": "high"
    }
  ],
  "sanitizers": [
    {
      "symbol": "mcpio.Validate",
      "sinks": ["mcp_response"],
      "argument": 0
    }
  ]
}
```

— Config file, fields and flags all named with a runnable example v1.7.21

02 **`gograph uncovered` command for coverage gaps** Not checked yet

NEW

95

New `gograph uncovered` command lists changed lines of non-test Go files that a `go test -coverprofile` run never executed, reporting exact line ranges per file and returning `gograph.uncovered.v1`. It switches from static callers to runtime evidence and supports `--uncommitted`, `--since <ref>`, `--coverprofile <file>`, and `--exclude-interface-methods` to skip methods implementing common interfaces such as `Error`, `String`, or `MarshalJSON`, reporting a count of skipped methods.

Find changed lines not exercised by your test suite so you can target new tests precisely.

```
$ gograph uncovered --uncommitted --coverprofile coverage.out
```

— Full flag set, output format and mechanism named product docs

03 **Architecture policy predicates for validate** Not checked yet NEW

90

New policy predicates `imports_only_from` (only allowlisted packages or files may import a package), `callers_only_from` (only allowlisted packages or files may call a function), and `call_arg_constant` (checks that argument N is always a compile-time constant, or never one of `forbidden_values`), each reporting every violating file and

line on failure; added as rules to the `validate` command for architecture enforcement.

Enforce that only your auth package is allowed to call os/exec by adding a callers_only_from policy, and define a custom MCP-response sink with high severity in gograph/flow.json.

```json
{
  "sinks": [
    {
      "symbol": "mcpio.WriteMCPResponse",
      "kind": "mcp_response",
      "argument": 1,
      "severity": "high"
    }
  ],
  "sanitizers": [
    {
      "symbol": "mcpio.Validate",
      "sinks": ["mcp_response"],
      "argument": 0
    }
  ]
}
```

— Names all three predicates with their exact behaviour v1.7.21

04 **`gograph tests` command output enhancements** Not checked yet

IMPROVED

83

The `tests` command gains `--format run|go-test`, printing an anchored, regex-escaped `-run` pattern per import path (`run`) or a ready-to-run `go test` command per package (`go-test`). Separately, `tests --transitive` output now shows the summary first and nearest results first, with bounded pages.

Identify which tests cover the functions you changed before merging, and get ready-to-paste go test commands for CI.

```
$ gograph tests --uncommitted --format go-test
```

— Exact flag values and a runnable example given v1.7.21

05 **Complexity command filtering flags** Not checked yet NEW

75

The `complexity` command gains `--no-tests`, `--top`, and `--path` filters to narrow cyclomatic-complexity output by test inclusion, ranking, and path.

Rank the 10 most complex non-test functions under a specific package path to focus refactoring effort.

```
$ gograph complexity --no-tests --top 10 --path internal/auth
```

— Exact flags plus a runnable command example v1.7.21

THINNER COVERAGE BELOW

06 **`whatsnew` and `version` info commands** Not checked yet NEW

50

New `gograph whatsnew` command displays release notes for the running binary. The `version` command now also reports the module version for `go install` builds.

Check the release notes bundled with the installed binary without visiting the web.

```
$ gograph whatsnew
```

— Two named commands with a runnable example for one v1.7.21

07 **Lower severity for constant-literal process execution**

Not checked yet IMPROVED

45

Process-execution findings are now marked `parameter_only` at most medium severity when the program is a constant literal that is not a shell or interpreter.

— States the rule but no example or scope detail product docs

08 **Committable policy file location** Not checked yet IMPROVED

45

Policy files are now read from `gograph/` (committable) first, then `.gograph/`, giving teams a checked-in policy location.

— Named paths but no further mechanism given v1.7.21

09 **Windows builds discontinued** Not checked yet BREAKING

35

Windows binaries are no longer built or published; macOS and Linux targets are unchanged.

— Clear statement but no migration guidance v1.7.21

└──▷ BREAKING ON UPGRADE

! Windows binaries are no longer built or published; macOS and Linux targets are unchanged.

WAS THIS USEFUL? 👍 👎

context7

Feedback ⓘ `@upstash/context7-mcp@4.7.3` NOTES ↗ RANK

Context7 maintains up-to-date documentation and context for AI agents through an MCP server.

> Context7 replaced its experimental `query-code` tool with four dedicated code-search tools, added per-user control over which experimental tools an MCP server exposes, simplified on-premise deployment configuration with a new environment variable, and gave the CLI a command for fetching library migration guides.

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

01 `ctx7 get-upgrade-guide` command

Not checked yet NEW

95

Fetch a migration guide between two library versions from the CLI with `ctx7 get-upgrade-guide <libraryId> <fromVersion> [toVersion]`, using `--page` to paginate long guides or `--json` for machine-readable output; exits with code 75 while the guide is still being prepared.

Pull a migration guide for a specific library version bump — paginate a long guide or pipe JSON output into another tool.

```
$ ctx7 get-upgrade-guide react 17.0.0 18.0.0 --json
```

— Full command syntax, flags, and exit code with a runnable example. ctx7@0.12.0

02 CONTEXT7_BASE_URL environment variable for on-premise deployments

Not checked yet

91

BREAKING

Set your on-premise deployment URL once with the `CONTEXT7_BASE_URL` environment variable instead of passing `--base-url` on every command; `--base-url` still takes precedence when both are set. Authenticate against a custom deployment using `CONTEXT7_API_KEY` only — a custom URL from either `CONTEXT7_BASE_URL` or `--base-url` disables the update check, telemetry, and `ctx7 login` session handling for that host, so on-premise users must set `CONTEXT7_API_KEY` to avoid auth failures.

Avoid repeating `--base-url` on every command for an on-premise context7 deployment by exporting the URL once in your shell environment.

```
$ export CONTEXT7_BASE_URL=https://context7.internal.example.com
export CONTEXT7_API_KEY=<your-api-key>
ctx7 search 'kubernetes ingress'
```

Persist the on-premise deployment URL and API key for all sessions by adding them to your shell profile or `.env` file.

```text
CONTEXT7_BASE_URL=https://context7.internal.example.com
CONTEXT7_API_KEY=<your-api-key>
```

— Exact env vars, precedence rule, and auth fallback with runnable examples. `ctx7@0.12.1`

03 **Per-user experimental tool selection and flags** Not checked yet

NEW

80

Pick which experimental tools your MCP server lists by going to Settings > MCP Tools in the Context7 dashboard; selecting All experimental tools also includes tools added in future releases. Enable all experimental tools with `--experimental`, `CONTEXT7_EXPERIMENTAL=1`, or `?experimental=1`; the older `--research-preview`, `CONTEXT7_RESEARCH_PREVIEW=1`, and `?research_preview=1` flags still work.

Start the MCP server with every experimental tool enabled — useful in a CI or local dev environment where you want access to all preview features without touching the dashboard.

```
$ CONTEXT7_EXPERIMENTAL=1 npx @upstash/context7-mcp
```

— Names dashboard path, flags and env vars with a runnable example.

`@upstash/context7-mcp@4.7.1`

04 **Code search tools replace query-code** Not checked yet

BREAKING

77

context7 replaces the experimental `query-code` tool with four new tools — `grep-code`, `glob-code`, `read-code`, and `find-symbol` — which accept the same parameters as the agents' native Grep, Glob, and Read tools; a dashboard choice that previously enabled `query-code` now enables these four instead. Before being replaced, `query-code` had been updated to accept `pattern` as a regex matched within one line (up to 1000 characters, with a one-word `query` or `fixed_strings`

phrase still supported), and both `ctx7 code` and `query-code` told callers to retry for up to 3 minutes while a library was being indexed for the first time.

Search for a regex pattern inside a library's source code with `query-code` — use `pattern` instead of a multi-word `query` to get real matches.

```
$ query-code --library react --pattern "use[A-Z][a-zA-Z]+"
```

— Names all tools and prior behavior but no example of new tools.

`@upstash/context7-mcp@4.7.3` `@upstash/context7-mcp@4.7.1`

└──▷ BREAKING ON UPGRADE

! `query-code` is removed; a dashboard choice that previously enabled it now enables `grep-code`, `glob-code`, `read-code`, and `find-symbol` instead.

! When a custom deployment URL is set via `CONTEXT7_BASE_URL` or `--base-url`, `ctx7` no longer uses the `ctx7 login` session or refreshes it against that host — authentication falls back to `CONTEXT7_API_KEY` only, so on-premise users must set that key to avoid auth failures.

WAS THIS USEFUL? 👍 👎

claudexor

Feedback ⓘ v6.2.0 NOTES ↗ RANK

Claudexor is a multi-harness control plane for Claude Code, Codex, Cursor, and OpenCode with quota-aware rotation across multiple subscriptions.

> Claudexor 6.2.0 adds explicit workspace-preparation garbage collection, splits out an ultrafast processing tier, lets account-hop recovery survive verification refusals, tightens ACP serve's JSON-RPC error handling, and fixes version probes to honor adapter update-control flags.

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

01 **Workspace preparation garbage collection commands** Not checked yet

NEW

85

`claudexor gc preparations` lists held workspace preparations, distinguishing `saved_cleanup` from machine proof and operator confirmation, and `claudexor gc release-preparation` reuses a generation's saved basis after fresh checks. While preparation resources are held, existing thread Apply, turn and file-read routes can now return retryable HTTP 409 `termination_unconfirmed` or `preparation_cleanup_failed`.

Inspect held workspace preparations and distinguish operator-confirmed cleanup from machine-only proof, then release a stalled preparation using its saved basis.

```
$ claudexor gc preparations
claudexor gc release-preparation
```

— Exact subcommands, error codes and a runnable example given v6.2.0

02 **Ultrafast processing tier** Not checked yet BREAKING

80

`processingPreference: ultrafast` is now its own capability and receipt value, separate from the existing `fast` and `priority` tiers, so Codex Fast no longer accidentally selects Ultrafast when it appears first in an account's list. Local and SSH runtime manifests now require Claudexor.app 6.2.0: older apps cannot decode a reviewer panel naming `ultrafast`, and over SSH an older app reports `appUpdateRequired` and cannot connect to the host; strict readers must also accept `ultrafast` in processing capability and receipt enums.

— Names config value and version requirement, no direct command shown v6.2.0

03 **ACP serve JSON-RPC error handling** Not checked yet IMPROVED

80

`claudexor acp serve` now rejects requests arriving before a successful `initialize` with `-32600`, and malformed `mcpServers` or `additionalDirectories` in session creation, load or resume with `-32602`; JSON-RPC batches remain unsupported and also receive `-32600`.

— Exact subcommand and JSON-RPC error codes named `v6.2.0`

---

04 **Account-hop continuation past verification refusals**

Not checked yet NEW 60

In Agent and Ask runs, saved work can now pass a verification or organization subscription refusal from an in-progress account hop to the next eligible account, provided the refused attempt did no work of its own; with no eligible account left the attempt ends `credential_pool_exhausted`. Strict readers must accept `verification_required` in harness events.

— Behavior explained but no config or command to invoke it `v6.2.0`

THINNER COVERAGE BELOW

05 **Version probes honor adapter update-control flags** Not checked yet

IMPROVED 55

Version probes now respect the adapter's declared update controls, including Copilot's `--no-auto-update` and Cursor's `--disable-auto-update`, so a newer cached Copilot no longer answers in place of the installed copy during version checks.

— Named flags but no reproduction steps given `v6.2.0`

└─▷ BREAKING ON UPGRADE

! Local and SSH runtime manifests require Claudexor.app 6.2.0: older apps cannot decode a reviewer panel naming `ultrafast`; over SSH an older app reports `appUpdateRequired` and cannot connect to the host.

! Strict readers must accept `verification_required` in harness events and `ultrafast` in processing capability and receipt enums.

! Existing thread Apply, turn and file-read routes can now return retryable HTTP 409 `termination_unconfirmed` or `preparation_cleanup_failed` while preparation resources are held.

! Older engines (including 6.1.2) refuse a data root served by 6.2.0 with `root_authority_floor_regression`; returning to an older engine requires the complete stopped backup described under v6.1.0, not just an older executable.

WAS THIS USEFUL? 👍 👎

gram

Feedback ⓘ server@2.35.0 NOTES ↗ RANK

Gram provides MCP servers and AI-agent tools from APIs with access controls and observability.

> Gram shipped project-scoped Gateway endpoints with a matching dashboard page, data export test events with per-signal telemetry, direct sign-in provider linking on MCP servers, and audit-logged temporary access grants, alongside a tsnet-based tailnet gateway, Okta sync for MCP servers, and several breaking API removals (external MCP fields, the Polar checkout endpoint, and old hosted MCP server routes).

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

01 Test events for data export destinations

Not checked yet NEW

95

Send a test event to any data export destination via the `dataExports.testRoute` management API endpoint or the `send_data_export_test_event` Platform MCP tool, returning plain-language results (accepted, refused, not found, or unreachable) with per-signal telemetry (one log, one trace, one metric); test signals are tagged `speakeasy.test_event=true` with a matching `speakeasy.test_event_id`, and every test log is named `speakeasy.test_event` with a message starting 'Speakeasy test event:'. In the dashboard, every saved data export gains a Send test event button in its settings for org admins, which shows the response including the test event ID attribute for destinations that nest dotted names such as Datadog.

Verify a data export destination is reachable before relying on it in production — get back a plain-language result per signal type.

```
$ POST /management/dataExports.testRoute
{ "export_id": "<export_id>" }
```

— Endpoint, MCP tool, UI button, and tagging scheme all named. server@2.35.0 dashboard@0.149.0

02 Project-scoped Gateway endpoints and dashboard page

Not checked yet

90

BREAKING

New project-scoped gateway endpoints `organizations.getProjectGateway`, `gatewaySessions.list`, `gatewaySessions.listFacets` (each accepting `project_id`), `telemetry.getProjectGatewayOverview`, and `telemetry.getProjectGatewayUsage` replace organization-wide gateway activity

endpoints; `list_gateway_sessions` in the Platform MCP gains a `project` scope. In the dashboard, each project's MCP Servers list now leads with its Project Gateway, and a dedicated Project Gateway page shows the project's endpoint, installation steps, connections, and activity.

— Every new endpoint and the dashboard surface are named. `server@2.35.0` `dashboard@0.149.0`

---

03 **Link sign-in provider directly to MCP servers** Not checked yet

90

BREAKING

Link a sign-in provider to an MCP server with `mcpServers.update` and `user_session_issuer_id`, or detach it with `clear_user_session_issuer`; this replaces the removed `toolsets.setUserSessionIssuer` endpoint, which callers must migrate away from.

Attach an OIDC sign-in provider to a specific MCP server so session authentication is handled per-server.

```
$ PATCH /mcp-servers/<server_id>
{
  "user_session_issuer_id": "<issuer_id>"
}
```

— Exact replacement endpoint, field, and a runnable example given. `server@2.35.0` `dashboard@0.149.0`

---

04 **tsnet-based tailnet gateway for private network ingress**

Not checked yet IMPROVED

85

Private network ingresses are now hosted via the tsnet-based tailnet gateway instead of the Tailscale Kubernetes operator; the `networkIngress.create` hostname is now optional and defaults to `speakeasy`, making new ingresses reachable at `speakeasy.<tailnet>.ts.net`.

— Exact endpoint, default value, and resulting address given. `server@2.35.0`

---

05 **Temporary access grants with challenge resolution** Not checked yet

NEW

80

Resolve a pending authorization denial while granting temporary access in one atomic call via `access.createTemporaryGrant`, which accepts `challenge_ids`; in the dashboard, org admins can grant themselves temporary access to restricted pages for 30 minutes to 24 hours, logged in the audit log.

Resolve a pending authorization denial at the same time as granting temporary access, keeping the two actions atomic.

```
$ POST /management/access.createTemporaryGrant
{ "scope": "<scope>", "recipient_id": "<user_id>", "resource_id":
"<resource_id>", "challenge_ids": ["<challenge_id>"] }
```

— Endpoint and dashboard duration range both specified. server@2.35.0 dashboard@0.149.0

06 **Disable organization encryption via environment variable**

Not checked yet 80

BREAKING

Set `GRAM_ORG_ENCRYPTION_DISABLED` to run without organization encryption; all other processes that handle OAuth credentials outside local and test environments now refuse to start without encryption built in.

— Exact environment variable and enforced startup behavior named. server@2.35.0

07 **Hosted MCP servers reachable only via their endpoints**

Not checked yet 80

BREAKING

Hosted MCP servers are now reachable only through their endpoints; `/mcp/<slug>` requests, sign-in metadata, sign-in routes, and install pages at the old address return not found, and access tokens issued for the toolset before hosted servers had endpoints are rejected.

— Old-address behavior and token rejection fully specified. server@2.35.0

08 **Platform MCP tools identify by mcp_id and project_slug**

Not checked yet 75

BREAKING

Platform MCP tools now identify MCP servers by `mcp_id` and projects by `project_slug` (as returned by `list_projects`) instead of project IDs; this affects `list_gateway_sessions`, `get_organization_gateway_activity`, `select_assistant_runtime`, and `set_assistant_reasoning_effort` — calls using project IDs will no longer work for these tools.

— Affected tools named; no explicit migration steps given. server@2.35.0

09 **Dry-run preview for Platform MCP write tools** Not checked yet

NEW 75

Pass `confirmed: false` on Platform MCP write tools that take a `confirmed` field to preview the operation without committing it, covering data export pause, resume, level

changes, and risk policy audience changes.

– Exact flag and covered operations are named. `server@2.35.0`

10 **Mark MCP server setup guide as completed** Not checked yet NEW

70

The new `mcpServers.setSetupGuideCompleted` endpoint marks an MCP server's upstream setup guide as done; MCP servers now report `setup_guide_completed_at` to everyone in the project.

– Exact endpoint and reported field named, no usage example. `server@2.35.0`

11 **External MCP fields removed from API** Not checked yet BREAKING

70

The external MCP fields are removed from the API: deployments no longer list external MCP sources, `exclude_external_mcps` is ignored, toolsets no longer carry external MCP header definitions, and `externalmcp` is no longer a valid tool filter type.

– Every removed field and filter type is named. `server@2.35.0` `dashboard@0.149.0`

12 **Recoverable sign-in failures in MCP connection logs**

Not checked yet IMPROVED

65

At the Warnings and errors or Info, warnings and errors export level, MCP connection log exports now also include failed sign-in steps that may still recover, arriving as warnings tagged `speakeasy.mcp_connection.final=false`.

– Exact field and export levels named, no UI path. `server@2.35.0`

13 **Detect drift in remote session client bindings** Not checked yet

DEPRECATED

65

The new `listRemoteSessionClientBindings` endpoint surfaces drift between an MCP Authentication Provider's bound remote session clients and the clients its MCP servers actually use; it is marked deprecated from the start as part of a migration.

– Endpoint named but deprecation limits long-term actionability. `server@2.35.0`

14 **Sync Okta applications as MCP servers** Not checked yet NEW

65

Sync Okta applications as MCP servers using the new Sync from Okta option in Add MCP server, enabled behind the `okta-mcp-import` rollout flag.

– Feature flag and UI entry point named. `dashboard@0.149.0`

THINNER COVERAGE BELOW

15 **Enterprise Managed Auth for MCP Authentication Providers**

Not checked yet NEW 55

Enable Enterprise Managed Auth from an MCP Authentication Provider's page by choosing Okta or any identity provider that publishes signing keys.

— UI path given, no deeper mechanism described. dashboard@0.149.0

16 **Overview and Configuration tabs in MCP Sessions** Not checked yet

NEW 55

MCP Sessions gains an Overview tab listing connections and a Configuration tab listing every MCP Authentication Provider available to the project's servers.

— UI tabs named, no mechanism beyond listing. dashboard@0.149.0

17 **Polar checkout endpoint removed** Not checked yet BREAKING 45

The `usage.createCheckout` endpoint, which created Polar checkout links, is removed.

— Endpoint named but no migration path given. server@2.35.0

└──▷ **BREAKING ON UPGRADE**

! The deprecated `toolsets.setUserSessionIssuer` endpoint has been removed; use `mcpServers.update` with `user_session_issuer_id` to link a sign-in provider, or `clear_user_session_issuer` to detach it.

! Platform MCP tools now require `mcp_id` to identify MCP servers and `project_slug` to identify projects; calls using project IDs will no longer work for `list_gateway_sessions`, `get_organization_gateway_activity`, `select_assistant_runtime`, and `set_assistant_reasoning_effort`.

! Hosted MCP servers are now reachable only through their endpoints; `/mcp/<slug>` requests, sign-in metadata, sign-in routes, and install pages at the old address return not found, and access tokens issued for the toolset before hosted servers had endpoints are rejected.

! The `usage.createCheckout` endpoint (created Polar checkout links) is removed.

! The external MCP fields are removed from the API: deployments no longer list external MCP sources, `exclude_external_mcps` is ignored, toolsets no longer carry external MCP header definitions, and `externalmcp` is no longer a valid tool filter type.

! Organization-wide gateway activity endpoints are replaced by `telemetry.getProjectGatewayOverview` and `telemetry.getProjectGatewayUsage`; callers of the old endpoints must migrate.

! The `toolsets.setUserSessionIssuer` endpoint has been removed; use `mcpServers.update` with `user_session_issuer_id` to link a sign-in provider, and `clear_user_session_issuer` to detach it.

! The external MCP fields are removed from the API: deployments no longer list external MCP sources, `exclude_external_mcps` is ignored when evolving a deployment, toolsets no

longer carry external MCP header definitions, and the `externalmcp` tool type is no longer a valid tool filter.

! `telemetry.getProjectGatewayOverview` and `telemetry.getProjectGatewayUsage` replace the previous organization-wide gateway activity endpoints.

WAS THIS USEFUL? 👍 👎

Composio

Feedback ⓘ @composio/cli@0.6.2 NOTES ↗ RANK

Composio connects AI agents to external tools and services with managed authentication and integrations.

> Composio's CLI now signals transient, retry-safe failures with a dedicated exit code, making it easier to automate retries in scripts and orchestrators.

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

01 Retryable failure exit code in CLI

Not checked yet NEW

72

The Composio CLI now returns exit code `75` when a failed request is safe to resend, letting scripts and orchestrators distinguish transient failures from permanent errors and retry automatically (e.g. `composio <command>; if [ $? -eq 75 ]; then ...; composio <command>; fi` or a shell `until` loop that retries on `75` and exits on other codes).

Retry a Composio CLI command automatically in a shell script when the failure is transient (exit code 75).

```
$ composio <command>; if [ $? -eq 75 ]; then echo 'Transient failure — safe to retry'; composio <command>; fi
```

Wrap a Composio CLI call in a shell retry loop — exit code 75 signals a transient failure that is safe to resend, so the loop retries automatically while other non-zero codes fail fast.

```
$ until composio <command>; do
  rc=$?
  [ $rc -eq 75 ] || exit $rc
  echo 'Transient failure, retrying...'
  sleep 5
done
```

— Names exact exit code and shows runnable retry scripts @composio/cli@0.6.2 @composio/cli@0.6.2-beta.559

WAS THIS USEFUL? 👍 👎

tabularis

Feedback ⓘ v0.33.0 NOTES ↗ RANK

Tabularis is a desktop SQL workspace with built-in drivers for SQL Server, DuckDB, ClickHouse, Redis and an MCP server for Claude, Cursor and Devin.

> tabularis v0.33.0 adds the ability to query flat files directly as tables, enforces read-only connections, tunnels SOCKS5 over SSH, and rounds out the release with AI-assisted editing, a routine debugger, and several PostgreSQL/MySQL data-handling improvements.

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

01 **SOCKS5 proxy over SSH tunnel** Not checked yet NEW 75

Open a SOCKS5 proxy over SSH for drivers that expose the `socks5_proxy` capability, enabled via the SOCKS5 proxy option in SSH tunnel settings.

Route a database driver through an SSH tunnel using SOCKS5 when the driver supports the `socks5_proxy` capability.

📍 In the SSH tunnel settings for the connection, enable SOCKS5 proxy ; tabularis will open a SOCKS5 proxy over the SSH tunnel and pass it to the driver automatically.

— Names exact capability flag and config location v0.33.0

02 **Open data files as queryable tables** Not checked yet NEW 70

Open CSV, TSV, JSON, NDJSON, or Parquet files directly as queryable tables without importing them into a database, using the Open data file option in the connection panel.

Query a Parquet or CSV file as a table in a one-off session without loading it into a database.

📍 In the connection panel, choose Open data file , select your file (CSV, TSV, JSON, NDJSON, or Parquet), then run queries against it in the SQL editor as a normal table.

— Names formats and UI path but no mechanism detail v0.33.0

03 **Enforced read-only mode per connection** Not checked yet NEW 65

Enforce read-only mode per connection via the Enforced read-only mode connection setting so that write queries are blocked before they reach the server.

Block accidental writes on a shared or production connection by enforcing read-only mode at the connection level.

📍 In the connection settings for the target connection, enable `Enforced read-only mode`; tabularis will block any write query before it reaches the server.

— Clear config toggle and blocking behaviour described v0.33.0

04 **Cross-connection row copy between PostgreSQL and SQLite**

Not checked yet NEW 60

Copy table rows between PostgreSQL connections or from PostgreSQL to SQLite, with a transfer plan showing column mapping and conversions, as demonstrated moving an orders table to an orders_archive table.

— Screenshot shows transfer plan mechanism and example tables v0.33.0

THINNER COVERAGE BELOW

05 **Import connections from project config files** Not checked yet

NEW 55

Import connection candidates from project configuration files, as shown with candidates discovered in a .env file and a compose.yaml.

— Screenshot shows concrete file types supported v0.33.0

06 **MySQL per-connection query timeout** Not checked yet NEW

50

Set a per-connection query timeout for MySQL with server-side cancellation.

— Names driver and mechanism but no config key shown v0.33.0

07 **Preview affected rows before UPDATE or DELETE** Not checked yet

NEW 45

Preview the rows an UPDATE or DELETE would affect before running the statement, for PostgreSQL connections.

— Describes the safeguard but no UI trigger given v0.33.0

08 **AI inline completion and API Route provider** Not checked yet NEW

40

Enable opt-in AI inline completion in the SQL editor, and add an `API Route` as an optional AI provider alongside declarative external assistants with built-in providers.

— Names the two additions but no setup steps shown v0.33.0

09 **Routine debugging workspace** Not checked yet NEW 30

Add a capability-driven routine debugging workspace for stored procedures and functions.

— Bare description with no mechanism or UI path v0.33.0

└──▷ **ALSO FROM THESE RELEASES**

WAS THIS USEFUL? 👍 👎

mcp-server-plugin Feedback ⓘ 0.228.vfa_37b_2127d97 NOTES ↗

RANK

> The Jenkins mcp-server-plugin now records the authenticated user identity in MCP-triggered build causes.

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

01 **User identity recorded in MCP build causes** Not checked yet

IMPROVED 40

The MCP server now exposes the authenticated user identity in build causes, so MCP-triggered builds record who initiated them.

— Describes what changed but no mechanism or UI path given 0.228.vfa_37b_2127d97

WAS THIS USEFUL? 👍 👎

magic

Feedback ⓘ v24.4.23 NOTES ↗ RANK

Magic is an instant SECURE Full Stack Apps and AI Agents

> Magic's latest release tightens slot argument validation and expression parsing while adding a new chatbot theme option.

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

01 **Explicit bounds for math.random slot** Not checked yet IMPROVED

65

magic now reads `[math.random]` bounds only from `[min]` and `[max]` child nodes, making random-number slot invocations explicit and consistent.

— Names the exact slot and child nodes, no example given v24.4.23

THINNER COVERAGE BELOW

02 **Childless slot validation in verify-slots** Not checked yet

IMPROVED

55

magic now declares childless slots so the `verify-slots` check flags any ignored children, catching misconfigured slot invocations at validation time.

— Explains mechanism but no concrete command shown v24.4.23

03 **Modern-pro chatbot theme** Not checked yet NEW

25

Adds a `modern-pro` theme to the chatbot widget for a refreshed visual style.

— Only a name and cosmetic description, no detail on usage v24.4.23

WAS THIS USEFUL? 👍 👎

◆ Agent Memory

claude-mem

Feedback ⓘ v15.21.0 NOTES ↗ RANK

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

> claude-mem's biggest addition this window is a Mem0/Supermemory import pipeline, extended with a stop endpoint and duplicate-record fixes, alongside five new IDE/agent integrations, four new provider presets, and tuned auto-clear context-compaction behavior.

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

01 Memory import from Mem0 and Supermemory

Not checked yet NEW

95

Import existing memories from Mem0 or Supermemory directly into a project with `npx claude-mem import --from mem0` or `npx claude-mem import --from supermemory`, using `--dry-run` to preview counts, `--max-records` (up to 100,000), `--timeout-ms` (up to 120,000), `--platform <host>`, `--documents`, and `--source-account` flags; imported memories are filed for Claude Code by default. An in-progress import can be stopped without deleting already-imported data via the `POST /api/setup/import/stop` endpoint, which returns a receipt of what was added. Supermemory `--documents` reruns no longer produce duplicate records, and the previously unsourced 100 MiB total and 10,000-record defaults have been removed (`--max-records` remains).

Migrate your existing Mem0 memories into claude-mem before switching agents, checking counts first with `--dry-run` before committing the write.

```
$ # Preview what would be imported
MEM0_API_KEY=your_key npx claude-mem import --from mem0 --source-account personal-mem0 --user my-user-id --project imported-memory --dry-run

# Run the actual import
MEM0_API_KEY=your_key npx claude-mem import --from mem0 --source-account personal-mem0 --user my-user-id --project imported-memory
```

Import from Supermemory and file the memories under a specific host (e.g. Cursor) so that agent's memory search finds them.

```
$ SUPERMEMORY_API_KEY=your_key npx claude-mem import --from
supermemory --source-account personal-supermemory --namespace
my_namespace --project imported-memory --platform cursor
```

Cancel a long-running import mid-flight without losing already-imported records or triggering an automatic restart.

```
$ curl -X POST http://localhost:<port>/api/setup/import/stop
```

— Full flag list, endpoint, and before/after fix detail given. v15.20.0 v15.19.0

---

02 **New IDE and agent integrations** Not checked yet NEW 90

Added install commands for five new integrations: Hermes Desktop MCP via `claude-mem hermes install --mcp-only --project-cwd <absolute-directory>` (removable with `claude-mem hermes uninstall`) without making claude-mem the default memory provider; GitHub Copilot in VS Code via `npx claude-mem install --ide copilot-vscode`, which writes an MCP entry including `"tools": ["*"]` to `~/.copilot/mcp-config.json`; Zed via `npx claude-mem install --ide zed`, registering claude-mem's MCP server in Zed's `context_servers` for search, timeline, details and explicit saves; per-project Aider memory via `npx claude-mem aider install`, which writes `.claude-mem/aider-memory.md` to the `read:` list in `.aider.conf.yml` and is refreshed with `npx claude-mem aider sync`; and Grok Build via `claude-mem grok install`, which writes `.grok/rules/claude-mem-context.md` and refreshes it after each session summary.

Connect Hermes Desktop to claude-mem's MCP search and save without making claude-mem the default memory provider for the project.

```
$ claude-mem hermes install --mcp-only --project-cwd
/home/user/myproject
```

Add Aider memory to a project so recalled notes are available on each Aider session, then refresh the file after new sessions accumulate.

```
$ # One-time setup inside the project
npx claude-mem aider install

# Refresh the memory file after new sessions
npx claude-mem aider sync
```

— Names every install command and config file written. v15.21.0 v15.19.0

03 **Work-state write guard** Not checked yet NEW 75

Gate `work_state_write` calls with a local check command by configuring `~/.claude-mem/work-state-write-guard.json`; a refused write returns HTTP 409 and nothing is saved.

— Names config path and response code, no example contents. v15.21.0

04 **New OpenAI-compatible provider presets** Not checked yet NEW 65

Added four new provider presets for the OpenAI-compatible provider: `modelgo` pointing to `https://api.modelgo.ai/v1`; AtmoRouter (default model `zai/glm-5.3-flash`); TrustedRouter (default `deepseek/deepseek-v4-flash`, with `trustedrouter/auto` as an opt-in); and HopBase (default `deepseek-v4.1-flash`).

— Names presets and defaults but not the selection command. v15.21.0 v15.19.0

05 **Per-request extra body for API agent** Not checked yet NEW 65

Add extra JSON fields to every API-agent request by setting `CLAUDE_MEM_API_AGENT_EXTRA_BODY` to a JSON object, with reserved fields like `model`, `messages`, and `max_tokens` dropped with a warning.

— Names env var and reserved fields but no sample payload. v15.21.0

06 **Auto-clear continuation prompt and threshold** Not checked yet IMPROVED 65

Auto-clear now resumes with a full continuation prompt ('You are working on "<list>" and recently completed "<task>". Pick back up on our overarching goal "<goal>" from where you left off.') instead of the bare word `continue`, and now triggers at 40% of the context window instead of 33%; override with `CLAUDE_MEM_AUTO_CLEAR_PERCENT` in `settings.json` or the environment.

— Gives before/after values and the override key. v15.21.0

07 **Continue conversations from My Work** Not checked yet NEW 65

Continue a Claude Code or Codex conversation from My Work by typing a message in the new message box under each conversation; `Copy resume command` copies `claude --resume <id>` or `codex resume <id>`, and on macOS `Open in Claude` opens the conversation in the Claude desktop app.

— Names resume commands and UI actions, no worked example. v15.19.0

08 **Export to stdio streams and FIFOs** Not checked yet IMPROVED

60

Export to `/dev/stdout`, `/dev/stderr`, a symlink to either, or a FIFO now writes clean JSON there, with progress lines and warnings sent to the other stream.

— Names supported targets, no example command shown. v15.21.0

THINNER COVERAGE BELOW

09 **Agent-attribution tags in memory search** Not checked yet IMPROVED

45

MCP search and detail results in the server runtime now show `[via <agent>]` tags to identify which agent created each memory, and `/api/sessions/observations` no longer labels every client as Claude Code.

— Names tag format and endpoint, minimal usage detail. v15.19.0

10 **French language support** Not checked yet NEW

20

French is now available in the language picker for My Work, the viewer, first-run setup and Observation TV.

— Bare mention of a new language option. v15.19.0

WAS THIS USEFUL?  

Volcengine OpenViking

Feedback ⓘ react@0.1.13 NOTES ↗

RANK

OpenViking is a context database that manages agent memory, knowledge retrieval, and skills.

> OpenViking's react package ships a batch of FileBrowser host APIs for controlling preview/tree navigation, file-save snapshots and dashboard request overrides, alongside named compile workflows backed by Git snapshot history and improved Wiki conflict handling.

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

01 FileBrowser preview and tree-expansion controls

Not checked yet

NEW

90

Hosts can call openResource(uri, { reveal: false }) to update only the preview without changing the tree selection, or update the controlled `uri` prop to both open the preview and reveal the resource in the directory tree; the new `defaultExpandedUri` prop sets which directory is expanded by default on first render.

Reveal a resource in the FileBrowser tree and open its preview when navigating programmatically from the host application.

javascript

```
// Update the controlled uri prop to open preview and reveal in
tree
<FileBrowser uri={selectedUri} />

// Or, to update the preview without changing the tree selection:
fileBrowserRef.current.openResource(uri, { reveal: false });
```

Pre-expand a specific directory in FileBrowser on first render, so users land with the relevant subtree already open.

javascript

```
<FileBrowser defaultExpandedUri="viking://project/src/components"
/>
```

— Named props with exact call signatures and runnable code examples react@0.1.13

02 busy prop for CommandConsole

Not checked yet NEW

80

Passing the new `busy` prop to `CommandConsole` locks input and submission while a command is pending; the existing `disabled` prop still shows its own unavailable

message.

Lock CommandConsole input while an async command is running so users cannot submit a second command before the first completes.

```javascript
<CommandConsole busy={isCommandRunning} />
```

— Named prop with clear mechanism and runnable example `react@0.1.13`

THINNER COVERAGE BELOW

03 **File-save snapshots in FileBrowser** Not checked yet NEW

55

After saving an editable ordinary file in `FileBrowser`, users can choose whether to keep a snapshot, enter a description, and save it; if the snapshot fails, the entered description is retained so the user can retry without undoing the underlying file save.

— Describes behaviour and failure handling but no API name `react@0.1.13`

04 **Named compile workflows with Git snapshot history** Not checked yet

NEW

55

OpenViking react adds named compile workflow configurations that can be switched through directory-style tabs, backed by native Git snapshot history, plus a CLI guide for adding a new named compile workflow.

— Names the mechanism but no command or flag given `react@0.1.13`

05 **Wiki conflict marker propagation** Not checked yet IMPROVED

55

Wiki conflict markers now propagate from deeply nested files up to every ancestor directory, including collapsed subtrees, and refresh after saving.

— Clear behavioural scope but no UI path or API named `react@0.1.13`

06 **FileBrowser selection reporting for VikingBot** Not checked yet

NEW

45

`FileBrowser` now reports selected tree entries so hosts can pass directory references to `VikingBot`.

— Names the integration target but no API or example `react@0.1.13`

07 **FileBrowser dashboard request overrides** Not checked yet IMPROVED

40

FileBrowser hosts can now override individual experience dashboard requests while the built-in UI, pagination, retries, and cancellation are retained.

— No named API or example given for the override mechanism `react@0.1.13`

---

08 **Explicit snapshots for Wiki edits** Not checked yet NEW

25

Wiki edits now support explicit snapshots.

— One-line mention with no mechanism or example `react@0.1.13`

WAS THIS USEFUL? 👍 👎

◆ Agent Sandboxes

NVIDIA OpenShell

Feedback ⓘ v0.1.9 NOTES ↗ RANK

OpenShell runs autonomous AI agents in policy-controlled sandboxes with isolated filesystem, network, and credential access.

> OpenShell v0.1.9 adds a ServiceDiscovery gRPC API for workspace route discovery, configurable sandbox provisioning timeouts, and workspace output format flags, alongside a wide set of breaking changes spanning the Go SDK's module path, Helm chart metrics and OIDC CA configuration, network policy validation, middleware selectors, and OCSF audit record formats.

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

01 **OIDC CA bundle scoping** Not checked yet BREAKING 95

Control OIDC CA scope outside Helm with `[openshell.gateway.oidc] ca_bundle`, `--oidc-ca-bundle`, or `OPENSHELL_OIDC_CA_BUNDLE`. The Helm chart no longer sets a process-wide `SSL_CERT_FILE` for a private OIDC CA; the CA from `server.oidc.caConfigMapName` now applies only to OIDC discovery and JWKS requests, so other TLS needs must use the new config options.

— Named config key, flag, env var and behavior change v0.1.9

02 **Configurable sandbox provisioning timeouts** Not checked yet NEW 90

Set how long the CLI waits for sandbox provisioning with `openshell sandbox create --timeout <SECONDS>`, overriding the 300-second default and the `OPENSHELL_PROVISION_TIMEOUT` environment variable. Slow isolation backends can also be tuned with three new gateway config keys, `provisioning_repair_window_seconds`, `runtime_bootstrap_grace_seconds`, and `sandbox_attach_request_timeout_seconds`, whose five-minute defaults are unchanged.

Override the default 300-second wait when provisioning is slow in a CI pipeline or on an under-resourced backend.

```
$ openshell sandbox create --name ci-agent --timeout 600
```

— Names flag, env var and config keys with defaults v0.1.9

03 **Helm chart Prometheus metrics toggle** Not checked yet BREAKING

90

Enable Prometheus metrics on the Helm chart with `metrics.enabled` instead of a port value; charts that set `service.metricsPort` to `0` or `false` will fail to render — set `metrics.enabled=false` and keep a valid port instead.

— Exact config keys and migration guidance given `v0.1.9`

---

04 **Go SDK module path change** Not checked yet BREAKING

85

The Go SDK has moved to `github.com/ai-openshell/sdk-go`; `github.com/NVIDIA/OpenShell/sdk/go` is deprecated and its last version is v0.1.8 — update module imports to the new path.

— Exact old and new module paths given `v0.1.9`

---

05 **Output format flags for workspace commands** Not checked yet NEW

80

Get workspace output in table, JSON, or YAML with `-o/--output table|json|yaml` on `openshell workspace create` and `openshell workspace get`.

Inspect workspace policy results in a script without parsing table output.

```
$ openshell workspace get my-workspace -o json
```

— Exact flag and runnable example given `v0.1.9`

---

06 **Higher request-ID replay limits in gateway** Not checked yet

IMPROVED

80

Raise the per-caller request-ID limit and replay window with `mutation_admission_limit_per_caller` (up to 1,000,000) and `mutation_replay_window_seconds` in the gateway config.

— Named config keys and concrete limit value `v0.1.9`

---

07 **OCSF audit record classification change** Not checked yet BREAKING

80

OCSF audit records for authorization denials of write methods now report `class_uid: 6003` / `status_id: 2` with Create, Update, or Delete instead of `type_uid: 600399` / Other (99) — update any queries or alerts that matched on `type_uid: 600399`.

— Exact old and new field values given `v0.1.9`

---

08 **Streaming HTTP middleware hooks capability** Not checked yet

DEPRECATED

75

Opt in to new streaming HTTP middleware hooks with the `openshell.supervisor-middleware.http-hooks` capability; the legacy `EvaluateHttpRequest` and `HttpResponsePreReturn` hooks are deprecated and will be removed in 0.2.0.

— Named capability and deprecated hooks with removal version v0.1.9

---

09 **Wildcard network policy private IP restriction** Not checked yet

75

BREAKING

Wildcard network policy endpoints no longer allow private IP literals on their own; add an exact endpoint or an `allowed_ips` grant for any private address sandboxes need to reach.

— Names config field and required migration action v0.1.9

---

10 **ServiceDiscovery gRPC API** Not checked yet NEW

70

A new `openshell.v1.ServiceDiscovery` gRPC service adds `GetServiceRouteSnapshot` and `WatchServiceRoutes` so proxies and controllers can discover a workspace's service routes.

— Named service and methods but no usage example v0.1.9

---

11 **Postgres migration 010 index build** Not checked yet

BREAKING

70

On Postgres, gateway migration `010` builds indexes on the `objects` table with a plain `CREATE INDEX`, which blocks writes while it runs — large deployments should plan for downtime during upgrade.

— Names migration, table, and operational impact v0.1.9

---

12 **Middleware selector port validation removal** Not checked yet

60

BREAKING

Middleware selectors can no longer include a port; selectors with a port now fail validation — remove the port from any selector.

— Clear migration step, limited mechanism detail v0.1.9

THINNER COVERAGE BELOW

13 **Request ID header for audit tracing** Not checked yet NEW

55

Inspected HTTP responses now carry an `X-OpenShell-Request-ID` header that links them to OCSF records and supervisor traces.

— Named header but no usage example v0.1.9

14 **Richer gateway delete audit records** Not checked yet IMPROVED

45

Gateway delete records now include the workspace and resource name, and report when a delete found nothing to remove or is still in progress.

— Describes change but no named field or endpoint v0.1.9

└──▷ **BREAKING ON UPGRADE**

! The Go SDK has moved to `github.com/ai-openshell/sdk-go`; `github.com/NVIDIA/OpenShell/sdk/go` is deprecated and its last version is v0.1.8 — update module imports to the new path.

! Wildcard network policy endpoints no longer allow private IP literals on their own; add an exact endpoint or an `allowed_ips` grant for any private address sandboxes need to reach.

! On Postgres, gateway migration `010` builds indexes on the `objects` table with a plain `CREATE INDEX`, which blocks writes while it runs — large deployments should plan for downtime during upgrade.

! The Helm chart now controls the gateway's Prometheus endpoint with `metrics.enabled` instead of a port value; charts that set `service.metricsPort` to `0` or `false` will fail to render — set `metrics.enabled=false` and keep a valid port instead.

! Middleware selectors can no longer include a port; selectors with a port now fail validation — remove the port from any selector.

! OCSF audit records for authorization denials of write methods now report `class_uid: 6003` / `status_id: 2` with Create, Update, or Delete instead of `type_uid: 600399` / Other (99) — update any queries or alerts that matched on `type_uid: 600399`.

! The Helm chart no longer sets a process-wide `SSL_CERT_FILE` for a private OIDC CA; the CA from `server.oidc.caConfigMapName` now applies only to OIDC discovery and JWKS requests — set `[openshell.gateway.oidc] ca_bundle`, `--oidc-ca-bundle`, or `OPENSHELL_OIDC_CA_BUNDLE` for other TLS needs.

WAS THIS USEFUL? 👍 👎

cplt

Feedback ⓘ 2026.11.17-141547-aa8caf6 NOTES ↗ RANK

cplt is a kernel-level sandbox for AI coding agents, running Copilot CLI, Claude Code, OpenCode, Gemini CLI, and others with git and gh guards.

> cplt's MCP broker now handles OAuth login flows end-to-end, launching browsers via shims and automatically retrying expired tokens.

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

01 **OAuth login support in MCP broker** Not checked yet NEW

65

The MCP broker adds an `open_url` operation plus `open` and `xdg-open` shims to handle OAuth login flows that require launching a browser. It also accepts gated OAuth remote specs and adds a refresh retry policy to recover from token expiry without manual intervention.

— Names ops and shims but no config/flag usage shown 2026.11.17-141547-aa8caf6

WAS THIS USEFUL? 👍 👎

Docker Sandboxes

Feedback ⓘ v0.54.0 NOTES ↗ RANK

Docker Sandboxes runs AI agents in isolated Docker environments with controlled filesystem, network, and tool access.

> Docker Sandboxes v0.54.0 adds dedicated port management commands and new settings keys for restricted networks, alongside three breaking security and behavior changes to SSH setup, root home directories, and clone-mode Git remotes.

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

01 New settings keys for TLS and Git remote compatibility

Not checked yet NEW 84

Two new settings address restricted-network environments: `tls.postQuantumKeyExchange` can be set to `false` via `sbx settings set tls.postQuantumKeyExchange false` followed by `sbx daemon restart` for firewalls or TLS-inspecting proxies that can't handle post-quantum key exchange; `clone.legacyTcpRemote` can be set to `true` via `sbx settings set clone.legacyTcpRemote true` to enable TCP-compatible Git remote transport for clone-mode sandboxes whose clients ignore `uploadpack`.

Work around a TLS-inspecting proxy that drops post-quantum key exchange so cloud sandbox connections stop stalling.

```
$ sbx settings set tls.postQuantumKeyExchange false
sbx daemon restart
```

— Both settings keys and exact commands given v0.54.0

02 Port management subcommands for sandboxes Not checked yet NEW

83

New `sbx port` subcommands let you manage published ports on a sandbox: `sbx port ls <SANDBOX>` lists current mappings, `sbx port publish <SANDBOX> <HOST_PORT>:<SANDBOX_PORT>` publishes a port (omitting the host port auto-assigns one), and `sbx port unpublish <SANDBOX> <SANDBOX_PORT>` removes mappings, optionally scoped to a single `<HOST_PORT>:<SANDBOX_PORT>` pair.

Publish a specific host-to-sandbox port mapping on a running sandbox, then confirm what is exposed.

```
$ sbx port publish my-sandbox 8080:80
sbx port ls my-sandbox
```

— Exact subcommand syntax with runnable example v0.54.0

03 **Root execution now uses root's home directory** Not checked yet

80

BREAKING

On Unix, running `sbx` as root now uses root's registered home directory; environment overrides `HOME`, `XDG_CONFIG_HOME`, `XDG_STATE_HOME`, `XDG_CACHE_HOME`, `SANDBOXES_STORAGE_ROOT`, and `DOCKER_SANDBOXES_API` are ignored when running as root. Any older root daemon using custom storage must be stopped before upgrading.

— Specific env vars named, clear before/after and migration note v0.54.0

04 **SSH config regenerated against shell-injection risk**

Not checked yet

78

BREAKING

Generated SSH configuration now prevents shell-command injection via crafted host or user names; after upgrading, existing config is stale and must be refreshed by rerunning `sbx setup ssh` (local) or `sbx --cloud setup ssh` (cloud).

— Clear security fix with exact migration commands v0.54.0

05 **Clone-mode Git remote now uses daemon socket** Not checked yet

72

BREAKING

Fetching from a clone-mode sandbox's `sandbox-<name>` Git remote now goes through the daemon socket instead of an unauthenticated localhost TCP port; starting the sandbox refreshes its remote configuration, and the sandbox must be running to fetch.

— Named remote and transport change, clear security rationale v0.54.0

06 **Stable diagnostic IDs in JSON/YAML output** Not checked yet

IMPROVED

62

Running `sbx diagnose --format json` or `--format yaml` now returns a stable `id` per check, intended for reliable script matching.

— Exact flags and purpose stated v0.54.0

THINNER COVERAGE BELOW

07 **MCP mode and server inspection** Not checked yet NEW

55

`sbx inspect` now shows a sandbox's MCP mode and connected servers, and `sbx mcp ls` lists host-level MCP registrations.

— Named commands but no further mechanism detail v0.54.0

---

08 **Network policy decision checking** Not checked yet NEW

53

`sbx policy check network` lets you check network policy decisions and see which rule allowed a given target.

— Named command, thin on mechanism v0.54.0

---

09 **Scheduled run working directory and duration units**

Not checked yet IMPROVED

53

Scheduled runs now accept a `--workdir` or `-w` flag, and schedule durations now accept case-insensitive units.

— Named flag but minimal behavioral detail v0.54.0

---

10 **Skills delete alias** Not checked yet NEW

35

Installed skills can now be removed using the `sbx skills delete` alias.

— Single named command, no further context v0.54.0

└──▷ BREAKING ON UPGRADE

! Rerun `sbx setup ssh` (local) and `sbx --cloud setup ssh` (cloud) after upgrading — generated SSH configuration now prevents shell-command injection via crafted host or user names, and existing config must be refreshed.

! On Unix, running `sbx` as root now uses root's registered home directory; environment overrides `HOME`, `XDG_CONFIG_HOME`, `XDG_STATE_HOME`, `XDG_CACHE_HOME`, `SANDBOXES_STORAGE_ROOT`, and `DOCKER_SANDBOXES_API` are ignored when running as root. Stop any older root daemon using custom storage before upgrading.

! Fetching from a clone-mode sandbox's `sandbox-<name>` Git remote now uses the daemon socket instead of an unauthenticated localhost TCP port; starting the sandbox refreshes its remote configuration and the sandbox must be running to fetch.

WAS THIS USEFUL? 👍 👎

◆ AI Agent Communication

AG-UI

Feedback ⓘ release/2026-11-17 NOTES ↗ RANK

AG-UI: the Agent-User Interaction Protocol. Bring Agents into Frontend Applications.

> AG-UI ships a new `@ag-ui/cloudflare-agents` package that serves the protocol from Cloudflare Durable Objects over HTTP (SSE) and WebSocket, adds stateless tool-approval resumption and reasoning round-tripping for agent runs, and tightens event validation and history compaction in `@ag-ui/client` with accompanying breaking changes.

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

01 Cloudflare Agents Durable Object serving

Not checked yet NEW

95

The new `@ag-ui/cloudflare-agents` package lets any Cloudflare Agents SDK agent serve the AG-UI protocol over HTTP (SSE) and WebSocket by wrapping it with withAGUI(Agent), with one AG-UI thread per Durable Object and the agent's own state exposed as shared thread state. Input members the protocol doesn't define are stripped with a warning, malformed known fields are refused before the run starts (400 over HTTP), and the package requires the `ai` peer dependency at `^6 || ^7` (up from `^5`) per Agents SDK 0.33.

Wrap a Cloudflare Agents SDK agent to serve the AG-UI protocol over HTTP (SSE) and WebSocket from a Durable Object.

typescript

```typescript
import { withAGUI } from '@ag-ui/cloudflare-agents';
import { Agent } from 'agents';

export class MyAgent extends withAGUI(Agent) {
  async onRun(run) {
    const result = await run.streamTextOptions();
    // result spreads instructions, messages, client tools and
abort signal into streamText
  }
}
```

— Full mechanism, named config surfaces, runnable example release/2026-11-17

02 Stateless tool-approval resumption

Not checked yet NEW

85

Tool approval resumption now maps AI SDK `toolApproval` / `needsApproval` to AG-UI interrupts, recording pending approvals in the Durable Object so a resume cannot approve a call the model never made, and an approved tool runs at most once. Rejections carry the user's reason in `payload.reason`, and a resume that leaves a waiting approval unanswered, omits its call, or sends a result for it is refused with `RUN_ERROR` `INVALID_RESUME`.

— Rich named mechanism and error codes, no runnable sample `release/2026-11-17`

---

03 **streamText options helper for agent runs** `Not checked yet` NEW

70

The new run.streamTextOptions() helper spreads a run's instructions, messages, client tools and abort signal directly into `streamText` in one call, and warns when a client-sent field the handler never reads is dropped.

Wrap a Cloudflare Agents SDK agent to serve the AG-UI protocol over HTTP (SSE) and WebSocket from a Durable Object.

typescript

```
import { withAGUI } from '@ag-ui/cloudflare-agents';
import { Agent } from 'agents';

export class MyAgent extends withAGUI(Agent) {
  async onRun(run) {
    const result = await run.streamTextOptions();
    // result spreads instructions, messages, client tools and
abort signal into streamText
  }
}
```

— Named helper with described behaviour and code sample `release/2026-11-17`

---

04 **Reasoning delta compaction and ordering change** `Not checked yet`

70

BREAKING

`@ag-ui/client` now compacts `REASONING_MESSAGE_CONTENT` deltas like text messages to reduce stored history size. In `@ag-ui/client@1.0.5` this changes `compactEvents` output ordering and reasoning-delta compaction, so any code depending on prior event ordering or raw reasoning deltas must be re-verified.

— Named function and version with explicit migration note `release/2026-11-17`

---

05 **Reasoning round-trip through Cloudflare Agents** `Not checked yet`

NEW

65

Reasoning now round-trips through Cloudflare Agents, including `REASONING` events, reasoning messages carried in the snapshot, and provider thinking/thought signatures plus OpenAI reasoning items preserved via `encryptedValue`.

— Named events and fields described, no example release/2026-11-17

06 **CloudflareAgentsClient protocol change and adapter deprecation**

Not checked yet DEPRECATED 65

In `@ag-ui/cloudflare-agents@0.1.0`, `CloudflareAgentsClient` now speaks the `ag-ui:*` WebSocket frames that `withAGUI` serves; the older `AgentsToAGUIAdapter` and SSE/NDJSON helpers remain available but are deprecated.

— Named client and deprecated helpers give clear migration path release/2026-11-17

07 **Request access inside a run** Not checked yet NEW 60

The originating HTTP request is now reachable inside a run via `AGUIRun.request`, enabling header forwarding, auth checks and tracing before processing continues.

Forward auth headers or add tracing inside a run by reading the originating request via `AGUIRun.request`.

typescript

```
import { withAGUI } from '@ag-ui/cloudflare-agents';
import { Agent } from 'agents';

export class MyAgent extends withAGUI(Agent) {
  async onRun(run) {
    const authHeader = run.request.headers.get('Authorization');
    // forward to downstream services or validate before
proceeding
  }
}
```

— Named property with runnable example, thin mechanism release/2026-11-17

08 **Stricter event discriminator validation** Not checked yet BREAKING 60

In `@ag-ui/client@1.0.5`, event enforcement now rejects missing, non-string, or inherited union discriminators that were previously accepted; teams must verify all emitted events conform.

— Specific validation change with verification guidance release/2026-11-17

THINNER COVERAGE BELOW

09 **Pending tool-call tracking in RUN_FINISHED** Not checked yet NEW

55

Client tool calls that the agent did not answer are now tracked via a `pendingToolCallIds` field on `RUN_FINISHED` (the field is omitted for pre-1.0 clients).

— Named field documented, no usage example given release/2026-11-17

10 **Message lookup index for replayed threads** Not checked yet

IMPROVED

50

`@ag-ui/client` speeds up large replayed threads with a new `MessageLookup` index that replaces full-list scans with lookups by message and tool-call id.

— Named index and mechanism, no example or benchmark release/2026-11-17

└──▷ **BREAKING ON UPGRADE**

! In `@ag-ui/client@1.0.5`, event enforcement now rejects missing, non-string, or inherited union discriminators that were previously accepted; verify all emitted events conform.

! In `@ag-ui/client@1.0.5`, `compactEvents` output ordering and reasoning-delta compaction changed; re-verify any code depending on prior event ordering or raw reasoning deltas.

! In `@ag-ui/cloudflare-agents@0.1.0`, the `ai` peer dependency is now `^6 || ^7` (was `^5`), as Agents SDK 0.33 requires.

! In `@ag-ui/cloudflare-agents@0.1.0`, `CloudflareAgentsClient` now speaks the `ag-ui:*` WebSocket frames that `withAGUI` serves; `AgentsToAGUIAdapter` and the SSE/NDJSON helpers remain but are deprecated.

! In `@ag-ui/cloudflare-agents@0.1.0`, input members the protocol does not define are stripped with a warning, and malformed known fields are refused before the run starts (400 over HTTP).

! In `@ag-ui/cloudflare-agents@0.1.0`, a resume that leaves a waiting approval unanswered, omits its call, or sends a result for it is refused with `RUN_ERROR` `INVALID_RESUME`.

WAS THIS USEFUL? 👍 👎

A2A Python SDK

Feedback ⓘ v1.3.1 NOTES ↗ RANK

The A2A Python SDK implements the Agent2Agent protocol for discovering agents and exchanging tasks and messages.

> A2A Python SDK's single release in this window adds httpx2 client support and extends DatabaseTaskEventStream with server call context.

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

01 **ServerCallContext in DatabaseTaskEventStream** Not checked yet

IMPROVED 55

Adds a `ServerCallContext` parameter to `DatabaseTaskEventStream` so cluster-aware task event streams can now receive per-call server context.

— Names exact class and parameter but no usage detail v1.3.1

02 **httpx2 client support** Not checked yet NEW 40

Adds httpx2 support, allowing the SDK client to run against applications built on the httpx2 HTTP library.

— States capability but no config or migration detail v1.3.1

WAS THIS USEFUL? 👍 👎

◆ Applied AI Agents

OpenClaw

Feedback ⓘ v2026.11.5 NOTES ↗ CODE ↗ RANK

OpenClaw is a personal AI assistant platform that connects messaging channels, tools, and model providers through a gateway.

> OpenClaw's latest release focuses on Gateway resiliency and plugin API changes, adding restart-safe conversation queues, interactive Telegram widgets, a configurable compaction cap, and several breaking changes to internal plugin and harness interfaces.

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

01 Compaction summary character cap

Not checked yet NEW

78

The `maxSummaryChars` setting controls the compaction summary character cap, letting operators keep context windows manageable in long-running conversations.

Cap the size of compaction summaries to keep context windows manageable in long-running conversations.

yaml

```yaml
maxSummaryChars: 4000
```

— Named config key with a runnable example value v2026.11.5

02 Gateway plugin and harness API changes

Not checked yet BREAKING

76

Third-party harnesses must now register session-keyed active runs inside the host `runAttempt` callback and retain the returned cleanup function. Plugin RPC handlers must replace `chatAbortControllers` / `chatQueuedTurns` with authorized Gateway method dispatch, and start-owner observations must use `startDeadlineAtMs` instead of `expiresAtMs`.

— Names exact callbacks and fields needing migration v2026.11.5

03 IRC empty allowFrom now denies all senders

Not checked yet

62

BREAKING

IRC channels with an explicitly empty `allowFrom` list now act as a deny rule for all senders in that channel, even when a wildcard channel or global sender list would

otherwise permit them; existing configurations with an empty per-channel list should be reviewed.

— Explains override behavior and flags configs needing review v2026.11.5

THINNER COVERAGE BELOW

04 **Restart-safe conversation queues in Gateway** Not checked yet

IMPROVED 45

Queued messages, send permissions, and final chat writes now survive Gateway restarts without requiring a manual resend.

— Describes behavior but no config or command to invoke it v2026.11.5

05 **Interactive Telegram widgets** Not checked yet NEW 41

Interactive widgets now render in Telegram messages, including within private-chat topics.

— States what renders but no setup steps given v2026.11.5

06 **Gemini Decision model support** Not checked yet NEW 18

OpenClaw adds support for the Gemini Decision model, per the release summary.

— Only mentioned by name, no detail provided v2026.11.5

└──▷ **BREAKING ON UPGRADE**

! Third-party harnesses must register session-keyed active runs inside the host `runAttempt` callback and retain the returned cleanup function.

! Plugin RPC handlers must replace `chatAbortControllers` / `chatQueuedTurns` with authorized Gateway method dispatch.

! Start-owner observations must use `startDeadlineAtMs` instead of `expiresAtMs`.

! IRC channels with an explicitly empty `allowFrom` list now act as a deny rule for all senders in that channel, even when a wildcard channel or global sender list would otherwise permit them — existing configurations with an empty per-channel list should be reviewed.

WAS THIS USEFUL? 👍 👎

Feynman

Feedback ⓘ v0.5.71 NOTES ↗ RANK

Feynman is a personal AI research assistant that searches papers and the web, builds literature reviews, and produces cited research reports.

> Feynman expands its web search configuration with three new providers and makes search and page-fetching more resilient to failures.

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

01 **Three new web search providers** Not checked yet NEW 75

SpecStream, Autom, and cloro can now be selected as web search providers using `feynman search set <provider>`, for example `feynman search set cloro` to switch away from a rate-limited DuckDuckGo.

Switch to a newly supported provider when DuckDuckGo is rate-limited or returning incomplete results.

```
$ feynman search set cloro
```

– Exact command and three provider names given, no further detail on differences. v0.5.71

THINNER COVERAGE BELOW

02 **Automatic search provider failover** Not checked yet IMPROVED 35

Search now automatically fails over to the next provider when one times out, instead of failing the whole search.

– Describes behavior change but no config or trigger details. v0.5.71

03 **More complete page extraction fallback** Not checked yet IMPROVED 25

Fetched pages now return more complete text when the main extractor misses content.

– Thin description with no mechanism or scope named. v0.5.71

WAS THIS USEFUL? 👍 👎

◆ Browser & Computer Use

browser-use

Feedback ⓘ 1.0.6 NOTES ↗ RANK

The browser-use Python library automates web browsers for AI agents.

> browser-use 1.0.6 introduces selectable BU 4.0 model tiers with a costlier default, plus a smoother first-run authentication experience with cloud live-view links.

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

01 Selectable BU 4.0 model tiers

Not checked yet NEW 88

Pass `llm='bu-4.0-mini'`, `llm='bu-4.0'`, `llm='bu-4.0-pro'`, or `llm='bu-fast'` to Agent() and ChatBrowserUse() to choose cost and reasoning level, from cheapest/fastest to highest-reasoning.

Run a complex task at maximum reasoning, or keep costs low for simple tasks, by explicitly selecting a BU 4.0 tier.

python

```python
from browser_use import Agent

# Hardest tasks — highest reasoning
result = Agent('Negotiate and book the cheapest available flight to Tokyo next week', llm='bu-4.0-pro').run_sync()
print(result.final_result())

# Simple tasks — lowest cost
result = Agent('What is the title of the browser-use GitHub repo?', llm='bu-4.0-mini').run_sync()
print(result.final_result())
```

— Named flag with concrete values and runnable example 1.0.6

02 Default model switched to bu-4.0

Not checked yet BREAKING 80

Agent() and ChatBrowserUse() now default to `bu-4.0` instead of `bu-fast`; this raises input/output token costs 15× ($1.80/$9.00 per 1M vs $0.12/$0.60). Pass `llm='bu-fast'` to restore the previous model and pricing.

— Exact pricing and migration path given 1.0.6

THINNER COVERAGE BELOW

03 **Automatic login on first run** Not checked yet IMPROVED

36

The first run without an API key now automatically opens Browser Use login and continues after approval, with cloud browsers printing their live-view link.

– Behavior described but no config detail given 1.0.6

└──▷ **BREAKING ON UPGRADE**

! Agent() and ChatBrowserUse() now default to `bu-4.0` instead of `bu-fast`; input/output tokens cost 15× more ($1.80/$9.00 per 1M vs $0.12/$0.60). Pass `llm='bu-fast'` to restore the previous model and price.

WAS THIS USEFUL?  

DEPLOY

◆ Inference Providers & GPU Clouds

Modal

Feedback ⓘ changelog · Nov 18, 2026 CHANGELOG ↗ RANK

Modal is a cloud infrastructure for teams that develop, train, and serve AI applications at scale.

> Modal's 1.6.4 release adds a suite of CLI tooling for managing Images, tokens, and usage stats without code, alongside smaller API and Sandbox refinements.

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

01 Token lifecycle management commands

Not checked yet NEW 88

New token lifecycle controls: Workspace.member_tokens.disable() or `modal workspace member-tokens disable` disables API tokens without hard-deleting them (with a corresponding re-enable method and CLI command), and `modal token rotate` rotates the active profile token by creating a replacement and deleting the old one.

Disable a compromised team member token immediately without deleting it, so it can be re-enabled later if needed.

```
$ modal workspace member-tokens disable <token-id>
```

— Named methods, CLI commands, and a runnable usage example changelog-20261118-1349b992

02 Image management CLI commands

Not checked yet NEW 86

New CLI commands let you manage named Modal Images end-to-end: `modal server deploy` deploys an app directly from a named Modal Image or third-party registry URI with no code required; `modal image build` builds a Docker image from a local Dockerfile and publishes it as a named Modal Image; `modal image pull` pulls an image from a third-party registry and publishes it as a named Modal Image; and `modal image names delete` deletes an image name without removing the underlying image.

— Four named commands with exact behaviour, no usage example given

changelog-20261118-1349b992

03 App and Sandbox stats methods

Not checked yet NEW 79

New stats methods expose usage data programmatically and via CLI: App.stats() or `modal app stats` returns high-level Function, Server, and Sandbox stats for an app, while App.sandboxes.stats() queries Sandbox-specific stats including costs, with flexible aggregation by Sandbox tags.

— Named methods and CLI command, aggregation detail, no example

changelog-20261118-1349b992

04 **Sandbox CPU env vars no longer set by default** Not checked yet

67

BREAKING

Sandboxes no longer set `OMP_NUM_THREADS`, `MKL_NUM_THREADS`, `OPENBLAS_NUM_THREADS`, `BLIS_NUM_THREADS`, `ORT_INTRA_OP_NUM_THREADS`, or `ORT_INTER_OP_NUM_THREADS` environment variables by default; the visible CPU count inside the Sandbox now reflects the CPU limit.

– Names all six env vars and the CPU-count change precisely changelog-20261118-1349b992

05 **Function lookup by ID** Not checked yet NEW

62

The new Function.from_id() method looks up a Function by its ID, supporting observability and autoscaling APIs for Modal Cls objects.

– Named method and purpose, but no usage shown changelog-20261118-1349b992

THINNER COVERAGE BELOW

06 **Image.build() without an App** Not checked yet IMPROVED

50

Image.build() can now be called without providing an App — it automatically creates an ephemeral App to run the build.

– Clear before/after behaviour but no example or limits changelog-20261118-1349b992

07 **Endpoint library listing command** Not checked yet NEW

50

The new `modal endpoint library` CLI command lists models that can be served with Modal Endpoints.

– Named runnable command, but no further mechanism detailed changelog-20261118-1349b992

08 **Python 3.10 deprecation warning** Not checked yet DEPRECATED

40

Modal now issues deprecation warnings for Images using Python 3.10; support will be dropped as of version 1.7.0.

– Clear removal timeline but no migration steps given changelog-20261118-1349b992

└──▷ BREAKING ON UPGRADE

! Modal now issues deprecation warnings for Images using Python 3.10; support will be dropped as of version 1.7.0.

WAS THIS USEFUL?  

Cloudflare Workers AI

Feedback ⓘ changelog · Nov 18, 2026

CHANGELOG ↗ RANK

Cloudflare Workers AI runs open-weight models, including the Clef decision models, as serverless inference on Cloudflare's global network.

> Cloudflare pushed a 1.0 release candidate of its `cf` CLI with live log tailing, dry-run previews, and a long list of breaking changes, alongside custom CA/mTLS support for Workers VPC Services, AI Search file summarization, per-appliance WAN gateway addressing for Magic WAN HA pairs, and TIFF input support in Cloudflare Images.

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

01 **Custom CA and mTLS certificates for VPC Services** Not checked yet

NEW 95

Upload a private CA certificate with `npx wrangler cert upload certificate-authority --ca-cert <path> --name <name>`, then reference it on a VPC Service via `--certificate-authority` and `--cert-verification-mode verify_full` to verify private services using your own CA. Upload an mTLS client certificate with `npx wrangler cert upload mtls-certificate --cert <cert> <key> --name <name>` and attach it with `--client-certificate` to authenticate to services that require mutual TLS.

Connect a Workers VPC Service to a PostgreSQL database that uses a private CA and requires a client certificate for mTLS.

```
$ npx wrangler cert upload certificate-authority --ca-cert
./ca.pem --name private-ca
npx wrangler cert upload mtls-certificate --cert ./client.pem
./client-key.pem --name private-client
npx wrangler vpc service private-db \
  --type tcp \
  --tcp-port 5432 \
  --app-protocol postgresql \
  --hostname db.internal.example \
  --tunnel-id TUNNEL_ID \
  --certificate-authority CA_CERTIFICATE_ID \
  --client-certificate CLIENT_CERTIFICATE_ID \
  --cert-verification-mode verify_full
```

— full command sequence with every flag named changelog-20261118-b68d143d

02 **Live log streaming via cf observability tail** Not checked yet

NEW 90

Stream live Workers observability events as newline-delimited JSON with `cf observability tail`, scoping to a Worker with `--worker`, a zone with `--zone`, or the whole account with `--all`, and filtering with `--filters` as inline JSON or from a file.

Tail live log events for a specific Worker, narrowing results by a saved filter file, to monitor production traffic in real time.

```
$ cf observability tail --worker my-worker --filters
@filters.json
```

— runnable command with all scoping flags named `changelog-20261118-b68d143d`

03 **Per-appliance WAN gateway addressing for Magic WAN HA**

Not checked yet NEW 90

Configure independent static addressing for each appliance in a Magic WAN high-availability pair via `PATCH https://api.cloudflare.com/client/v4/accounts/{account_id}/magic/sites/{site_id}/wans/{wan_id}` with `static_addressing` fields `address`, `gateway_address`, `secondary_address`, and `secondary_gateway_address`, so failover can use an independent circuit per appliance.

Give each appliance in a Magic WAN high availability pair its own ISP gateway so failover uses an independent circuit.

```
$ curl --request PATCH \

https://api.cloudflare.com/client/v4/accounts/{account_id}/magic/
sites/{site_id}/wans/{wan_id} \
  --header "Authorization: Bearer <API_TOKEN>" \
  --header "Content-Type: application/json" \
  --data '{
    "static_addressing": {
      "address": "203.0.113.2/30",
      "gateway_address": "203.0.113.1",
      "secondary_address": "198.51.100.2/30",
      "secondary_gateway_address": "198.51.100.1"
    }
  }'
```

— named endpoint, method, and all fields with a runnable example

changelog-20261118-b68d143d

04 **cf CLI 1.0 breaking changes and migrations** Not checked yet

90

BREAKING

`cf workers secrets bulk` is renamed to `cf workers secrets bulk-edit` with no alias; Build Output moves to version 1 at `.cloudflare/output/v1/` and `cf deploy` automatically creates a new build there on first run; Vite-based projects now require `2.0.0-rc.0` or later and legacy Wrangler-based projects require Wrangler `4.163.0` or later; API commands generated from camelCase or snake_case names now use kebab-case (e.g. `cf accounts billing getCredits` is now `cf accounts billing get-credits`); `cf previews deploy` now takes the Preview name as `--preview-name` instead of a positional argument; commands run with `cf` store state in the project's `.cloudflare/state` directory; `cf` no longer saves your account choice between runs, so set `accountId` in `cloudflare.config.ts` or `CLOUDFLARE_ACCOUNT_ID` to skip the prompt; streaming tail consumers move to `experimental.streamingTailConsumers` in `cloudflare.config.ts`; destructive commands run without a terminal now exit with an error unless you pass `--force`; prompts, progress, and status messages are written to stderr, so scripts reading messages from stdout must read stderr instead.

— enumerates all ten migration changes verbatim changelog-20261118-b68d143d

05 **New cf CLI deploy, dry-run, and account options** Not checked yet

NEW

78

`cf deploy ./dist` deploys a directory of static assets directly, skipping the build step and creating an assets-only Worker. `cf deploy --dry-run` previews the project setup changes a deploy would make, and `--dry-run` can be passed to any API command to preview what it would do, with sensitive fields redacted in request previews. `--temporary-account` creates a short-lived Cloudflare account without signing in, and `CLOUDFLARE_API_HEADERS` sets custom headers to send with all API requests.

— four distinct flags named but each thinly described changelog-20261118-b68d143d

06 **AI Search file summaries** Not checked yet NEW

70

Set `indexing_options.file_summary` when creating or updating an AI Search instance to have the LLM generate a summary, topics, and keywords for each text document that produces at least three chunks; the result is stored in the item's `context` metadata.

— names the config key and storage location changelog-20261118-b68d143d

07 **TIFF input support in Cloudflare Images** Not checked yet NEW

60

Cloudflare Images now accepts TIFF as an input format, enabling dynamic transcoding of TIFF master files to web formats such as WebP or AVIF with per-request resizing.

— describes capability but no concrete API call shown `changelog-20261118-b68d143d`

└──▷ BREAKING ON UPGRADE

! `cf workers secrets bulk` is renamed to `cf workers secrets bulk-edit` with no alias.

! Build Output moves to version 1 at `.cloudflare/output/v1/` ; `cf deploy` will automatically create a new build on first run in this location.

! Vite-based projects now require `2.0.0-rc.0` or later; legacy Wrangler-based projects require Wrangler `4.163.0` or later.

! API commands generated from camelCase or snake_case names now use kebab-case (e.g. `cf accounts billing getCredits` is now `cf accounts billing get-credits`).

! `cf previews deploy` now takes the Preview name as `--preview-name` instead of a positional argument.

! Commands run with `cf` store state in the project's `.cloudflare/state` directory.

! `cf` no longer saves your account choice between runs; set `accountId` in `cloudflare.config.ts` or `CLOUDFLARE_ACCOUNT_ID` to skip the prompt.

! Streaming tail consumers move to `experimental.streamingTailConsumers` in `cloudflare.config.ts` .

! Destructive commands run without a terminal now exit with an error unless you pass `--force` .

! Prompts, progress, and status messages are written to stderr; stdout holds only command output, so scripts that read messages from stdout must read stderr instead.

WAS THIS USEFUL? 👍 👎

Eigen Labs Darkbloom

Feedback ⓘ v0.10.14 NOTES ↗

CODE ↗ RANK

Darkbloom is a private Inference Network on Idle Macs

> Darkbloom's v0.10.14 release drops MDM enrollment in favor of App Attest-only device verification, adds finer environment-variable control over Qwen and Gemma inference limits, and tightens limits on the local batch API while widening inline video format support.

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

01 Qwen and Gemma inference limit controls

Not checked yet IMPROVED

93

`DARKBLOOM_MAX_IMAGES_PER_REQUEST` now allows Qwen models up to 64 and Gemma 4 26B up to 24 inline images per request (up from 16). `DARKBLOOM_QWEN_VIDEO_MAX_FRAMES` and `DARKBLOOM_QWEN_VIDEO_MAX_TOKENS` let Qwen models read up to 32 timestamped frames per video. `DARKBLOOM_LOOP_EXIT_THINKING`, set to a list excluding `qwen3_5_moe`, lets Qwen 3.6 35B end repeating thoughts with an answer instead of running out of tokens.

Tune Qwen video ingestion to cap frames and tokens when running on memory-constrained hardware.

```
$ DARKBLOOM_QWEN_VIDEO_MAX_FRAMES=8
DARKBLOOM_QWEN_VIDEO_MAX_TOKENS=1024 darkbloom start
```

— Every env var named with exact limits and defaults. v0.10.14

02 Local API `$ref` resolution for tool schemas

Not checked yet

NEW

86

By default, the local API now inlines shared `$ref` schema references for tool schemas on Gemma 4 26B and Qwen 3.8 27B. Setting `DARKBLOOM_TOOLSCHEMA_REF_RESOLVE_MODELS=off` before `darkbloom start` or `darkbloom cluster` disables this and passes `$ref`s through untouched.

Opt out of automatic `$ref` inlining for tool schemas on the local API — useful if your tooling sends schemas with intentional `$ref` pass-through.

```
$ DARKBLOOM_TOOLSCHEMA_REF_RESOLVE_MODELS=off darkbloom start
```

— Named env var, models, and exact opt-out command. v0.10.14

03 **MDM-free App Attest device verification** Not checked yet IMPROVED

75

`darkbloom unenroll` removes the old Darkbloom Provider Enrollment MDM profile and opens System Settings; Macs are now verified with App Attest only instead of requiring an MDM profile.

Remove the legacy MDM enrollment profile after upgrading, so the Mac is verified by App Attest only.

```
$ darkbloom unenroll
```

— Exact command and profile name given, outcome clear. v0.10.14

04 **Stricter limits on chat completions batch endpoint**

Not checked yet

73

BREAKING

The local API's `/v1/chat/completions/batch` endpoint now enforces a limit of 64 requests per batch, returning `400 batch_too_large` if exceeded, alongside a 32 MiB cap on kept Responses history and a 32 MiB relay backlog limit.

— Endpoint, error code and exact limits all specified. v0.10.14

05 **Broader video MIME acceptance for inline uploads** Not checked yet

IMPROVED

68

Inline MP4 and QuickTime videos are now accepted regardless of their MIME label (e.g. `video/mov`, `application/octet-stream`, `video/x-m4v`); other containers now return `unsupported_media` instead of `invalid_media`.

— Exact MIME types and new error code named. v0.10.14

THINNER COVERAGE BELOW

06 **Two-Mac pairing restored for Gemma 4 26B** Not checked yet IMPROVED

40

Two-Mac pairs now serve Gemma 4 26B again, in both layer-split and tensor-parallel layouts.

— States what works again but no setup steps. v0.10.14

07 **Disk space check in `darkbloom update`** Not checked yet IMPROVED

40

`darkbloom update` now reports when the disk is too full for an update and states how much space is needed.

— Names command but no exact numbers given. v0.10.14

WAS THIS USEFUL? 👍 👎

AWS Amazon Bedrock Feedback ⓘ

Distribute already-accepted private offers from AWS Marketplace (2026-11-16)

BLOG ↗

Amazon Bedrock provides managed foundation-model inference, customization, knowledge bases, and APIs for generative AI applications.

Amazon Bedrock adds a Distribution status column on the Manage subscriptions page to surface undistributed AWS Marketplace agreements.

NEW INTEGRATION Distribute private offers for third-party models from AWS Marketplace agreements

WAS THIS USEFUL? 👍 👎

Nebius AI Cloud

Feedback ⓘ November 9–15, 2026 (2026-11-16)

BLOG ↗ RANK

Nebius AI Cloud provides GPU compute, managed Kubernetes, distributed training infrastructure, and serverless AI workloads.

> Nebius AI Cloud's Observability Agent gained JSON structured log parsing, Managed SkyPilot extended to external Kubernetes clusters, and default quotas for B200/B300 GPUs were cut to zero without reservations.

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

01 **Zero default quota for B200/B300 GPUs without reservation**

Not checked yet BREAKING 60

Default Compute quotas on NVIDIA B200 and B300 GPUs for regular VMs without reservations are now zero in all quota tiers, in all applicable regions — existing setups relying on non-zero default quotas for these GPUs will no longer have access without an explicit reservation.

– Clear before/after and migration note to reserve capacity

November 9–15, 2026 (2026-11-16)

THINNER COVERAGE BELOW

02 **JSON structured log parsing in Observability Agent**

Not checked yet IMPROVED 55

Nebius Observability Agent v1.0.10 now parses structured Pod logs in JSON before sending them, surfacing `message`, `severity`, and other fields as separate columns in Observability Logs.

– Names mechanism and fields but no config steps November 9–15, 2026 (2026-11-16)

03 **New tutorials for Saturn Cloud and vLLM on Managed Kubernetes**

Not checked yet NEW 55

A new tutorial covers deploying Saturn Cloud on a Managed Service for Kubernetes cluster to run JupyterLab and RStudio workspaces, scheduled jobs, and model deployments. Another new tutorial covers running vLLM inference on a Managed Kubernetes cluster with Managed Service for SkyPilot, including launching a vLLM server, serving a model on a single GPU, and sending inference requests.

– Documentation walkthroughs give a clear starting point November 9–15, 2026 (2026-11-16)

04 **Managed SkyPilot support for external Kubernetes clusters**

Not checked yet IMPROVED 45

Managed SkyPilot now connects to Kubernetes clusters outside of Nebius AI Cloud, in addition to Google Cloud, AWS, and Runpod.

— States scope expansion but no setup detail November 9–15, 2026 (2026-11-16)

└──▷ **BREAKING ON UPGRADE**

! Default Compute quotas on NVIDIA B200 and B300 GPUs for regular VMs without reservations are now zero in all quota tiers, in all applicable regions — existing setups relying on non-zero default quotas for these GPUs will no longer have access without an explicit reservation.

WAS THIS USEFUL? 👍 👎

AWS Amazon SageMaker AI Feedback ⓘ November 16, 2026

NOTES ↗ RANK

Amazon SageMaker AI provides managed tools and infrastructure for building, training, fine-tuning, and deploying machine-learning models.

> SageMaker HyperPod adds cluster operations tracking for monitoring scaling and deep health check requests.

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

01 **Cluster operations tracking in HyperPod** Not checked yet NEW 45

The new cluster operations feature lets users track scaling and deep health check requests on SageMaker HyperPod clusters, and filter cluster events by request.

— Describes feature but no UI path or API name given November 16, 2026

WAS THIS USEFUL? 👍 👎

Fireworks AI

Feedback ⓘ

Dataset creation, upload, and download APIs retired · NOTES ↗ · RANK

Fireworks AI provides hosted inference, model fine-tuning, and deployment APIs for open-weight models.

> Fireworks AI retired its managed Dataset API in favor of direct-from-storage training data loading and local JSONL batch workloads, breaking any workflow that creates or uploads datasets.

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

01 Dataset creation and management APIs retired

Not checked yet

86

BREAKING

Creating, updating, uploading, validating, previewing, splitting, and downloading datasets now return `501 Not Implemented` across the REST/gRPC API, `firectl`, and the Python SDK — any workflow that creates or uploads datasets will break on upgrade. Legacy dataset records can still be inspected and cleaned up using `firectl dataset list`, `firectl dataset get`, and `firectl dataset delete`, which remain active after the retirement.

Delete an existing managed dataset record after migrating training data to local or cloud storage.

```
$ firectl dataset list
firectl dataset delete <dataset-id>
```

— Names exact status code, affected surfaces, and surviving commands

Dataset creation, upload, and download …

THINNER COVERAGE BELOW

02 Direct-from-storage training data loading

Not checked yet · NEW

50

The Training API can now load and tokenize training data directly from local files or your own cloud storage, with no dataset resource created or stored on Fireworks.

— Explains mechanism but no concrete command shown

Dataset creation, upload, and download …

03 Local JSONL batch workloads on preemptible deployments

Not checked yet · NEW

45

Fireworks AI now supports running bulk asynchronous batch workloads directly from local JSONL files on preemptible deployments.

— Describes capability but no command or config given

Dataset creation, upload, and download …

└──▷ **BREAKING ON UPGRADE**

! Creating, updating, uploading, validating, previewing, splitting, and downloading datasets now return 501 Not Implemented across the REST/gRPC API, `firectl`, and the Python SDK — any workflow that creates or uploads datasets will break on upgrade.

WAS THIS USEFUL? 👍 👎

RunPod

Feedback ⓘ BreakingREST API v1 retired NOTES ↗ RANK

RunPod is an AI developer cloud platform for experimenting, training, fine-tuning, deploying, and scaling AI workloads across 31 global regions.

> RunPod retired REST API v1 and launched REST API v2 covering Pods, Serverless, storage, and billing, alongside a one-command automated migration path and new scoped API keys for finer-grained access control.

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

01 REST API v1 retired, REST API v2 launched

Not checked yet

95

BREAKING

REST API v1 is retired as of November 15, 2026 and is no longer available — calls to v1 endpoints lose access until migrated to v2. The new REST API v2 base URL `https://api.runpod.io/v2` supports creating and managing Pods, querying Serverless endpoints, provisioning Network volumes, and retrieving billing data. The full OpenAPI specification is available via `GET https://api.runpod.io/v2/openapi.json` for generating typed clients or validating requests.

Fetch the RunPod v2 OpenAPI schema to generate a typed client or validate requests in your toolchain.

```
$ curl --request GET --url https://api.runpod.io/v2/openapi.json
```

— Exact endpoint, base URL, resources and retirement date given

BreakingREST API v1 retired

02 Automated migration command for v1 and GraphQL integrations

Not checked yet NEW

80

Existing REST v1 or GraphQL integrations can be migrated to v2 automatically by running `/runpod:migrate` in a coding agent with the RunPod skills plugin installed, avoiding manual rewriting of each call.

Automatically migrate an existing REST v1 or GraphQL integration to v2 without manually rewriting each call.

```
$ /runpod:migrate
```

— Exact command given though plugin setup details are unspecified

BreakingREST API v1 retired

THINNER COVERAGE BELOW

03 **Scoped API keys for RunPod services** Not checked yet NEW

45

New scoped API keys grant per-scope access to RunPod services, including granular access for REST API v2. There is no default full-access option, so every key is issued only the scopes it needs.

— Describes behavior but gives no key-creation steps or UI path

BreakingREST API v1 retired

└──▷ **BREAKING ON UPGRADE**

! REST API v1 is retired as of November 15, 2026 and is no longer available — integrations that call v1 endpoints will lose access until migrated to v2.

WAS THIS USEFUL? 👍 👎

◆ **AI / LLM Gateways**

OpenRouter

Feedback ⓘ updated · Nov 18, 2026 NOTES ↗

OpenRouter is a routing gateway that exposes hundreds of models from many providers behind one OpenAI-compatible API with failover and unified billing.

NEW INTEGRATION Hugging Face Inference Providers routing via OpenRouter

WAS THIS USEFUL? 👍 👎

BerriAI LiteLLM

Feedback ⓘ v1.108.11 NOTES ↗ RANK

LiteLLM is an AI gateway that calls over 100 model APIs in OpenAI format with cost tracking, guardrails, load balancing, and logging.

> LiteLLM added native generateContent support for custom providers registered via custom_provider_map.

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

01 **Native generateContent for custom providers** Not checked yet NEW

55

Custom providers registered in `custom_provider_map` can now call the native `generateContent` API directly.

– Names config key but lacks mechanism detail or examples v1.108.11

WAS THIS USEFUL? 👍 👎

Mozilla AI Otari

Feedback ⓘ v0.49.0 NOTES ↗ RANK

Mozilla AI Otari is an open-source AI gateway you run yourself. One endpoint for 40+ providers, with virtual keys, budgets, and usage tracking.

> Mozilla AI Otari shipped workspace-level MCP server governance, scoped API keys, organization invitation workflows, and new Gemini decision provider and Brave/Tavily search adapters, alongside several breaking changes touching config, email formats, and service-key file access.

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

01 **MCP server catalog and governance in workspaces** Not checked yet

NEW

70

Adds a workspace-level MCP server catalog with sign-in, health checks, and member request handling, a web UI page covering the catalog, sign-in, and member requests, and per-key access controls so administrators can see which API keys may use a workspace's MCP servers.

Navigate to the MCP servers page to browse the catalog, sign in a server, and review member access requests for a workspace.

> 📍 In the web UI, go to Workspace › MCP Servers, select a server from the catalog, sign it in, and review pending member requests.

Check which API keys have permission to use a workspace's MCP servers before granting a new integration access.

> 📍 In the web UI, go to Workspace › MCP Servers, then select a server and open the Keys tab to see the list of authorized keys.

– Names catalog, UI page, and per-key controls with exact navigation steps. v0.49.0

02 **Service key management for end users** Not checked yet BREAKING

65

Service keys can now own files as their end users, with unreachable file IDs refused instead of causing silent failures; a service key can read its end users against its cap via the API, and the dashboard shows its end users against its cap. This is a breaking change — workflows that previously passed unresolvable file IDs will now receive errors.

– Explains mechanism and before/after but no named endpoint. v0.48.0

03 **API key endpoint scoping** Not checked yet BREAKING

60

API keys can now be scoped to the specific endpoints they are permitted to call. This is a breaking change: existing keys with unrestricted access will need to be updated.

— Names the behavior change and migration impact, no exact flag shown. v0.48.0

04 **Brave and Tavily search adapters for any-search** Not checked yet

60

BREAKING

Adds Brave and Tavily search adapters to the `any-search` integration. The Tavily adapter is marked BREAKING — existing `any-search` configurations may require updates to accommodate the new adapter.

— Names both adapters and the integration, plus the breaking note. v0.48.0

05 **Plain email address requirement in core module** Not checked yet

60

BREAKING

Email addresses must now be a plain address only — the `core` module no longer accepts display-name or angle-bracket formats (e.g. User <user@example.com>), breaking any integration that sends that form.

— Names the module and exact format change with example. v0.48.0

THINNER COVERAGE BELOW

06 **Organization invitation workflows** Not checked yet NEW

55

Create, list, cancel, and resend organization invitations, with automatic email delivery and retry on failure.

— Describes the full workflow but no endpoint or UI path named. v0.48.0

07 **Removal of file_understanding_enabled config field**

Not checked yet

55

BREAKING

The `file_understanding_enabled` config field is removed — configurations referencing it will break on upgrade.

— Names the exact config key and upgrade impact. v0.48.0

08 **OAuth browser-app support via CORS** Not checked yet NEW

45

Browser apps can now exchange OAuth app tokens and call web tools using their own key, enabled via new CORS support.

— Explains mechanism briefly but no config or endpoint named. v0.49.0

09 **Public Providers page** Not checked yet NEW 45

Adds a public Providers page listing each provider's terms and privacy links.

– Names the page but not its exact location or URL. v0.49.0

10 **Gemini decision provider** Not checked yet NEW 40

Adds Gemini as a decision provider alongside existing options.

– Names the provider but no config detail on enabling it. v0.48.0

11 **Bootstrap operator role handover** Not checked yet NEW 35

An administrator can now hand over the bootstrap operator role to another user.

– States capability but no steps or UI location given. v0.49.0

12 **Billing-linked refusals and misspelled model name surfacing** Not checked yet IMPROVED 30

Refused requests are now linked to billing, and misspelled model names are surfaced in the response.

– Thin prose description with no fields or endpoints named. v0.49.0

13 **Otari.ai welcome screen option** Not checked yet NEW 30

The dashboard welcome screen now offers self-hosted operators an otari.ai option.

– Bare description of a UI addition with no further detail. v0.48.0

14 **Self-service account deletion** Not checked yet NEW 25

Users can now delete their own account from the identity settings page.

– Short description, no detail on data handling or confirmation steps. v0.49.0

└──▷ BREAKING ON UPGRADE

! API keys are now scoped to the endpoints they may call — existing keys with unrestricted access will need to be updated.

! The `file_understanding_enabled` config field is removed — configurations referencing it will break on upgrade.

! Service keys that own files now refuse unreachable file IDs — workflows that passed unresolvable IDs will receive errors instead of silent failures.

! The `any-search` Tavily adapter is marked BREAKING — existing `any-search` configurations may require updates to accommodate the new adapter.

! Email addresses must now be a plain address only — the `core` module no longer accepts display-name or angle-bracket formats (e.g. `User <user@example.com>`), breaking any integration that sends that form.

WAS THIS USEFUL?  

OmniRoute

Feedback ⓘ v3.8.54 NOTES ↗

OmniRoute puts hundreds of model providers behind one endpoint, with quota-aware fallback and prompt compression.

OmniRoute adds per-key usage and session APIs, a PII token vault, `jev` and `expiry-first` combo strategies, agentic pipeline tool-loop detection and context compaction, and rolling 5-hour per-key budgets.

└──▷ GET THIS VERSION ▾

Docker

```
$ docker pull ghcr.io/cheaper-inference/omniroute:3.8.54
```

git

```
$ git clone --branch v3.8.54 https://github.com/cheaper-
inference/OmniRoute.git
# already have the repo? check out this version:
$ git checkout v3.8.54
```

published under the project's GitHub account · checked 18 Nov Wrong command?

New `/v1/me/status` API endpoint reports today's and this week's USD, request and token usage, every per-key limit, masked account labels, and cache-first `accountQuotas`; also accepts `x-api-key` and a per-key `anthropicRateLimitHeaders` mode (`auto` / `forward` / `strip`) to control whether pooled Anthropic account-identity and quota headers reach the client. Not checked yet

New `/v1/me/sessions` API endpoint lists and inspects agent sessions with project, client, date, and sorting filters. Not checked yet

New `POST /api/keys/import` endpoint for importing API keys. Not checked yet

New `POST /api/intelligence/sync` endpoint (with `{"rebuildTier": true}`) rebuilds the `models_dev_tier` intelligence source on demand; `GET /api/intelligence/sync` now reports per-source health in `sourcesDetail`; `POST` / `DELETE /api/intelligence/overrides` manage `user_override` entries. Not checked yet

Adds a new `jev` combo routing strategy — a TypeSafe System One `choice` call admits the winner plus up to 4 fallbacks at or above 0.15 confidence; enabled via the `COMBO_JEV_STRATEGY_ENABLED` feature flag (default off). Not checked yet

+13 more

NEW INTEGRATION New provider integrations

NEW INTEGRATION Keenable and codex-search providers

NEW SUPPORTED PLATFORM European Portuguese (PT-PT) locale

WAS THIS USEFUL? 👍 👎

◆ ML Platforms & SDKs

Hugging Face

Feedback ⓘ Inference Providers Powered by OpenRouter

NOTES ↗ RANK

Hugging Face is an open platform for building, sharing, and deploying machine-learning models, datasets, and applications.

> Hugging Face rolled out OpenRouter-backed inference routing for language models, letting users pick providers via model ID suffixes, alongside new provider visibility and control tools — though custom provider keys for language models were broken in the process.

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

01 **OpenRouter-backed inference routing via model ID suffixes**

Not checked yet NEW 80

Language model inference requests can now be routed through OpenRouter with automatic fallback by appending `:fastest`, `:cheapest`, or a named provider such as `:novita` to the model ID.

— Names exact suffix syntax and mechanism, but no runnable command example

Inference Providers Powered by OpenRout…

02 **Provider visibility and control for inference** Not checked yet

NEW 60

Model pages gained two new tabs — Providers and Metrics — in the inference widget showing per-provider price, speed, and uptime, and users can disable individual providers for their account via Settings > Inference Providers.

Disable a specific provider for your account so it is never used in automatic routing.

📍 Go to your Hugging Face account Settings › Inference Providers and toggle off the provider you want to exclude.

— Names both tabs and gives exact settings navigation path

Inference Providers Powered by OpenRout…

THINNER COVERAGE BELOW

03 **Custom provider keys broken for language models** Not checked yet

25

BREAKING

Custom provider keys no longer work for language models as part of this change.

— Bare statement with no migration guidance or scope detail

Inference Providers Powered by OpenRout…

└──▷ **BREAKING ON UPGRADE**

! Custom provider keys no longer work for language models.

WAS THIS USEFUL? 👍 👎

OpenAI Python SDK

Feedback ⓘ v3.54.0 NOTES ↗ RANK

The OpenAI Python SDK provides typed Python clients for OpenAI model APIs, streaming, and asynchronous requests.

> The OpenAI Python SDK added typed models for Managed Agents budget events, new GPT-6.2 model identifiers with longer background windows for the Responses API, and changes to how Agents session errors and schedule pause reasons are surfaced.

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

01 **New GPT-6.2 model identifiers and background windows**

Not checked yet NEW 73

Adds the `GPT-6.2 Astra`, `GPT-6.2 Astra Law`, and `GPT-6.2 Sol` model identifiers usable in Responses and Chat Completions calls (Astra Law targeted at law-domain Responses calls). Also adds longer-running Responses background windows, settable by passing `'2h'` or `'8h'` as the background window value.

— Names exact model IDs and parameter values, usable directly v3.53.0

THINNER COVERAGE BELOW

02 **Agents session errors and schedule pause reasons** Not checked yet

IMPROVED 52

In the beta Agents API, session errors are now delivered as stream events rather than raised exceptions, allowing streaming code to handle them inline. Separately, schedule pause reasons are now surfaced in the API response, letting callers inspect why a scheduled run was paused.

— Describes behavior change but no field/endpoint names given v3.52.0

03 **Typed Managed Agents budget event models** Not checked yet NEW

35

Adds typed Python models for Managed Agents budget events, giving code that processes agent spending events structured type definitions and IDE autocompletion.

— Names the feature but no mechanism or usage detail v3.54.0

WAS THIS USEFUL? 👍 👎

OpenAI TypeScript SDK

Feedback ⓘ v7.59.0 NOTES ↗

RANK

The OpenAI TypeScript SDK provides typed JavaScript and TypeScript clients for OpenAI APIs and streaming responses.

> The OpenAI TypeScript SDK added typed budget events for Managed Agents, three new GPT-6.2 model identifiers, longer background processing windows for the Responses API, and new fields for inspecting paused fine-tuning or evaluation schedules.

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

01 **New GPT-6.2 model identifiers** Not checked yet NEW 65

Three new model identifiers are available for Responses workloads: `gpt-6.2-astra` for general-purpose GPT-6.2 Astra workloads, `gpt-6.2-astra-law` for legal-domain Responses workloads, and `gpt-6.2-sol` for GPT-6.2 Sol workloads.

— Exact model identifiers named, directly usable in code v7.58.0

THINNER COVERAGE BELOW

02 **Longer Responses background windows** Not checked yet NEW 40

Responses API calls that need more processing time can now use 2-hour and 8-hour background windows.

— Durations named but no config key or flag given v7.58.0

03 **Schedule pause reason fields** Not checked yet NEW 30

New fields let callers inspect why a fine-tuning or evaluation schedule was paused.

— Described only in prose, no field names given v7.58.0

04 **Typed Managed Agents budget events** Not checked yet NEW 30

TypeScript consumers now get typed budget events for Managed Agents, providing compile-time visibility into agent budget lifecycle data.

— No type or endpoint name given, just prose description v7.59.0

WAS THIS USEFUL? 👍 👎

Anthropic Python SDK

Feedback ⓘ v2.10.0 NOTES ↗ RANK

The Anthropic Python SDK provides typed Python clients for Claude APIs, streaming responses, and tool use.

> This window adds beta billing policy support with new billing error codes and webhook events for external key failure/recovery, introduces an Update Bundle method for Claude Tag Bundles, and ships several breaking changes around beta header handling and the C# SDK.

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

01 **Deprecated beta header values** Not checked yet DEPRECATED

80

The beta header values `prompt-caching-2024-07-31`, `pdfs-2024-09-25`, `token-counting-2024-11-01`, `token-efficient-tools-2025-02-19`, `extended-cache-ttl-2025-04-11`, `cache-diagnosis-2026-04-07`, and `ce-user-management-2026-07-13` are deprecated and will be removed in a later release — remove them from requests now.

— Seven exact header values named with migration action v2.9.0

02 **Batch requests default header behavior change** Not checked yet

75

BREAKING

`beta.messages.batches` no longer sends `message-batches-2024-09-24`; if your client sets a default `anthropic-beta` header, batch requests now send that header instead of the old value being replaced.

— Names exact header values and affected method v2.9.0

03 **Billing policy support and billing error codes** Not checked yet

NEW

65

Message requests can now pass billing policy details via the beta `billing_policy` request parameter, and message calls return new billing-related error codes: `contract_spend_limit_reached`, `linked billing refusal`, and `user_profile_billing_not_authorized`.

— Names the parameter and error codes but no usage example v2.10.0

04 **C# AnthropicBeta enum renumbering** Not checked yet BREAKING

60

In C#, the `AnthropicBeta` enum's members have new numbers, so code compiled against an earlier SDK version silently sends a different beta value with no error until you rebuild it.

— Explains silent failure mode but no specific enum values v2.9.0

THINNER COVERAGE BELOW

05 **Update Bundle method for Claude Tag Bundles** Not checked yet NEW

50

Adds an Update Bundle method that sets a Claude Tag Bundle's instructions and returns them on the Bundle object.

— Named method but no endpoint or code example v2.9.0

06 **Webhook events for external key failure and recovery**

Not checked yet NEW

25

The SDK now delivers webhook events for external key failure and recovery scenarios.

— No event names, payload, or schema given v2.10.0

└──▷ **BREAKING ON UPGRADE**

! In C#, the `AnthropicBeta` enum's members have new numbers, so code compiled against an earlier SDK version sends a different beta value with no error until you rebuild it.

! The beta header values `prompt-caching-2024-07-31`, `pdfs-2024-09-25`, `token-counting-2024-11-01`, `token-efficient-tools-2025-02-19`, `extended-cache-ttl-2025-04-11`, `cache-diagnosis-2026-04-07`, and `ce-user-management-2026-07-13` are deprecated and will be removed in a later release — delete them from your requests now.

! `beta.messages.batches` no longer sends `message-batches-2024-09-24`; if your client sets a default `anthropic-beta` header, batch requests now send that header instead of the old value being replaced.

WAS THIS USEFUL? 👍 👎

Anthropic TypeScript SDK

Feedback ⓘ sdk-v1.10.0

NOTES ↗ RANK

The Anthropic TypeScript SDK provides typed JavaScript and TypeScript clients for Claude APIs, streaming, and tool use.

> The Anthropic TypeScript SDK added beta billing controls and new billing error codes, richer RBAC role metadata, a Tag Bundle update method, and deprecated a batch of legacy beta headers while changing how the batches endpoint sends its beta header.

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

01 **Deprecated beta header values** Not checked yet DEPRECATED

85

The beta header values `prompt-caching-2024-07-31`, `pdfs-2024-09-25`, `token-counting-2024-11-01`, `token-efficient-tools-2025-02-19`, `extended-cache-ttl-2025-04-11`, `cache-diagnosis-2026-04-07`, and `ce-user-management-2026-07-13` are deprecated and will be removed from the SDKs in a later release; delete them from requests now. `files-api-2025-04-14` and `skills-2025-10-02` must follow their respective migration guides before removal.

– Lists every deprecated header with migration instruction sdk-v1.9.0

02 **Batches endpoint beta header behavior change** Not checked yet

80

BREAKING

`beta.messages.batches` no longer sends the `message-batches-2024-09-24` beta header; if your client sets a default `anthropic-beta` header, batch requests now send that default where some SDK versions previously replaced it with the old value.

– Exact header names and before/after behavior given sdk-v1.9.0

THINNER COVERAGE BELOW

03 **New billing error codes for messages** Not checked yet NEW

55

Adds the `contract_spend_limit_reached`, `add_linked_billing_refusal`, and `user_profile_billing_not_authorized` error codes to messages.

– Exact codes named but trigger conditions unclear sdk-v1.10.0

04 **Update Bundle method for Claude Tag Bundles** Not checked yet NEW

55

Adds an Update Bundle method to set a Claude Tag Bundle's instructions and return them on Bundle responses.

— Describes purpose but not exact method signature `sdk-v1.9.0`

05 **Beta `billing_policy` parameter on messages** Not checked yet NEW

45

Adds beta support for the `billing_policy` request parameter on the Messages API.

— Names the param but not its accepted values or effect `sdk-v1.10.0`

06 **Source and error code fields on RBAC Role reads** Not checked yet

IMPROVED

45

Adds `source` and `error codes` fields to the RBAC Role read endpoints, exposing richer role metadata.

— Names fields but not the exact endpoint path `sdk-v1.9.0`

└──▷ **BREAKING ON UPGRADE**

! In C#, the `AnthropicBeta` enum's members have new numbers, so code compiled against an earlier SDK version sends a different beta value with no error until you rebuild it.

! `beta.messages.batches` no longer sends the `message-batches-2024-09-24` beta header; if your client sets a default `anthropic-beta` header, batch requests now send that default where some SDK versions previously replaced it with the old value.

! The beta header values `prompt-caching-2024-07-31`, `pdfs-2024-09-25`, `token-counting-2024-11-01`, `token-efficient-tools-2025-02-19`, `extended-cache-ttl-2025-04-11`, `cache-diagnosis-2026-04-07`, and `ce-user-management-2026-07-13` are deprecated and will be removed from the SDKs in a later release; delete them from your requests now. `files-api-2025-04-14` and `skills-2025-10-02` must follow their respective migration guides before removal.

WAS THIS USEFUL? 👍 👎

SkyPilot

Feedback ⓘ v0.14.3 NOTES ↗ RANK

SkyPilot runs AI training and inference jobs across clouds and Kubernetes clusters, picking cheaper or available capacity and handling spot preemption.

> SkyPilot's v0.14.3 release adds machine-readable output for managed job launches, dependencies and runtime-owned recovery for managed jobs, Nebius spot pricing, and passwordless ACR authentication via Azure VM managed identity, alongside Slurm permission and monitoring refinements.

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

01 **Machine-readable output for `sky jobs launch`** Not checked yet

NEW

83

`sky jobs launch` now accepts `--output id` or `--output json` to return a machine-readable job ID or the full JSON job record directly from the CLI, making it easy to capture job IDs or full results for scripts and downstream pipelines.

Capture a newly launched managed job's ID in a script so you can poll or depend on it programmatically.

```
$ sky jobs launch --output id my_task.yaml
```

Get the full JSON record of a launched job for downstream processing or logging pipelines.

```
$ sky jobs launch --output json my_task.yaml
```

– Exact flag and values with runnable example commands v0.14.3

THINNER COVERAGE BELOW

02 **ACR authentication via VM managed identity** Not checked yet NEW

50

SkyPilot can now authenticate to Azure Container Registry (ACR) using the VM's managed identity, removing the need for stored credentials.

– Names mechanism and benefit but no config key or flag v0.14.3

03 **Sudo permission controls for Slurm clusters** Not checked yet NEW

45

SkyPilot adds executable sudo allowlists for user impersonation on Slurm clusters, and supports reduced sudo permissions that no longer require container snapshots.

— Two mechanisms described but no flags or config named v0.14.3

04 **Slurm GPU visibility in Helm compute overview** Not checked yet

IMPROVED

40

The Helm compute overview now includes Slurm GPU allocation, showing 0% when the cluster is idle.

— Names dashboard and behaviour but no exact path given v0.14.3

05 **Dependencies between managed jobs** Not checked yet NEW

35

Managed jobs can now declare dependencies on other managed jobs, so one job waits for another to finish before it starts.

— No config key or command shown for declaring a dependency v0.14.3

06 **LiteLLM proxy metrics federation** Not checked yet NEW

35

LiteLLM proxy metrics are now federated so they are collected alongside other cluster metrics.

— Names integration but no endpoint or config surface v0.14.3

07 **Runtime-owned recovery for managed jobs** Not checked yet NEW

30

Managed jobs now support multiple runtimes and runtime-owned recovery, where the runtime itself handles recovery for a job rather than relying solely on SkyPilot's controller.

— Mechanism named briefly, no flags or config given v0.14.3

08 **Nebius spot pricing support** Not checked yet NEW

30

SkyPilot adds support for Nebius spot pricing, enabling cost-optimised workloads on Nebius infrastructure.

— Names provider but no pricing mechanism or command v0.14.3

09 **Stalled managed job reporting** Not checked yet NEW

20

SkyPilot now reports managed job tasks that stalled before they ever ran.

— Bare description with no detection mechanism or command v0.14.3

WAS THIS USEFUL? 👍 👎

Google Gen AI Python SDK Feedback ⓘ v3.4.0 NOTES ↗

RANK

The Google Gen AI Python SDK provides Python interfaces to Gemini model generation, streaming, tools, and embeddings.

> Google Gen AI Python SDK v3.4.0 adds tool search controls and static file processing options to the Interactions SDK, plus batch embedding support for the Gemini Enterprise API.

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

01 **Interactions SDK tool search and file options** Not checked yet

NEW 50

The Interactions SDK gains tool search types for controlling how tools are discovered during agent interactions, and static file processing options exposed in `google.genai.interactions` for controlling how files are handled in interaction flows.

— Names module and capability but no exact flags or signatures v3.4.0

02 **Batch embedding inputs for Gemini Enterprise API** Not checked yet

NEW 40

Gemini embedding models on the Gemini Enterprise API now support batch inputs and usage metadata.

— States capability but no API signature or usage example v3.4.0

WAS THIS USEFUL? 👍 👎

Google Gen AI JavaScript SDK

Feedback ⓘ v3.4.0

NOTES ↗ RANK

The Google Gen AI JavaScript SDK provides TypeScript and JavaScript interfaces to Gemini APIs and live model sessions.

> Google Gen AI JavaScript SDK adds batch embedding support for Gemini Enterprise and expands the Interactions SDK with tool search types and static file processing options.

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

01 **Batch embeddings for Gemini Enterprise API** Not checked yet NEW

65

Supports batch inputs and usage metadata for Gemini embedding models by calling `ai.batches.createEmbeddings` on the Gemini Enterprise API.

– Names exact method but no further mechanism detail. v3.4.0

THINNER COVERAGE BELOW

02 **Tool search types and static file options in Interactions SDK**

Not checked yet NEW

30

Adds tool search types to the Interactions SDK and exposes static file processing options in the Interactions SDK.

– Bare mentions with no named flags or mechanism. v3.4.0

WAS THIS USEFUL? 👍 👎

AWS Amazon Bedrock AgentCore

Feedback ⓘ

Web search: Free monthly queries (2026-11-16) BLOG ↗ RANK

Amazon Bedrock AgentCore provides managed agent runtimes, memory, identity, gateways, observability, and evaluation services.

> Amazon Bedrock AgentCore now includes a monthly allotment of free web search queries shared across its Web Search Tool and API.

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

01 **Free monthly web search queries** Not checked yet NEW 45

Amazon Bedrock AgentCore now includes 5,000 free web search queries each month per AWS account, shared between the Web Search Tool and the Web Search API.

— Names exact quota and surfaces but no setup detail

Web search: Free monthly queries (2026-…

WAS THIS USEFUL? 👍 👎

Snowflake Cortex AI

Feedback ⓘ

Nov 16, 2026: Create, evaluate, and optimize Cortex AI Functions with Cortex AI Function

NOTES ↗ RANK

Snowflake Cortex AI provides managed language-model functions, search, and agents that work with data governed in Snowflake.

> Snowflake Cortex AI reached general availability for Cortex AI Function Studio, a new workspace for building, benchmarking, and optimizing AI Functions directly from natural-language task descriptions.

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

01 **Cortex AI Function Studio for building and optimizing AI Functions**

Not checked yet NEW 75

Cortex AI Function Studio lets users build Cortex AI Functions by describing a task in natural language through the Cortex Code CLI or Snowsight AI Studio. It automatically benchmarks and optimizes prompts, model selection, and workflow strategies using configurable evaluation metrics and managed optimization workflows, and supports multimodal AI Functions covering text, documents, images, audio, and video based on supported model capabilities.

When analysts want to build and evaluate an AI Function through a guided no-code experience, navigate to AI Studio in Snowsight.

> 📍 In Snowsight, go to AI Studio › AI Function Studio › New Function, describe your task in the text field, then select Optimize to run automated benchmarking across prompts and models.

— Names CLI/UI entry points and optimization mechanism, with step-by-step navigation example.

Nov 16, 2026: Create, evaluate, and opt…

WAS THIS USEFUL? 👍 👎

◆ Inference Serving & Optimization

NVIDIA Triton Inference Server

Feedback ⓘ v2.75.0

NOTES ↗ RANK

Triton Inference Server is an open-source platform for deploying and serving AI models at scale across CPUs and GPUs.

> Triton Inference Server v2.75.0 adds BF16 input/output support for AOTI in the PyTorch backend and now ships third-party license notices in containers.

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

01 **AOTI BF16 input/output in PyTorch backend** Not checked yet

IMPROVED 40

The PyTorch backend now supports AOTI BF16 input and output tensors.

— Names backend and data type but no usage detail v2.75.0

WAS THIS USEFUL?  

Experiential Labs Experiential

Feedback ⓘ

v0.7.250 NOTES ↗ RANK

Experiential is the open source, zero markup gateway for BYOK, self-hosted and 1000+ marketplace models. It learns from your traffic to cut costs, recommend better models, and train a specialized model you own.

> Experiential shipped async video generation, more granular reasoning-token reporting across Azure AI Foundry and thinking-capable models, fine-grained thinking controls for GLM models on Vertex and Z.ai, plus a tunable input-moderation wait, an org-level prompt-caching switch, and clearer retry signaling for unavailable routes.

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

01 **Thinking controls for GLM models on Vertex and Z.ai**

Not checked yet NEW 95

Turn off thinking for Vertex MaaS GLM models by passing `chat_template_kwargs {enable_thinking: false}`, working around Vertex's rejection of `reasoning_effort: none` and its ignoring of `thinking {type: disabled}`. Z.ai GLM-5 and GLM-5.1 now use Z.ai's own thinking switch to disable reasoning, while GLM-5.2 and later use the documented `reasoning_effort`.

Disable thinking for a Vertex MaaS GLM model when `reasoning_effort: none` is rejected by Vertex.

```
$ curl -X POST https://<gateway>/v1/chat/completions -H
'Authorization: Bearer <token>' -H 'Content-Type:
application/json' -d '{"model": "vertex-glm",
"chat_template_kwargs": {"enable_thinking": false}, "messages":
[{"role": "user", "content": "Summarise this document."}]}'
```

— Exact params named with a runnable example request. v0.7.198

02 **Reasoning-token reporting across Azure and estimate-billed models**

Not checked yet NEW INTEGRATION IMPROVED 80

Experiential's gateway now reads the top-level Chat reasoning count that Azure AI Foundry reports, surfacing it alongside other token metrics. It also surfaces estimated reasoning-token counts in `completion_tokens_details.reasoning_tokens` (Chat) and `output_tokens_details` (Responses) for lanes that bill thinking without

counting it, such as Claude, with these estimates flagged via `estimated_token_details`.

— Named response fields fully, but no example call shown. v0.7.198 v0.7.181

03 **Retryable 503 for unavailable routes** Not checked yet IMPROVED

70

An unavailable route at admission now escalates as a retryable `503` `unavailable_route` with a `Retry-After: 2` header instead of an anonymous `500`.

— Exact status code and header named with before/after behavior. v0.7.198

04 **Async video generation via /v1/videos** Not checked yet NEW

60

Access async video generation through the `/v1/videos` jobs endpoint.

— Endpoint path named but no example request given. v0.7.198

THINNER COVERAGE BELOW

05 **Configurable wait for parallel input moderation** Not checked yet

NEW

57

Experiential adds an explicit additional wait for parallel input moderation, configurable between 0 and 5000 ms, with a default of 250 ms.

— Numeric range and default given, but no config key named. v0.7.250

06 **Org-level prompt-caching switch at admission** Not checked yet

NEW

35

Enable or disable prompt caching at the organization level via a switch at native gateway admission.

— No config key or endpoint named for the switch. v0.7.181

v0.7.184 NOTES ↗ STABLE

Experiential adds thinking-level controls (none, low, medium, high) for Gemini 4 Argon in the gateway, with `none` transmitted as `thinkingBudget 0`.

└──▷ GET THIS VERSION ▾

```
$ git clone --branch v0.7.184
https://github.com/experientiallabs/experiential.git
# already have the repo? check out this version:
$ git checkout v0.7.184
```

Control Gemini 4 Argon reasoning effort in the gateway by selecting a thinking level — none, low, medium, or high — where `none` is sent as `thinkingBudget 0`

Not checked yet

WAS THIS USEFUL? 👍 👎

vLLM

Feedback ⓘ v0.34.0 NOTES ↗ CODE ↗

The vLLM inference engine runs large language models with high-throughput, memory-efficient GPU execution.

vLLM adds checkpoint/restore endpoints, extensible KV cache on by default, fault-tolerant DP+EP scale-down, Rust frontend Responses API, suffix decoding on MRV2, and new speculative decoding plugin API.

└──▷ GET THIS VERSION ▾

pipx

```
$ pipx install vllm==0.34.0
# already installed? upgrade to this version:
$ pipx install --force vllm==0.34.0
```

git

```
$ git clone --branch v0.34.0 https://github.com/vllm-project/vllm.git
# already have the repo? check out this version:
$ git checkout v0.34.0
```

named in the project's README · checked 18 Nov Wrong command?

Enable checkpoint and restore of live serving instances using the new `/checkpoint/prepare`, `/checkpoint/status`, and `/checkpoint/restore` endpoints, available in both the Python and Rust frontends. Not checked yet

Opt out of the new extensible KV cache (now on by default) with `--gpu-memory-utilization 0.9` or `--no-enable-extensible-kv-cache` to preserve old memory sizing behaviour. Not checked yet

Register custom speculative decoding methods via the new `register_speculative_method` plugin API, with declarative configs. Not checked yet

vLLM now adds suffix decoding on Model Runner V2 with a GPU speculator, removing the dependency on Arctic Inference. Not checked yet

vLLM now adds fault-tolerant DP+EP scale-down: dead ranks are removed and their experts redistributed without a restart. Not checked yet

+7 more

NEW SUPPORTED PLATFORM Mistral Large 4

NEW SUPPORTED PLATFORM NVIDIA Nemotron-3.5-Transcribe

NEW SUPPORTED PLATFORM Ricursive Ripple-65B Preview

NEW SUPPORTED PLATFORM Zyphra ZAYA1

NEW SUPPORTED PLATFORM Baidu Unlimited-OCR-2

NEW SUPPORTED PLATFORM Liquid AI LFM3, LFM3-VL, LFM3-Audio

NEW SUPPORTED PLATFORM NVIDIA Canary-1B-v2

NEW SUPPORTED PLATFORM Ming3-ASR

NEW SUPPORTED PLATFORM NVIDIA Parakeet TDT

UPDATED SUPPORTED PLATFORM ROCm support updated to ROCm 10.0 only

RETIRED SUPPORTED PLATFORM ROCm 7.2 artifacts dropped

└──▷ **BREAKING ON UPGRADE**

- ! Model Runner V1 is removed; `VLLM_USE_V2_MODEL_RUNNER=0` now fails at startup. V1 logits processors must be ported to `vllm.v1.worker.gpu.sample.logits_processor`.
- ! Default `--block-size` changed from 16 to 64; pass `--block-size 16` to keep the old behaviour.
- ! Extensible KV cache is on by default and `gpu_memory_utilization` defaults to 1.0; pass `--gpu-memory-utilization 0.9` or `--no-enable-extensible-kv-cache` to restore old sizing.
- ! `huggingface_hub >= 2.0.0` is now required, along with `tokenizers >= 0.23.3` and `timm >= 1.0.30`.
- ! QuTLASS / `fp_quant` is removed; `fp_quant` checkpoints no longer load.
- ! Marlin W4A8 is removed; use `--quantization humming` with `VLLM_HUMMING_INPUT_QUANT_CONFIG` instead of `VLLM_MARLIN_INPUT_DTYPE`.
- ! Keye-VL and Keye-VL-1.5 model support is removed.

WAS THIS USEFUL? 👍 👎

Niko1221 Strata

Feedback ⓘ v0.2.14 NOTES ↗ RANK

Strata is a local inference engine with one-click install for Windows and Linux, providing an OpenAI/Anthropic-compatible API on localhost for models like Qwen and DeepSeek.

> Strata's v0.2.14 release replaces llama.cpp's prompt-reading path with new Strata-native kernels for NVIDIA and AMD GPUs by default, and tunes prompt chunking and multi-card memory layout for faster processing.

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

01 Strata-native prompt reading kernels for NVIDIA/AMD

Not checked yet NEW 85

Strata now reads prompts using its own kernels on NVIDIA and AMD cards on Windows and Linux, replacing the llama.cpp path for that step as the new default. The old llama.cpp prompt-reading path can still be enabled by building with `-DSTRATA_MMQ=ON`.

Build Strata from source using the legacy llama.cpp prompt-reading kernels instead of the new Strata-native ones, for comparison or rollback.

```
$ cmake -DSTRATA_MMQ=ON <path>
```

— Names exact build flag and platforms; runnable command given v0.2.14

THINNER COVERAGE BELOW

02 Two-card memory layout optimization for drafter head

Not checked yet IMPROVED 40

Two-card configurations now keep only each card's own layers' dense weights and leave room for the drafter's head on the last card, increasing expert slot availability.

— Describes specific memory layout change but no way to invoke it directly v0.2.14

03 Equal-sized prompt chunking for single-card processing

Not checked yet IMPROVED 35

Long prompts on a single card are now split into equal-sized pieces instead of full pieces plus a short remainder, reducing prompt processing time for mid-length contexts.

— Explains before/after behavior but no config or command to use it v0.2.14

WAS THIS USEFUL?  

◆ Local LLM Runtimes

Mintplex Labs AnythingLLM

Feedback ⓘ v1.20.0

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's v1.20.0 release adds identity forwarding, timeout control, and a management API for MCP servers, brings the browser extension to full chat parity, rebuilds the agent PowerPoint tool, and lets Gmail/Outlook email approvals be edited inline.

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

01 MCP server configuration and management API

Not checked yet NEW

92

Per-server config keys let you opt in to forwarding workspace and user identity to an MCP server with `anythingllm.forwardIdentity`, and cap how long its tool calls may run with `anythingllm.timeout` (seconds, default 60). New `/v1/mcp-servers` API endpoints let you list, add, update, delete, start, stop, and reload MCP servers, and toggle individual tools on or off.

Prevent an MCP server that makes slow external calls from hanging indefinitely by capping its tool call runtime.

json

```
{
  "mcpServers": {
    "my-slow-server": {
      "command": "npx",
      "args": ["my-slow-mcp"],
      "anythingllm": {
        "timeout": 120
      }
    }
  }
}
```

Pass the calling workspace and user through to an MCP server so it can apply per-user access controls or logging.

```json
{
  "mcpServers": {
    "my-audited-server": {
      "command": "npx",
      "args": ["my-audited-mcp"],
      "anythingllm": {
        "forwardIdentity": true
      }
    }
  }
}
```

— Names exact config keys, default value, and full endpoint capability set v1.20.0

02 **Metrics endpoint opt-out** Not checked yet NEW 75

Set the `DISABLE_METRICS_ENDPOINT` environment variable to disable the `/api/utils/metrics` endpoint, for deployments that should not expose telemetry.

Stop AnythingLLM from exposing the metrics endpoint in a production deployment where you do not want telemetry accessible.

```env
DISABLE_METRICS_ENDPOINT=true
```

— Exact env var and endpoint given, but no further mechanism detail v1.20.0

03 **Editable email approval cards** Not checked yet IMPROVED 65

Email approval cards for the Gmail and Outlook agent tools now let you edit recipients, Cc, Bcc, subject, and body inline before approving, with the approval timer paused during editing.

— Names exact editable fields and timer behaviour v1.20.0

04 **Full chat sidebar in browser extension** Not checked yet NEW 60

The browser extension now works as a full chat sidebar in Chrome, with right-click Add context for text and images and direct chat, reaching feature parity with the traditional chat UI.

— Describes behaviour and parity claim but no config or API surface v1.20.0

THINNER COVERAGE BELOW

05 **Rebuilt PowerPoint agent tool** Not checked yet IMPROVED

55

The agent PowerPoint tool now produces presentations with color themes, content-aware slide layouts that vary from slide to slide, and live slide previews shown in chat as the deck is built.

— Describes new output qualities but no invocation detail v1.20.0

06 **DeepSeek vision and web search support** Not checked yet NEW

35

DeepSeek now supports vision and can be used as an agent web search provider.

— Bare statement with no mechanism or configuration detail v1.20.0

WAS THIS USEFUL? 👍 👎

mlxtop

Feedback ⓘ v3.4.0 NOTES ↗ RANK

The mlxtop terminal UI monitors local LLMs on Apple Silicon, showing loaded models, memory use, GPU activity, and generation speed.

> mlxtop v3.4.0 adds machine-readable output modes (JSON snapshots, streaming, and a Prometheus server), plus deeper diagnostics for decode efficiency, slowdown causes, thermal pressure, and cross-session speed regressions.

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

01 JSON, streaming, and Prometheus output modes

Not checked yet

NEW

95

`mlxtop --once --json` gives a one-line JSON snapshot of the full static report, `mlxtop stream` emits a continuous stream of one JSON object per sample, and `mlxtop serve` exposes Prometheus metrics plus `/snapshot` and `/health` endpoints on `127.0.0.1:10063` (loopback only).

Scrape mlxtop's Prometheus endpoint from another machine via an SSH tunnel — useful when the inference host has no open ports.

```
$ ssh -L 10063:127.0.0.1:10063 user@inference-host 'mlxtop serve'
&
curl http://127.0.0.1:10063/metrics
```

Stream one JSON object per sample to a log file for offline analysis of a long inference run.

```
$ mlxtop stream >> ~/inference-session.jsonl
```

— Names exact flags, subcommands, endpoints, and port with runnable examples v3.4.0

02 Doctor view of model internals

Not checked yet NEW

80

Press `d` or run `mlxtop doctor` to see the weights read in one forward pass, bits per weight, quantization, KV bytes per token and format, context allowed and trained, draft details, and runtime version.

— Named key and command with concrete fields shown, no sample output v3.4.0

03 Replay-based slowdown attribution

Not checked yet NEW

70

`mlxtop replay` replays a recording and tallies each slowdown reason — slow samples, episodes, episodes per hour, cache misses, model loads, full KV caches, and overlap.

— Named command and full list of tallied reasons, no usage example v3.4.0

---

04 **Decode efficiency against memory bandwidth** Not checked yet NEW

70

mlxtop now reads the chip's peak memory bandwidth and estimates how close decode runs to that limit per verify pass, shown on the runtime card and in `--once` output as `DECODE` / `DRAFT` lines.

— Explains mechanism and output location but no runnable example v3.4.0

---

05 **Cross-session generation speed baselines** Not checked yet NEW

60

mlxtop now keeps each model's usual generation speed by context across sessions in a private file, naming regressions like 'This model now generates 31% below its usual from earlier sessions at this depth'.

— Describes behaviour and sample message but no config surface named v3.4.0

---

06 **In-flight request token headroom in memory panel** Not checked yet

IMPROVED

60

The memory panel now shows tokens each in-flight request can still add (for example `≈37k tokens at 2 requests`), sourced from the runtime's KV room, configured window, or remaining MLX memory.

— Concrete example value given but it's a passive UI display only v3.4.0

---

07 **Journal event attribution for first-token latency** Not checked yet

IMPROVED

60

The Journal now names `CACHE MISS`, `MODEL LOAD`, `KV CACHE FULL`, and `OVERLAP` events (noting `max_concurrent_requests`) to explain first-token latency, with time lost added to the session summary.

— Named event labels and config field, but just a display feature v3.4.0

---

08 **macOS thermal pressure monitoring** Not checked yet NEW

60

mlxtop now reads macOS thermal pressure level via `notifyutil` every 5 s and shows THERMAL heavy on the card and Thermal limit in the verdict from `heavy` up on Apple Silicon.

— Names mechanism, interval, and threshold but no interaction surface v3.4.0

THINNER COVERAGE BELOW

09 **Expanded log line fields** Not checked yet IMPROVED 45

Log lines now end with `pressure=`, `thermal=`, `power_mw=`, `draft=`, and `ts=` fields; recorded samples carry `ts`.

— Lists fields verbatim but no explanation of use v3.4.0

WAS THIS USEFUL? 👍 👎

LM Studio

Feedback ⓘ Introducing Agent Remote Control (2026-11-16)

BLOG ↗ RANK

LM Studio is a desktop application that runs local LLMs and provides model discovery, downloads, chat, and a local API server.

> LM Studio introduced Agent Remote Control, letting Bionic agent sessions be hosted and managed across laptops, dev boxes, and cloud VMs from any device over an encrypted LM Link network.

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

01 Remote control of Bionic agent sessions

Not checked yet NEW

88

LM Studio now supports desktop-to-desktop remote control of Bionic agent sessions, encrypted end-to-end over LM Link (built with Tailscale). Enable remote hosting on a machine via the Remote control button in the sidebar and toggling on Allow remote agent control; manage remote machines from the controlling device under `Settings > Devices > Remote Machines`. For display-less machines, Bionic can run headless using `llmster`, a new headless daemon.

Turn a headless cloud VM into a remote agent host so you can offload long-running tasks without keeping your laptop open.

> 📍 Install Bionic on the VM, then run: llmster On the VM (via SSH or remote session), enable remote control: click Remote control in the sidebar and turn on Allow remote agent control. On your laptop, go to Settings › Devices › Remote Machines and add the VM to your device network.

— Names exact UI toggles, settings path, and the llmster daemon command

Introducing Agent Remote Control (2026-…

02 Mid-session task handoff via @mention

Not checked yet NEW

68

Users can hand off work to another machine mid-session by typing `@` in the composer and picking a device — Bionic starts a session there and waits for it.

Hand off a long build-and-test job to a dedicated Windows machine mid-session while staying on your Mac.

> 📍 In the Bionic composer on your Mac, type @ and select your Windows machine. Bionic opens a new session there and waits for it, while showing you which session

dispatched the work on the remote machine.

— Exact composer interaction described, concrete steps

Introducing Agent Remote Control (2026-…

THINNER COVERAGE BELOW

03 **Remote project browsing and device selection** Not checked yet

NEW

57

Browse remote projects and chats from the sidebar under Remote machines, or hover over a machine to start a new project on it. At project creation time, users can choose which device a project runs on, with all machines in the device network listed along with online status.

— Describes UI navigation but no deeper mechanism

Introducing Agent Remote Control (2026-…

WAS THIS USEFUL? 👍 👎

◆ AI Chat Apps

Open WebUI

Feedback ⓘ v0.13.4 NOTES ↗ RANK

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.

> Open WebUI v0.13.4 introduces Tether-based bridging for local MCP servers and sandboxed per-user computers, adds video generation and multi-connection image generation, brings interactive visualizations and new coding agents into chat, and gives side chats full chat parity, while replacing Direct Connections with Tether.

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

01 Sandboxed per-user computers via Terminals

Not checked yet NEW

91

When Terminals 0.4.0 or later is connected, Open WebUI can provision sandboxed computers on demand for users, chats or models, configurable with image, CPU, memory, storage, network profile (defined by admins via `TERMINALS_NETWORK_PROFILES`), environment variables, start directory and setup script.

Grant all new chats a default sandboxed computer without per-model configuration.

📍 In Open WebUI, go to Admin Settings › Models › Model Defaults, select a terminal connection (requires Terminals 0.4.0 or later), then set `TERMINALS_NETWORK_PROFILES` on the Terminals server to define available network profiles for provisioned containers.

— Names config key, version requirement and configurable parameters. v0.13.4

02 Local MCP servers via Tether

Not checked yet NEW

88

Local MCP servers can run over stdio through Tether 0.1.16 or later with no proxy (such as mcpo) and no open port; add one in Tether under Settings > Integrations > Tools with its command, for example `npx -y @modelcontextprotocol/server-filesystem /tmp`, then select it from the + menu > Tools in any chat.

Add a local MCP server that runs as a stdio command through Tether so it appears in the Tools list with no open port.

📍 In the Tether app, go to Settings › Integrations › Tools, add a new entry with the server's command (for example: npx `-y` @modelcontextprotocol/server-filesystem `/tmp`), then pick it from the + menu › Tools in any Open WebUI chat.

— Exact command, version requirement and UI path given. v0.13.4

03 **Terminal sandbox lock via Tether** Not checked yet NEW

80

Under Settings > Open Terminal > Sandbox, Tether 0.1.13 or later can lock the terminal into a macOS or Windows sandbox, confining coding agents to one folder with no network access unless explicitly enabled and no incoming connections.

— Exact UI path and version but no further limits given. v0.13.4

04 **Multiple image generation connections** Not checked yet IMPROVED

78

Admins can add multiple image generation connections under Admin Settings > Images, each an OpenAI-compatible server such as LiteLLM, Gemini, ComfyUI or AUTOMATIC1111 with its own models, sizes, workflows and access; users with more than one image model pick one from the Image entry in the + menu.

— Names four providers and admin path and selection UI. v0.13.4

05 **Visualizations built-in tool in chat** Not checked yet NEW

75

Turning on the Visualizations built-in tool per model under Builtin Tools lets chat responses render interactive charts, diagrams and calculators; administrators can override its instructions with `VISUALIZATION_PROMPT_TEMPLATE` under Admin Settings > Interface > Visualization Prompt.

Let administrators override the system prompt that drives the Visualizations tool across all models from a single environment variable.

📍 In Admin Settings › Interface › Visualization Prompt, paste a custom prompt; or set the environment variable `VISUALIZATION_PROMPT_TEMPLATE` to your prompt text before starting Open WebUI.

— Names config key and exact admin path but lacks limits. v0.13.4

06 **Video generation in chat** Not checked yet NEW

75

Admins turn on Video Generation under Admin Settings > Videos, add OpenAI, LiteLLM or OpenRouter connections with a default video model, and grant the Video Generation permission; users then enable Video from the + menu to generate videos in chat.

— Names providers, admin path and permission but no limits. v0.13.4

07 **Coding agents and default terminal assignment** Not checked yet

NEW

75

Added coding agents Hermes, Antigravity, GitHub Copilot, Goose and Qwen Code as chat models alongside existing agents, with any ACP agent registered via `OPEN_TERMINAL_ACP_AGENTS` also usable once installed and signed in; administrators can also set a default terminal for all new chats without a model-specific terminal under Admin Settings > Models > Model Defaults.

– Names agents and env var but little mechanism depth. v0.13.4

---

08 **Full chat features in side chats** Not checked yet IMPROVED

65

Side chats now run full chat, adding message buttons, sources, token usage, follow-up suggestions, slash commands, queued messages, chat variables, web search settings, tool events, HTML and SVG previews and multi-model answers.

– Lists many added capabilities but no config or command. v0.13.4

THINNER COVERAGE BELOW

09 **Sidebar tabs for Notes, workspace and more** Not checked yet IMPROVED

55

Notes, channels, the workspace, automations, the calendar, the playground, the home page and new chats from pinned models can now be opened as tabs beside the current chat from the sidebar, user menu or in-chat links.

– Lists destinations but no mechanism beyond opening as tabs. v0.13.4

10 **Tether replaces Direct Connections** Not checked yet BREAKING

30

Direct Connections has been replaced by Tether as the mechanism for connecting local resources such as MCP servers and terminals to Open WebUI.

– Only states the replacement, no migration detail. v0.13.4

└──▷ **BREAKING ON UPGRADE**

! Tether replaces Direct Connections.

WAS THIS USEFUL? 👍 👎

LibreChat

Feedback ⓘ v0.8.9 NOTES ↗ RANK

LibreChat provides a self-hosted chat interface with multiple model providers, agents, tools, and retrieval over uploaded files.

> LibreChat v0.8.9's headline addition is Slack connectivity for agents, alongside a new Background Responses API, SDK-backed agent runtimes, and a wave of smaller platform, observability and model-support updates.

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

01 **Agent Channels for Slack** Not checked yet NEW 85

Agents can now be connected to Slack from the Agent Builder, linking organization Slack workspaces to agents with tenant-owned OAuth storage, admin-assigned bot managers, and Slack Block Kit rich message rendering. A new standalone Slack Channel Gateway can be deployed as an optional Docker or Helm service to route agent responses through Slack with linked accounts, approvals, and feedback.

— Names Agent Builder, OAuth storage, Block Kit, Docker/Helm gateway in detail v0.8.9

02 **Background Responses API** Not checked yet NEW 60

The new Background Responses API lets developers execute, persist, resume, review, cancel, and recover background agent responses, with saved results and decision retention.

— Named API with clear behaviors but no endpoint syntax given v0.8.9

03 **SDK-backed agents for native chat** Not checked yet NEW 60

Agents can be registered as SDK-backed for native chat, supporting local or remote runtimes, authorized host tools, and automatic connection recovery without repeating execution.

— Describes runtime options and recovery behavior, no exact config shown v0.8.9

THINNER COVERAGE BELOW

04 **MCP name reservation index migration required** Not checked yet

55

BREAKING

Deployments with durable MCP name reservations must run the targeted index migration: upgrade every writer and temporarily gate reservation/account-erasure writes before replacing the legacy global index; tenantless tombstones must not be deleted during this process.

— Gives concrete migration steps and constraints for operators v0.8.9

05 **New model support added** Not checked yet NEW

45

Added support for Gemini Hydrogen, GPT-6.1 Terra, Claude Fable 5.2, GPT Image 2.6 Sunburst, GPT Image 2.6 Flare, Claude Mythos 5.2, and GPT-6 Cyber models.

— Lists exact model names but no usage mechanism described v0.8.9

06 **Agent identity in provider and MCP headers** Not checked yet NEW

40

Agent identity can now be added to provider and MCP header templates, enabling identity-aware routing in multi-tenant deployments.

— States capability but no header field names or syntax given v0.8.9

07 **Native PostgreSQL tool favorites** Not checked yet NEW

40

Tool favorites are now natively supported in PostgreSQL, so tool preferences persist in Postgres deployments.

— Names Postgres persistence but no schema or key detail v0.8.9

08 **Deployment themes with welcome backdrop** Not checked yet NEW

35

Deployment themes can now apply a welcome backdrop option, configurable per theme.

— Names the option but lacks configuration path specifics v0.8.9

09 **Tool Search defers system tool loading** Not checked yet IMPROVED

30

System tools are now deferred behind Tool Search so the composer only loads tools on demand, reducing initial render cost.

— Brief mechanism note, no config or flag named v0.8.9

10 **OpenTelemetry tool execution logs** Not checked yet NEW

30

Agent tool calls now emit OpenTelemetry logs for observability of tool execution.

— Names OpenTelemetry but no setup detail given v0.8.9

---

11 **Inline Ask-User questions in chat** Not checked yet IMPROVED

20

Ask-User questions now display inline in the chat thread instead of overlaying the composer.

— Simple UI change described without further mechanism v0.8.9

└──▷ **BREAKING ON UPGRADE**

! Deployments with durable MCP name reservations must run the targeted index migration: upgrade every writer and temporarily gate reservation/account-erasure writes before replacing the legacy global index; do not delete tenantless tombstones.

WAS THIS USEFUL? 👍 👎

DATA

◆ Vector DBs & RAG

knowhere

Feedback ⓘ v1.2.60 NOTES ↗ RANK

Knowhere extracts, parses, and outputs structured chunks ready for AI Agents and RAG.

> knowhere's biggest changes this window are a page-count admission ceiling that blocks oversized or unreadable documents before billing or model calls, full PPTX embedded video/audio extraction with page-scoped media retrieval, and slide-deck parsing from page images, alongside new operator tooling for OpenSearch index limits and shadow backfill recovery.

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

01 Page-count admission ceiling before billing

Not checked yet

91

BREAKING

Document admission now rejects files exceeding a page-count ceiling, as well as documents whose page count is unreadable, before billing, PDF conversion, or model calls. The ceiling is set via `page_memory_config.max_pages`, which replaces the previous `MAX_PDF_PAGE_LIMIT` environment variable/config key and removes the 200-page split it controlled.

Prevent large documents from incurring billing or model calls by setting an admission ceiling — any document exceeding this page count is rejected before processing begins.

```yaml
page_memory_config:
  max_pages: 150
```

— Named config key, old env var it replaces, and example value. v1.2.60

02 Page-scoped media retrieval format

Not checked yet NEW

83

Page-level retrieval now includes video and audio assets embedded on matching pages: coarse hits surface page text, images, and media via `metadata.page_texts`, with `SAME-AS` pointers on shared pages; refined rows return the full node with each media item shown as marker, signed link, and transcript in `results[]`.

— Named response fields give a direct integration path. v1.2.59

03 Configurable OpenSearch open-index limit

Not checked yet NEW

80

The OpenSearch open-index limit is now configurable via the `KNOWHERE_OPENSEARCH_OPEN_INDEX_LIMIT` environment variable; production defaults render 600/540/480 (limit/idle-high/idle-low watermarks) and staging uses 100/90/80.

— Env var named with exact numeric defaults for two environments. v1.2.58

04 **Configurable LibreOffice conversion timeout** Not checked yet NEW

78

The LibreOffice conversion timeout used when ingesting PPTX/DOC files is now configurable via the `LIBREOFFICE_CONVERSION_TIMEOUT_SECONDS` environment variable, up from the previous 600-second default.

Raise the LibreOffice conversion timeout when ingesting large or complex PPTX/DOC files that exceed the 600-second default.

```
$ LIBREOFFICE_CONVERSION_TIMEOUT_SECONDS=1200 python -m knowhere.worker
```

— Env var and runnable command given directly. v1.2.59

05 **operator.py release-stopped-shadow command** Not checked yet NEW

70

Release a stopped shadow backfill issuer without stopping the shared OpenSearch node using the new `operator.py release-stopped-shadow` command.

Unblock a stalled shadow backfill in production without taking down the shared OpenSearch node.

```
$ python operator.py release-stopped-shadow
```

— Exact runnable command provided for the operation. v1.2.58

06 **PPTX/Slide Deck Parsing Enhancements** Not checked yet NEW

65

knowhere now parses slide decks from page images instead of TOC extraction, routing them through a dedicated slide path, and extracts embedded video and audio from PPTX files, separating them before PDF normalisation, transcribing them with timestamps, and publishing them as typed video/audio chunks linked to their source page.

— Describes the change but no config or flag to invoke it. Full pipeline described but no config surface to tune it.

v1.2.60 v1.2.59

THINNER COVERAGE BELOW

07 **Section hit completeness and page asset markers** Not checked yet

IMPROVED

54

Section hits on both the PostgreSQL and OpenSearch discovery paths now return every row the section owns (excluding Root-parked assets), and embedded media assets on a page are now listed under `metadata.assets` with one marker line in the page chunk's content.

— Named field and discovery paths but no example shown. v1.2.59

08 **Lexical compiler fingerprint change requires reparsing**

Not checked yet

BREAKING

47

The lexical compiler fingerprint changes in this release; any document parsed before the deploy but not yet indexed will fail indexing and must be reparsed.

— Clear migration note but no command to perform it. v1.2.59

09 **ShardPlan and probe.links removed** Not checked yet

BREAKING

40

The `ShardPlan` type, its validation and verdict tools, and `probe.links` have been removed.

— Names removed surfaces but gives no migration path. v1.2.60

└──▷ **BREAKING ON UPGRADE**

! The `MAX_PDF_PAGE_LIMIT` environment variable or config key is no longer read; the 200-page split it controlled has been removed.

! The `ShardPlan` type, its validation and verdict tools, and `probe.links` have been removed.

! The lexical compiler fingerprint changes: any document parsed before the deploy but not yet indexed will fail indexing and must be reparsed.

WAS THIS USEFUL? 👍 👎

ralforion OrionBelt Ontology Builder

Feedback ⓘ

v1.32.1 NOTES ↗ RANK

OrionBelt Ontology Builder is a browser-based ontology workbench for OWL ontologies and SKOS vocabularies. Streamlit + rdflib, no Java, no Protégé.

> OrionBelt Ontology Builder's single release in this window focused on list-page usability, adding name filtering with wildcard support and a compact page selector across the tool's list views.

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

01 **Name filters and pagination on list pages** Not checked yet

IMPROVED 75

All list pages — Classes, object and data properties, individuals, SKOS concepts, Relations, and Restrictions — gained a name/label filter that does case-insensitive plain text matching anywhere in the name, or whole-name matching with `*` for any text and `?` for one character. A narrow page selector now appears at the right of the filter row on any list with more than one page; changing the filter resets to page 1 and shows a matched-entry count.

Find all Classes whose names start with `address` by typing a wildcard pattern into the new name filter on the Classes list page.

📍 In the Classes list page, type `address*` into the filter box at the top of the list; matching entries appear immediately and the page resets to 1.

— Names exact wildcard syntax and affected list pages; no limits given v1.32.1

WAS THIS USEFUL? 👍 👎

liliang-cn CortexDB

Feedback ⓘ v2.156.0 NOTES ↗ RANK

CortexDB is a local-first AI agent memory with a knowledge graph, ontology constraints, and traceable reasoning—all in one SQLite file.

> CortexDB's latest releases center on standing rules that surface persistent instructions to agents at session start, plus a new Library view for browsing the knowledge graph and shared importflow tooling across both server binaries.

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

01 **Standing rules surfaced at session start and in Recalls**

Not checked yet NEW 44

Standing rules now reach the agent at session start and lead the recall, so persistent instructions are surfaced before the first prompt of every session; these standing rules are also now displayed on the Recalls page of the live view.

– Explains behaviour and timing but no config/flag named v2.156.0 v2.155.0

02 **Importflow tools shared across server binaries** Not checked yet

NEW 38

Both CortexDB binaries now expose importflow's tools, backed by a shared brain serving them, so structured-data import works across agents without needing a separate process.

– Describes mechanism but no commands or config named v2.153.0

03 **Library view on the Graph page** Not checked yet NEW 33

The live web interface adds a Library view that presents the Graph page as a browsable library.

– Names UI location but no further mechanism detail v2.154.5

WAS THIS USEFUL? 👍 👎

GrowGraph OntoCast

Feedback ⓘ v0.9.2 CODE ↗ RANK

OntoCast generates RDF knowledge graphs from documents by co-evolving domain ontologies and fact graphs with an agentic LLM pipeline.

> OntoCast shipped model-free facts validation with a new `check` CLI, a selectable and extensible facts vocabulary with automatic repairs and ontology planning, and a `curate` command for catalog consolidation — alongside breaking changes to ontology identification, extension-ontology storage, and property domain/range representation.

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

01 Model-free facts validation and check CLI

Not checked yet NEW

93

`check_facts` validates a facts graph against the catalog and its SHACL shapes without a model call, and never mutates the graph. The `ontocast check FILE…` CLI command runs this same check, accepting `--tenant` or `--project` to scope it without wiping the store or loading seed directories.

Validate a local Turtle facts file against your project catalog in CI, without touching the store or triggering any model calls.

```
$ ontocast check facts.ttl --tenant my-tenant --project my-
project
```

▶ View command output v0.9.2

Run the same SHACL-backed check programmatically and inspect the report before deciding whether to persist anything.

```python
from ontocast.check import check_facts
from ontocast.onto.rdfgraph import RDFGraph
from ontocast.tool import ToolBox

toolbox = ToolBox(...)
await toolbox.initialize(summarize_ontologies=False)
catalog = await toolbox.catalog_graphs()

facts = RDFGraph.from_file('facts.ttl')
report = await check_facts(facts, catalog)
print(report)
```

— Names the function, CLI command and flags precisely v0.9.2

02 **Facts vocabulary configuration** Not checked yet NEW 85

Adds `FACTS_VOCABULARY` to select the vocabulary facts extraction uses — `strict` (default) or `open` — backed by built-in well-known vocabularies including RDF, RDFS, OWL, XSD, SKOS, Dublin Core, PROV-O, SOSA/SSN, OWL-Time, CSVW, DCAT, ORG, vCard, FOAF, schema.org, and QUDT; `FACTS_ADDITIONAL_STANDARD_NAMESPACES` extends this pool with well-known terms now checked against their vocabulary with near-miss suggestions on unknown terms (`unknown_term`); and `FACTS_QUANTITY_FALLBACK_VOCABULARY` , with an optional `carrier` role, controls how quantities fall back when no vocabulary term matches.

— Full vocabulary list and both config keys named Config key named, mechanism only briefly sketched

v0.9.1

03 **Catalog curation via curate command** Not checked yet NEW 85

The new `ontocast curate` CLI command and `ontocast.curate_catalog` API review and optionally rewrite catalog ontology terms using the same consolidation decisions; pass `--apply` to write a new version without touching authored terms.

Review and rewrite the terms your pipeline has added to a catalog ontology, applying consolidation decisions without touching any hand-authored terms.

```
$ ontocast curate --apply
```

▶ View command output v0.9.1

— Runnable command and flag with clear scope v0.9.1

04 **Ontology resolution requires full IRI** Not checked yet BREAKING 80

An ontology is now identified by its IRI alone; a short name or prefix shared by several catalog ontologies no longer resolves and raises `AmbiguousOntologyRefError` , returned as HTTP 400 on `/ontologies/{id}/…` endpoints, listing their IRIs.

— Exact error, endpoint and status code named v0.9.1

05 **Extension ontology storage path changed** Not checked yet BREAKING 75

A document whose home ontology is external now writes its terms to `CURRENT_DOMAIN/<id>-ext-<hash>` ; extension ontologies stored under

`CURRENT_DOMAIN/<id>-ext` by an earlier release are not migrated.

— Old and new path keys both named, migration noted v0.9.1

06 **Facts graph repairs and findings** Not checked yet NEW

70

Adds automatic model-free repairs for facts graphs — reversed triples are corrected (`triple_reversed`), misrouted `sosa:hasSimpleResult` nodes move to `sosa:hasResult` (`paired_object_property`), and unknown terms matching a catalog term in another namespace are rewritten (`property_alias`, `class_alias`) — alongside a new mandatory finding `statement_without_subject` that flags observations or values stating nothing about a subject.

— Four repair types named but no invocation shown Single finding name, no usage detail given

v0.9.1

07 **ToolBox catalog retrieval and init skip** Not checked yet NEW

65

ToolBox.catalog_graphs() retrieves the merged catalog graph and its SHACL shapes, and passing `summarize_ontologies=False` to ToolBox.initialize() skips the model call that otherwise describes undescribed ontologies.

— Named methods and parameter but no example v0.9.2

08 **Ontology planning before extraction** Not checked yet NEW

60

`ONTOLOGY_PLAN_ENABLED` (on by default) runs one model call per document that lists the terms it will add with their domain and range, checked against the whole catalog before units render.

— Explains the mechanism and timing clearly v0.9.1

09 **Facts pipeline telemetry counters** Not checked yet NEW

60

New telemetry counters expose facts pipeline internals: `facts_statements_without_subject`, `facts_completion_detached_linked`, `facts_domain_repair_calls`, `facts_domain_violations_dropped`, `facts_date_values_added`, `facts_self_loops_dropped`, `facts_unlinked_dropped`, `facts_merge_rejections_by_reason`, and `facts_twin_labels`, plus context counters `ctx/facts_catalog_ontologies`, `ctx/admitted_planned_terms`, and `ctx/admitted_named_terms`.

— All counter names listed, no usage guidance v0.9.1

10 **Property domain/range representation changed** Not checked yet

60

BREAKING

A minted property's class domain and range are now written as `schema:domainIncludes` / `schema:rangeIncludes` instead of the previous representation; a datatype range stays `rdfs:range`.

— New predicates named, old representation unspecified v0.9.1

THINNER COVERAGE BELOW

11 **RDF file loading and literal validation in RDFGraph**

Not checked yet NEW 50

`RDFGraph.from_file` reads any RDF file by its file suffix, including Turtle 1.2 support, and `RDFGraph.is_well_formed_literal` checks whether an RDF literal is well-formed.

— Named methods but described only in one line each v0.9.2

└──▷ BREAKING ON UPGRADE

! An ontology is now identified by its IRI alone; a short name or prefix shared by several catalog ontologies no longer resolves and raises `AmbiguousOntologyRefError` (HTTP 400 on `/ontologies/{id}/…`, listing their IRIs).

! A document whose home ontology is external now writes its terms to `CURRENT_DOMAIN/<id>-ext-<hash>`; extension ontologies stored under `CURRENT_DOMAIN/<id>-ext` by an earlier release are not migrated.

! A minted property's class domain and range are now written as `schema:domainIncludes` / `schema:rangeIncludes` instead of the previous representation; a datatype range stays `rdfs:range`.

WAS THIS USEFUL?  

Qdrant

Feedback ⓘ v1.20.2 NOTES ↗ RANK

Qdrant is an open-source vector database and search engine that stores, indexes, and retrieves high-dimensional vectors for semantic search and retrieval-augmented generation.

> Qdrant v1.20.2 speeds up batch searches that share a filter and reverts HNSW graph construction to the classic algorithm by default.

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

01 **Classic HNSW graph construction restored as default**

Not checked yet

55

BREAKING

HNSW graph construction now uses the classic algorithm by default, so `ef_construct` tunes graph quality on large segments.

— Names config key and behaviour but no before/after detail v1.20.2

02 **Faster batch searches with shared filter** Not checked yet IMPROVED

25

Batch searches that use the same filter on the same vector are now significantly faster.

— No mechanism or numbers given for the speedup v1.20.2

WAS THIS USEFUL? 👍 👎

oxideterm

Feedback ⓘ v2.3.3 NOTES ↗ RANK

Oxideterm is a fast, native Rust + GPUI workspace for local shells, SSH/SFTP, and MCP tools—no WebView, no telemetry, open source..

> OxideTerm's v2.3.3 release adds CLI-driven batch SSH provisioning, BLE plugin host APIs, a paginated MCP terminal-history tool, and a reworked trusted host key manager, alongside session manager, paste-protection, and zen mode UI improvements.

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

01 Batch SSH connection creation via CLI

Not checked yet NEW

95

`oxideterm connections batch-create` creates up to 2,000 SSH connections from an OpenSSH-style `Host` template using placeholders `{host}`, `{profile}`, `{user}`, `{port}`, `{name}`, and `{group}`, one connection per line; writes are all-or-nothing transactional and saved templates are stored in the system credential store or an encrypted portable vault.

Provision dozens of SSH connections at once during an infrastructure onboarding — pipe a host list through the CLI instead of clicking through the UI.

```
$ oxideterm connections batch-create --template 'Host {host}\n
User {user}\n  Port {port}\n  ProxyCommand ssh bastion -W %h:%p'
--hosts hosts.txt
```

— Runnable command with full placeholder list and transactional limit v2.3.3

02 Remote Shell Integration v7 OSC handling

Not checked yet IMPROVED

85

Remote Shell Integration is now version 7, gating OSC 133 command markers and private OSC adapters behind verified SSH session markers (`OXIDETERM_PRIVATE_OSC=1`) while keeping OSC 7 working-directory tracking active even if servers strip `LC_OXIDETERM_SESSION` ; existing v6 scripts upgrade automatically.

— Named env vars and version with upgrade behaviour detailed v2.3.3

03 BLE plugin host APIs

Not checked yet NEW

75

New BLE plugin host APIs add `ble.read` and `ble.write` capabilities alongside `ble.sessions`, `ble.describe`, `ble.onNotification`, and

`ble.offNotification` interfaces, plus manifest-declared terminal toolbar chips and pane-bound floating overlays; GATT payloads never enter logs or error messages.

— Named APIs but no usage example or command shown v2.3.3

04 **Paginated terminal history MCP tool** Not checked yet NEW

65

New `terminals_history` MCP tool lets AI clients paginate through terminal scrollback or screen contents using snapshot slots and `next_cursor`, independent of the visible viewport and unaffected by user scrolling.

— Named tool and cursor field, no example invocation v2.3.3

05 **Session Manager sidebar redesign** Not checked yet IMPROVED

60

Session Manager now renders directly in the sidebar with local search (cleared and closed on `Esc`), whole-row drag-and-drop reordering, Ctrl/Cmd and Shift multiselect, and context-menu batch deletion.

— UI behaviors named but no config or command v2.3.3

06 **Editable preview for multiline paste protection** Not checked yet

IMPROVED

60

When multiline paste protection is enabled, pasted drafts now open in an editable preview with `Enter` for newlines, `Esc` to cancel, Undo/Redo, and an action to strip Markdown code fences before sending.

— Named keys and actions, no command or config v2.3.3

THINNER COVERAGE BELOW

07 **Trusted host key management in Settings** Not checked yet NEW

55

Manage trusted host keys in Settings under Connection Settings: search saved keys by host, port, or fingerprint, copy fingerprints, delete entries, open the trust file, and replace only the reviewed entry when a remote host key changes.

— Clear UI path but no config file name or API v2.3.3

08 **Zen mode exit button on hover** Not checked yet IMPROVED

35

Zen mode now shows an exit button when the pointer moves to the top window edge, removing the requirement to remember a keyboard shortcut.

— Simple UI tweak, thinly described v2.3.3

WAS THIS USEFUL? 👍 👎

EVALUATE

◆ AI Observability & Evals

UK AI Security Institute Inspect AI

Feedback ⓘ

0.3.291 CODE ↗ RANK

Inspect is the UK AI Security Institute's framework for large language model evaluations, covering datasets, solvers, scorers, and agentic eval tasks.

> Inspect AI ships a trio of log-management CLI commands for sharded eval runs, expands Sentinel's human-in-the-loop oversight to bridged agents, and adds support for new Hugging Face generation-config handling and OpenAI's GPT-6.2 Astra and Sol models.

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

01 Log merge, exclude, and zero CLI commands

Not checked yet NEW

90

Inspect adds `inspect log merge-shards` to merge a sharded evaluation run's logs from the command line, `inspect log exclude` to exclude samples of a finished log from its metrics (also available via the equivalent Python API), and `inspect log zero` to count samples of a finished log as failures via an empty-attempt run. Eval-set selection workers can now use empty_attempt() and zero_samples() to cover unstarted tasks.

Merge sharded logs from a large parallel eval run into a single consolidated log file.

```
$ inspect log merge-shards
```

Exclude specific samples from a finished log so they are dropped from its aggregate metrics.

```
$ inspect log exclude
```

Mark all unscored samples in a finished log as failures using an empty-attempt run.

```
$ inspect log zero
```

— Names exact runnable commands and their Python API equivalents 0.3.291

02 Human review escalation in Sentinel for bridged agents

Not checked yet IMPROVED 60

Sentinel now runs for bridged agents, extending oversight to agent bridge traffic. The sentinel sidecar can escalate to a human reviewer for a tool call, a model's whole turn, or a model's request before it is sent, and human() can now ask a person about a request before it is sent to the model, adding pre-send human review to sentinel workflows.

— Describes mechanism but no config key or flag to enable it 0.3.291

THINNER COVERAGE BELOW

03 **Ignore Hugging Face generation config option** Not checked yet

NEW 40

Add an option to ignore a Hugging Face model repository's generation config, and record template and generation settings in the eval log.

— Option described in prose without a named flag 0.3.291

04 **Support for GPT-6.2 Astra and Sol models** Not checked yet NEW

30

Adds support for GPT-6.2 Astra and GPT-6.2 Sol OpenAI models.

— Bare model names with no usage or integration detail 0.3.291

WAS THIS USEFUL? 👍 👎

Arize Phoenix

Feedback ⓘ @arizeai/phoenix-cli@1.23.0 NOTES ↗

RANK

Arize Phoenix is an open-source platform that monitors LLM application traces and evaluates their outputs.

> Arize Phoenix added a CLI command for managing prompt version tags and expanded OpenTelemetry sampling and span-limit controls in register(), along with safeguards against silent misconfiguration.

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

01 **Sampler and span-limit configuration in register()**

Not checked yet NEW 90

register() now accepts `sampler` and `spanLimits` parameters, with `AlwaysOnSampler`, `AlwaysOffSampler`, `ParentBasedSampler`, and `TraceIdRatioBasedSampler` re-exported for direct use. The library also warns at runtime when `OTEL_TRACES_SAMPLER`, `OTEL_TRACES_SAMPLER_ARG`, `OTEL_SPAN_*`, `OTEL_ATTRIBUTE_*`, or `OTEL_BSP_*` environment variables are set but have no effect, since OpenTelemetry JS 3 silently ignores them.

Sample only 20% of traces in production by passing a `TraceIdRatioBasedSampler` to register() instead of relying on ignored `OTEL_TRACES_SAMPLER` env vars.

```javascript
import { register, TraceIdRatioBasedSampler } from
'@arizeai/phoenix-otel';

register({
  projectName: 'my-llm-app',
  sampler: new TraceIdRatioBasedSampler(0.2),
});
```

Cap attribute and span counts at registration time using `spanLimits` to avoid over-sized traces.

```javascript
import { register } from '@arizeai/phoenix-otel';

register({
  projectName: 'my-llm-app',
  spanLimits: {
    attributeCountLimit: 64,
    eventCountLimit: 10,
  },
});
```

— Named parameters, classes, and env vars with runnable code `@arizeai/phoenix-otel@3.1.0`

02 **Prompt version tagging via CLI** Not checked yet NEW 80

The `px prompt tag <prompt-version-id> --tag <name>` command tags a prompt version or moves an existing tag, with an optional `--description <text>` argument; omitting `--description` clears any existing description.

Pin a stable prompt version as `production` so downstream consumers can resolve it by tag name instead of ID.

```
$ px prompt tag v1a2b3c4 --tag production --description 'Approved for prod rollout'
```

Move the `production` tag to a newer version and clear its old description in one step.

```
$ px prompt tag v9z8y7x6 --tag production
```

— Exact command and flags given, runnable as-is `@arizeai/phoenix-cli@1.23.0`

03 **Positional exporter misconfiguration now throws TypeError**

Not checked yet BREAKING 60

Re-exported `BatchSpanProcessor` and `SimpleSpanProcessor` now throw a `TypeError` when an exporter is passed positionally instead of as `{ exporter }`, catching a misconfiguration that previously dropped every span silently.

— Clear before/after behaviour but no code example given `@arizeai/phoenix-otel@3.1.0`

WAS THIS USEFUL? 👍 👎

Langfuse

Feedback ⓘ v4.82.0 NOTES ↗ RANK

Langfuse provides tracing, evaluation, and monitoring for LLM applications.

> Langfuse moved audit logging onto ClickHouse for higher-volume access, added Google Vertex AI as an AI gateway provider and Gemini support in decision-model evaluators, and normalized cost tracking across native integrations and OpenTelemetry events.

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

01 Audit logs backed by ClickHouse

Not checked yet IMPROVED

60

Audit logs are now read from ClickHouse, enabling higher-volume audit log access backed by a columnar store, and existing Postgres audit logs are backfilled into ClickHouse.

– Names the backend change but no reader-facing action v4.82.0 v4.81.0

THINNER COVERAGE BELOW

02 Developer preset for API keys

Not checked yet NEW

50

Langfuse adds a Developer preset for API keys and renames the existing Ingestion preset to Application.

– Names the presets but lacks deeper mechanism v4.82.0

03 Cost normalization across providers and OpenTelemetry

Not checked yet IMPROVED

50

Langfuse normalizes provider usage and supplied costs across native integrations for consistent cost tracking, and applies model pricing and enrichment to owned `OpenTelemetry` events.

– Describes scope but no concrete usage path v4.82.0 v4.81.0

04 Filtering in datasets and events UI

Not checked yet NEW

50

Adds filtering by status in the datasets UI and by numeric metadata values in the events table.

– Names two UI surfaces, clear navigation but no further detail v4.81.0

05 **Google Vertex AI in AI gateway** Not checked yet NEW 45

Adds Google Vertex AI as a provider in the AI gateway, alongside existing providers.

— Names the provider but no config detail given v4.81.0

06 **Gemini decision models in evaluators** Not checked yet NEW 40

Adds Gemini decision models as an option in decision-model evaluators.

— Bare addition with no further mechanism described v4.81.0

WAS THIS USEFUL? 👍 👎

Comet Opik

Feedback ⓘ 2.2.144 NOTES ↗ RANK

Opik traces, evaluates, and monitors LLM applications and agentic workflows, with datasets, an LLM-as-judge metric library, and production dashboards.

> Opik introduces a guardrails system spanning SDK, backend API, and UI for validating LLM outputs against policies.

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

01 **Guardrails API, policies UI, and SDK method** Not checked yet NEW 48

Opik adds a guardrails system: an SDK method to run guardrail rules and stored policies against the Opik backend, a new guardrail policies page in the UI for viewing and managing policies, and new guardrails validations and guardrail policies API contracts on the backend (these currently respond 501 while implementation lands).

— Names SDK method, UI page, API contracts but no concrete flags or paths given 2.2.144

WAS THIS USEFUL? 👍 👎

Braintrust

Feedback ⓘ updated · Nov 18, 2026 NOTES ↗

Braintrust provides evaluation, tracing, and improvement workflows for AI applications.

Braintrust adds on-device trace translation in Chrome/Edge, Loop pattern rules with SQL filters, scheduled project-activity digests, Terraform GA with new ephemeral token and user-lookup resources, and coding-agent trace continuation via `--parent`.

Translate an entire non-English trace to English in desktop Chrome or Edge using the browser's built-in API, keeping trace text on-device without consuming model credits. Not checked yet

Create Loop pattern rules to scope each scheduled Patterns run to specific traces with a SQL filter and define problems for Loop to watch for. Not checked yet

Send recurring project-activity summaries to your team with scheduled digests, choosing compact or detailed writing and setting a schedule and destinations. Not checked yet

NEW INTEGRATION Continue a coding-agent trace from an existing SDK span with `bt trace run --parent <export()>`, attaching the new session below that span instead of starting a new trace. Not checked yet

Run an MCP server tool with only its own arguments — `bt mcp run` no longer requires the telemetry argument. Not checked yet

+3 more

NEW INTEGRATION Google Antigravity SDK tracing

NEW INTEGRATION aiobotocore async Bedrock Runtime tracing

UPDATED SUPPORTED PLATFORM LiveKit Agents v1.8.11+ support

NEW INTEGRATION OpenCode 2 tracing in bt CLI

UPDATED INTEGRATION Service tier tracking across LLM providers

└──▷ BREAKING ON UPGRADE

- ! The Terraform provider now requires Terraform or OpenTofu 1.11.1 or later.
- ! `braintrust_ai_secret` now requires `secret_version` alongside `secret`, and `braintrust_env_var` now requires `value_version` alongside `value`; changing only the value no longer updates it, and references to `secret` or `value` return `null`.
- ! The `braintrust_org_automation` Terraform resource is removed; delete its blocks from your configuration and run `terraform state rm` on each address.
- ! LiveKit span names are renamed: `livekit_agent_session` → `session`, `user_speaking` → `user_speech`, `agent_speaking` → `agent_speech`,

`stt_processing` → `stt`, `tts_request` → `tts`, `llm_request_run` → `llm_response`.

- ! LiveKit `vad_endpointing` and `eou_detection` spans are removed and per-component metrics (`livekit_metrics`) are no longer recorded.
- ! LiveKit TTS time to first audio is renamed from `time_to_first_token` to `time_to_first_audio`.
- ! Pipecat span names are renamed: `pipecat_pipeline` → `session`, `pipecat_llm_response` → `llm_response`, `stt_transcription` → `stt`, `tts_response` → `tts`, `user_speaking` → `user_speech`.
- ! `BraintrustPipecatObserver` is no longer exported; use setup_pipecat(), wrap_pipeline_worker(), or auto_instrument() instead.
- ! `client.agents`, `client.org_automations`, and their types are removed from the `braintrust.api` REST client.

WAS THIS USEFUL? 👍 👎

GOVERN

◆ AI Security Posture & Discovery

Mend.io Mend AI

Feedback ⓘ changelog · Nov 18, 2026

CHANGELOG ↗ RANK

Mend AI is an AI security service that discovers AI components and helps teams govern their use in software applications.

> Mend AI expanded its reporting with a new report that surfaces AI entities across models, agents and tools, a dedicated report for chat system prompts, and org-wide report generation — while renaming the old AI Inventory report.

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

01 **AI Inventory report across models, agents, tools, MCP servers**

Not checked yet NEW 50

New AI Inventory report provides one row per detected AI entity across models, system prompts, agents, agent configurations, tools, and MCP servers.

– Names entity types covered but no UI path or command changelog-20261118-605c5559

02 **Conversational System Prompts report** Not checked yet NEW 50

New Conversational System Prompts report pinpoints system prompts behind chat-based AI applications and shows where each is defined in code, so security teams can focus review on user-facing AI.

– Explains mechanism (locates prompts in code) but no exact navigation

changelog-20261118-605c5559

03 **Organization-scope report generation** Not checked yet IMPROVED 40

All Mend AI reports can now be generated at the organization scope, in addition to the existing application and project scopes.

– Clear capability but no config key or command given changelog-20261118-605c5559

04 **AI Inventory report renamed to Models Inventory** Not checked yet 25

BREAKING

The previous AI Inventory report, which covered AI models only, is renamed to Models Inventory.

– Simple rename notice with no migration detail changelog-20261118-605c5559

└──▷ **BREAKING ON UPGRADE**

! The previous AI Inventory report (which covered AI models only) is renamed to Models Inventory.

WAS THIS USEFUL?  

◆ Model & Agent Protection

ToolHive

Feedback ⓘ v0.66.0 NOTES ↗ RANK

ToolHive runs MCP servers in isolated containers and enforces identity and access policies for local and Kubernetes deployments.

> ToolHive v0.66.0 adds container hardening with an opt-in read-only root filesystem for Docker workloads, introduces ID-JAG minting from OAuth exchange grants with actor validation, and extends vMCP embedder APIs for audit records and URL elicitation fallback.

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

01 **ID-JAG minting from OAuth exchange grants** Not checked yet NEW

60

ToolHive now mints ID-JAGs from OAuth exchange grants and validates the actor before issuance, with the acting agent named in the token.

— Explains mechanism and actor validation, no command shown v0.66.0

THINNER COVERAGE BELOW

02 **Read-only root filesystem for Docker workloads** Not checked yet

NEW

45

ToolHive can now harden Docker workload containers by enabling an opt-in read-only root filesystem.

— Names the hardening mechanism but not the exact flag v0.66.0

03 **vMCP embedder additions for audit and URL fallback**

Not checked yet NEW

40

vMCP embedders can now add audit event subjects to emitted audit records, and can set the legacy URL fallback text shown to MCP clients that declared URL elicitation.

— Describes two embedder API additions without exact signatures v0.66.0

04 **Server profile in anonymous usage metrics** Not checked yet

IMPROVED

20

Anonymous usage metrics now include the server profile.

— Bare statement with no further detail v0.66.0

WAS THIS USEFUL? 👍 👎

◆ AI Red Teaming & Testing

Giskard

Feedback ⓘ giskard-scan/v1.0.2 NOTES ↗ RANK

Giskard tests and evaluates LLM applications for performance, bias, and security issues.

> Giskard shipped a new v3 vulnerability scanner and RAG knowledge-base evaluator for agentic systems, alongside developer-convenience additions like synchronous eval execution and prompt-template validation, while formally deprecating v2's tabular scanning and test-suite tooling.

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

01 Red-team vulnerability scan for AI agents

Not checked yet NEW

86

Run automated red-team scans against an AI agent using `vulnerability_scan`, covering prompt injection, harmful content, stereotypes, misinformation, and other OWASP LLM Top-10 threat categories. Requires installing `giskard[scan]`.

Red-team an agent against OWASP LLM Top-10 threat categories, including prompt injection and harmful content, without writing individual test cases.

python

```python
import asyncio
from giskard.scan import vulnerability_scan

async def my_agent(inputs: str) -> str:
    return f'Echo: {inputs}'

async def main() -> None:
    await vulnerability_scan(
        target=my_agent,
        description='A customer support chatbot for an e-commerce
platform.',
        languages=['en'],
    )

asyncio.run(main())
```

— Names function, install extra, and threat categories with runnable example

giskard-scan/v1.0.2

02 Custom adversarial scenario generators

Not checked yet NEW

84

Extend vulnerability scans with custom adversarial scenarios by passing your own `ScenarioGenerator` instances to `generate_suite`, or by registering them on `vulnerability_suite_generator_registry`.

Add a domain-specific adversarial scenario to a vulnerability scan without a model provider or network access.

python

```python
from giskard.scan import generate_suite
from giskard.scan import ScenarioGenerator

class MyRiskGenerator(ScenarioGenerator):
    def generate(self, context, budget):
        # Return a deterministic scenario targeting a known
domain risk
        return [my_known_risk_scenario]

suite = await generate_suite(
    target=my_agent,
    description='A customer support chatbot for an e-commerce
platform.',
    generators=[MyRiskGenerator()],
)
await suite.run()
```

— Named classes/functions with a runnable deterministic example `giskard-scan/v1.0.2`

03 **Synchronous run_sync() for checks and suites** Not checked yet

NEW

70

Call run_sync() on a scenario or suite to execute evals from a plain script without managing an event loop.

Run a Groundedness eval from a plain Python script without an async event loop.

```python
from giskard.checks import Scenario, Groundedness

def get_answer(inputs: str) -> str:
    return "Paris"

scenario = (
    Scenario("test_france_capital")
    .interact(inputs="What is the capital of France?",
outputs=get_answer)
    .check(Groundedness(name="answer is grounded",
context="France is in Western Europe. Its capital is Paris."))
)
result = scenario.run_sync()
result.print_report()
```

— Named method with working script example, minimal further detail `giskard-core/v1.0.4`

04 **Telemetry opt-out via environment variable** Not checked yet NEW

65

Opt out of telemetry by setting `DO_NOT_TRACK=1` or `GISKARD_TELEMETRY_DISABLED=1` as an environment variable or in a `.env` file in the working directory.

— Exact env var names given, no further mechanism explained `giskard-scan/v1.0.2`

05 **RAG quality evaluator replacing v2 RAGET** Not checked yet NEW

62

Evaluate RAG knowledge-base quality with `quality_scan` and `KnowledgeBase`, replacing the v2 RAGET test-set generator. This now ships natively in `giskard-scan` with no dependency on v2.

— Names the functions and migration from v2 but no runnable example `giskard-scan/v1.0.2`

THINNER COVERAGE BELOW

06 **v2 tabular scan and test suite deprecated** Not checked yet

DEPRECATED

52

Giskard v2 is no longer actively maintained; the automatic tabular scan (`giskard.Model` + `giskard.Dataset`), the `giskard.testing` ML test suite, and Giskard Hub are v2-only and not planned for v3.

— Names deprecated classes/modules but no migration path given `giskard-scan/v1.0.2`

07 **validate_template for agent prompts** Not checked yet NEW

50

Validate agent prompt templates before execution with the new `validate_template` function in `giskard.agents`.

— Named function and module but no example or mechanism detail giskard-core/v1.0.4

---

08 **Graceful skip of unavailable Hub checks** Not checked yet IMPROVED

30

Giskard now warns and skips Hub checks that are unavailable at runtime, rather than failing the entire scan.

— Describes behaviour change but no named surface or example giskard-core/v1.0.4

└──▷ **BREAKING ON UPGRADE**

! Giskard v2 is no longer actively maintained; the automatic tabular scan (`giskard.Model` + `giskard.Dataset`), `giskard.testing` ML test suite, and Giskard Hub are v2-only and not planned for v3.

WAS THIS USEFUL?  

◆ AI Runtime Defense & Guardrails

FailproofAI

Feedback ⓘ fp-cloud-cli-v0.0.3 NOTES ↗ RANK

FailproofAI provides observability and enforcement for AI agent harnesses, capturing every run and blocking dangerous tool calls before they execute.

> FailproofAI's CLI gained browser and SSO-based login with email-code and headless fallbacks, while its telemetry pipeline became durable against outages, credential refusals, and partial server acknowledgements with dedicated retry budgets.

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

01 Durable telemetry delivery with retry budgets

Not checked yet

95

BREAKING

Telemetry batches that fail due to network errors or HTTP `408`, `429`, or `5xx` responses are now kept on disk and retried automatically once delivery recovers, for both NDJSON event batches and OTLP exports, without consuming the batch's retry budget. Permanently rejected batches (HTTP `400` or `event_too_large`) are preserved under `state/failed/` for manual inspection instead of being discarded. Credential refusals (HTTP `401` / `403`) now get their own bounded retry budget so repeated refusals no longer block other batches, and batches refused during an outage are redriven once a working key arrives. Partial acknowledgements (HTTP `200` that stores some events and skips others) have a separate retry budget from outage passes, preventing skipped events from being double-counted in metrics. The CLI's delivery status view now understands the `.sN` (skipped acknowledgements) and `.rN` (client rejections) filename suffixes on parked batch files. Batches marked `.poison` before v1.0.14 are not resent automatically and must be recovered by hand; OTLP exports refused with 401/403 require manual recovery even after the key is corrected — stop the daemon, back up the affected files under `<FAILPROOFAI_HOME>/state/failed/otlp/`, move them into `state/spool-otlp/` with fresh `otlp-<unique>.jsonl` names, then restart the daemon.

— Names exact HTTP codes, paths, suffixes, and a full manual recovery procedure v1.0.14

02 Browser and SSO login for fp login

Not checked yet IMPROVED

70

`fp login` now opens a browser automatically when one is available (including over SSH) and completes the terminal session once authorized, falling back to displaying the link where it cannot open one. Press `Tab` during login to switch to a 6-digit email code flow, press `Tab` again to return to the browser flow, or pass `--email` to go straight to

the code path. Passing `--json` to `fp login`, or running it without a terminal, keeps the existing code-based flow unchanged.

Authenticate from a remote machine or CI environment where a browser cannot open, so the link is displayed for you to open on another device.



— Names exact flags and Tab behavior but no deeper protocol detail `fp-cloud-cli-v0.0.3`

└──▷ BREAKING ON UPGRADE

! Batches marked `.poison` before v1.0.14 are not resent automatically and must be recovered by hand.

! OTLP exports refused with 401/403 require manual recovery even after the key is corrected: stop the daemon, back up the affected files under `<FAILPROOFAI_HOME>/state/failed/otlp/`, move them into `state/spool-otlp/` with fresh `otlp-<unique>.jsonl` names, then restart the daemon.

WAS THIS USEFUL? 👍 👎

Doberman-Core

Feedback ⓘ v0.21.0 NOTES ↗ RANK

Doberman-Core is a runtime security tool that blocks unsafe or unintended AI agent inputs, outputs, and tool calls before execution.

> Doberman-Core's v0.21.0 release introduces a Windows Hello-required 2FA mode that removes TOTP secrets from disk in favor of TPM-backed approval signatures, alongside breaking changes affecting the `winhello` packaging extra and Cursor hook installation.

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

01 **Windows Hello-required 2FA mode** Not checked yet NEW 90

New `doberman 2fa hello enroll` command, run from an admin terminal, enrolls a workstation in Hello-required 2FA mode: it deletes the TOTP secret from disk and replaces approvals with TPM-backed Windows Hello key signatures tied to each action. `doberman doctor` now reports whether Hello-required mode is active, and once enrolled, loosening commands require a Windows Hello gesture to proceed.

Enroll a Windows workstation in Hello-required 2FA mode so TOTP secrets are removed from disk and all approvals are backed by a TPM key gesture.

```
$ doberman 2fa hello enroll
```

— Exact enroll command plus mechanism and downstream effects given. v0.21.0

02 **Cursor hook wait time increased to 660 seconds** Not checked yet 60

BREAKING

Cursor users must re-run `doberman install-hooks --host cursor` after upgrading, because the hook now waits 660 seconds.

— Exact command and new wait value, but no explanation of cause. v0.21.0

THINNER COVERAGE BELOW

03 **`winhello` extra now empty** Not checked yet BREAKING 35

The `winhello` packaging extra is now empty, a side effect of the new Windows Hello-required 2FA mode being introduced in this release.

— States the change but gives no migration mechanism. v0.21.0

└──▷ **BREAKING ON UPGRADE**

! Cursor users must re-run `doberman install-hooks --host cursor` after upgrading so the hook waits 660 seconds.

! The `winhello` extra is now empty.

WAS THIS USEFUL? 👍 👎

AI MODELS

◆ Frontier Models

Anthropic

Feedback ⓘ updated · Nov 18, 2026 NOTES ↗ RANK

The Claude Developer Platform is the API for building applications on Anthropic's Claude models that generate, analyze, and automate work.

> Anthropic's Claude API gained beta linked billing for flexible usage payment and automatic text watermarking across its latest Claude models.

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

01 **Linked billing for API usage** Not checked yet NEW 81

Enable linked billing in beta by passing the `user-profiles-billing-2026-10-01` beta header, then set `billing_policy` on each `/v1/messages` or `/v1/chat/completions` request to choose whether your app or the customer's Claude organization pays for usage.

— Names exact beta header, field and both endpoints snapshot-20261118

THINNER COVERAGE BELOW

02 **Text watermarking for generated content** Not checked yet NEW 59

Text generated by Claude Opus 4.6, Claude Opus 4.5, Claude Sonnet 4.5, and Claude Haiku 4.5 now carries Anthropic's text watermark on supported platforms; setting `temperature`, `top_p`, or `top_k` to a non-default value may suppress the watermark.

— Names affected models and suppression conditions, no verification tool given

snapshot-20261118

WAS THIS USEFUL? 👍 👎

Mistral AI

Feedback ⓘ updated · Nov 18, 2026 NOTES ↗ RANK

Learn how to deploy and use Mistral AI's Large Language Models with our comprehensive documentation, guides, and tutorials.

> Mistral AI's Large 4 models gained a new reasoning-depth control, while the previous way to disable reasoning entirely was removed as a breaking change.

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

01 **Reasoning effort control in Mistral Large 4** Not checked yet NEW

78

`mistral-large-4` and `mistral-large-4-0` now support a `reasoning_effort` parameter with values `low`, `medium`, or `high` to control reasoning depth. The model always reasons, and `reasoning_effort` defaults to `medium` when omitted.

— Names exact param and values but no runnable example given. snapshot-20261118

02 **`reasoning_effort="none"` no longer accepted** Not checked yet

60

BREAKING

For `mistral-large-4` and `mistral-large-4-0`, setting `reasoning_effort="none"` is no longer valid; requests that still pass this value now return HTTP 400.

— Clear before/after but no migration guidance given. snapshot-20261118

└──▷ **BREAKING ON UPGRADE**

! `reasoning_effort="none"` is no longer accepted for Mistral Large 4 (`mistral-large-4`, `mistral-large-4-0`); requests that set `reasoning_effort="none"` now return HTTP 400.

WAS THIS USEFUL? 👍 👎

OpenAI GPT-6.2 Astra

Feedback ⓘ released · Nov 18, 2026

MODEL ↗ DOCS ↗

GPT-6.2 Astra from OpenAI was released on 17 Nov 2026. 1050K-token context; $8.00 in / $40.00 out per million tokens (OpenRouter). Not checked yet

Succeeds GPT-6 Astra. Not checked yet

OpenAI adds GPT-6.2 Astra and GPT-6.2 Sol models, available as `gpt-6.2-astra` in the API Not checked yet

Call the new `gpt-6.2-astra` model via the Responses API as OpenAI's most capable model for complex, end-to-end tasks. Not checked yet

Use the new GPT-6.2 Sol model for everyday workloads including coding and powering agents at scale. Not checked yet

WAS THIS USEFUL? 👍 👎

OpenAI GPT-6.2 Astra Pro

Feedback ⓘ

released · Nov 18, 2026 MODEL ↗

GPT-6.2 Astra Pro from OpenAI was released on 17 Nov 2026. 1050K-token context; $8.00 in / $40.00 out per million tokens (OpenRouter). Not checked yet

Succeeds GPT-6 GPT-6.2 Astra Pro. Not checked yet

WAS THIS USEFUL? 👍 👎

OpenAI GPT-6.2 Sol

Feedback ⓘ released · Nov 18, 2026

MODEL ↗

GPT-6.2 Sol from OpenAI was released on 17 Nov 2026. 1050K-token context; $2.00 in / $10.00 out per million tokens (OpenRouter). Not checked yet

Succeeds GPT-6.1 Sol. Not checked yet

WAS THIS USEFUL? 👍 👎

OpenAI GPT-6.2 Sol Pro

Feedback ⓘ

released · Nov 18, 2026 MODEL ↗

GPT-6.2 Sol Pro from OpenAI was released on 17 Nov 2026. 1050K-token context; $2.00 in / $10.00 out per million tokens (OpenRouter). Not checked yet

Succeeds GPT-6.1 GPT-6.2 Sol Pro. Not checked yet

WAS THIS USEFUL? 👍 👎

Particle AI Particle-1

Feedback ⓘ

price increase · Nov 15, 2026 MODEL ↗

Particle-1 is a multimodal model from Particle AI.

Price increase on OpenRouter (particle-ai/particle-1): $0.05 → $0.06 in, $0.50 → $0.50 out per million tokens (+20%), first seen 15 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

◆ Open-Weight Models

Tencent Hy4 flash

Feedback ⓘ released · Nov 18, 2026

MODEL ↗

Hy4 flash from Tencent was released on 16 Nov 2026. 1049K-token context; $0.035 in / $0.1405 out per million tokens (OpenRouter). Not checked yet

WAS THIS USEFUL? 👍 👎

Alibaba Qwen4 Max

Feedback ⓘ released · Nov 18, 2026

MODEL ↗ NOTES ↗

Qwen4 Max from Alibaba was released on 16 Nov 2026. 1M-token context; $1.425 in / $4.275 out per million tokens (OpenRouter). Not checked yet

Qwen Model Studio adds qwen4-max multimodal flagship and three MindLab Macaron models to its international deployment scope Not checked yet

Adds `qwen4-max`, a multimodal model supporting text generation, image and video understanding, and deep reasoning, with web search, tool calling, and structured outputs, compatible with OpenAI and Anthropic APIs. Not checked yet

Adds `mindlab/macaron-v1.1-venti` and `mindlab/macaron-v1.1-coding-venti`, MoE-architecture models with a 1M context window covering chat, agent, coding, and UI4A tasks, available in the Singapore region. Not checked yet

Adds `mindlab/macaron-v1-tall`, a small model with a 256K context window optimised for chat, agent, coding, and UI4A, available in the Singapore region. Not checked yet

WAS THIS USEFUL? 👍 👎

H2O.ai

Feedback ⓘ announced · Nov 17, 2026 PRESS ↗ RANK

H2O.ai is a H2O LLM Studio - a framework and no-code GUI for fine-tuning LLMs.
Documentation: https://docs.h2o.ai/h2o-llmstudio/

> H2O.ai introduced two new zero-shot foundation models — TimeH2O for time-series forecasting and TabH2O for tabular prediction — now generally available on H2O AI Cloud, with open research weights for TimeH2O released separately.

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

01 **TimeH2O zero-shot time-series forecasting model** Not checked yet

NEW 70

TimeH2O is a 108-million-parameter time-series foundation model that forecasts zero-shot without building a per-series model, and can be fine-tuned on private data for use cases including demand forecasting, inventory, financial planning, energy load, and healthcare resource planning.

— Names parameter count and use cases but no API or command. launch-20261117-56f29e0d

02 **Open research weights and licensing for TimeH2O** Not checked yet

NEW 70

TimeH2O-v1 open weights are available on Hugging Face under the H2O Model Non-Commercial License for research and evaluation; production use requires a commercial license from H2O.ai.

— Names exact weights release, host, and license terms for use. launch-20261117-56f29e0d

THINNER COVERAGE BELOW

03 **TabH2O tabular foundation model** Not checked yet NEW 55

TabH2O is a tabular foundation model for structured business data, supporting zero-shot prediction and fine-tuning for risk scoring, fraud detection, and customer analytics, using the same approach as TimeH2O.

— Describes purpose but lacks model size or technical detail. launch-20261117-56f29e0d

04 **MLOps deployment and governance for foundation models**

Not checked yet IMPROVED 55

Both TimeH2O and TabH2O are deployable, monitored, and governed through H2O MLOps on H2O AI Cloud, with support for managed cloud, private cloud, on-premises, air-gapped, and FedRAMP High certified government environments.

— Names deployment environments but no concrete setup steps. launch-20261117-56f29e0d

WAS THIS USEFUL? 👍 👎

Shisa.AI Shisa

Feedback ⓘ announced · Nov 16, 2026 PRESS ↗

RANK

Shisa.AIã¯â¦æ¥æ¬ã®âAIç
ççº¶ãã©ãã§ãã§ãâ¦æ¥æ¬èªãâãã¹ãâåãª
ãã«âãâãâãâãâãâ¼ãâ³ãâ¢ãâ«ã¨ãâãâãâãâãâ

> Shisa's Jouzu coding tool launched with same-day Qwen 4 Max support, multi-agent development workflows, a TextGuard web-content safety check, and Japanese voice input.

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

01 **Qwen 4 Max model support and mixing** Not checked yet NEW

45

Jouzu adds Qwen 4 Max as an available model on its release day, giving Japanese developers access to the latest Alibaba Cloud model from day one, and supports mixing Qwen 4 Max with other models based on task complexity and cost so teams can balance quality and efficiency.

— Names the model but no config/flag to select it launch-20261116-6fc475f1

02 **Multi-agent development workflows** Not checked yet NEW

35

Jouzu supports multi-agent workflows that split development, review, and verification tasks across multiple AI agents simultaneously.

— Describes behavior but no setup detail launch-20261116-6fc475f1

03 **TextGuard web content safety check** Not checked yet NEW

35

Jouzu includes a TextGuard feature that checks web pages before passing them to AI models, blocking malicious prompts and protecting sensitive data.

— Explains purpose but no mechanism or config launch-20261116-6fc475f1

04 **Voice input for Japanese prompts** Not checked yet NEW

35

Jouzu supports voice input using Shisa AI real-time speech recognition, allowing spoken Japanese prompts alongside keyboard input.

— Names the capability but no usage detail launch-20261116-6fc475f1

WAS THIS USEFUL? 👍 👎

DeepSeek V4.1 Flash

Feedback ⓘ price cut · Nov 15, 2026

MODEL ↗

DeepSeek V4.1 Flash is a multimodal model from DeepSeek.

Price cut on OpenRouter (deepseek/deepseek-v4.1-flash-1017): $0.30 → $0.15 in, $1.20 → $0.60 out per million tokens (-50%), first seen 15 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

Z.ai GLM 5.1

Feedback ⓘ retirement scheduled · Nov 15, 2026

MODEL ↗

GLM 5.1 is a chat model from Z.ai.

OpenRouter lists a retirement date of 20 Nov 2026 for z-ai/glm-5.1, first seen 15 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

Moonshot AI Kimi K3.1 Pro

Feedback ⓘ

price cut · Nov 15, 2026 MODEL ↗

Kimi K3.1 Pro is a multimodal model from Moonshot AI.

Price cut on OpenRouter (moonshotai/kimi-k3.1-pro): $2.00 → $0.2232 in, $10.00 → $8.4114 out per million tokens (-89%), first seen 15 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

Alibaba Qwen3.5-9B

Feedback ⓘ price cut · Nov 15, 2026

MODEL ↗

Qwen3.5-9B is a multimodal model from Alibaba.

Price cut on OpenRouter (qwen/qwen3.5-9b): $0.17 → $0.073 in, $0.25 → $0.108 out per million tokens (-57%), first seen 15 Nov 2026. Not checked yet

WAS THIS USEFUL? 👍 👎

// END OF TRANSMISSION

The AI Toolchain · curated daily · November 18, 2026
Releases & features from the open-source and commercial AI tools that practitioners actually run.

▷ Subscribe Feedback on this issue