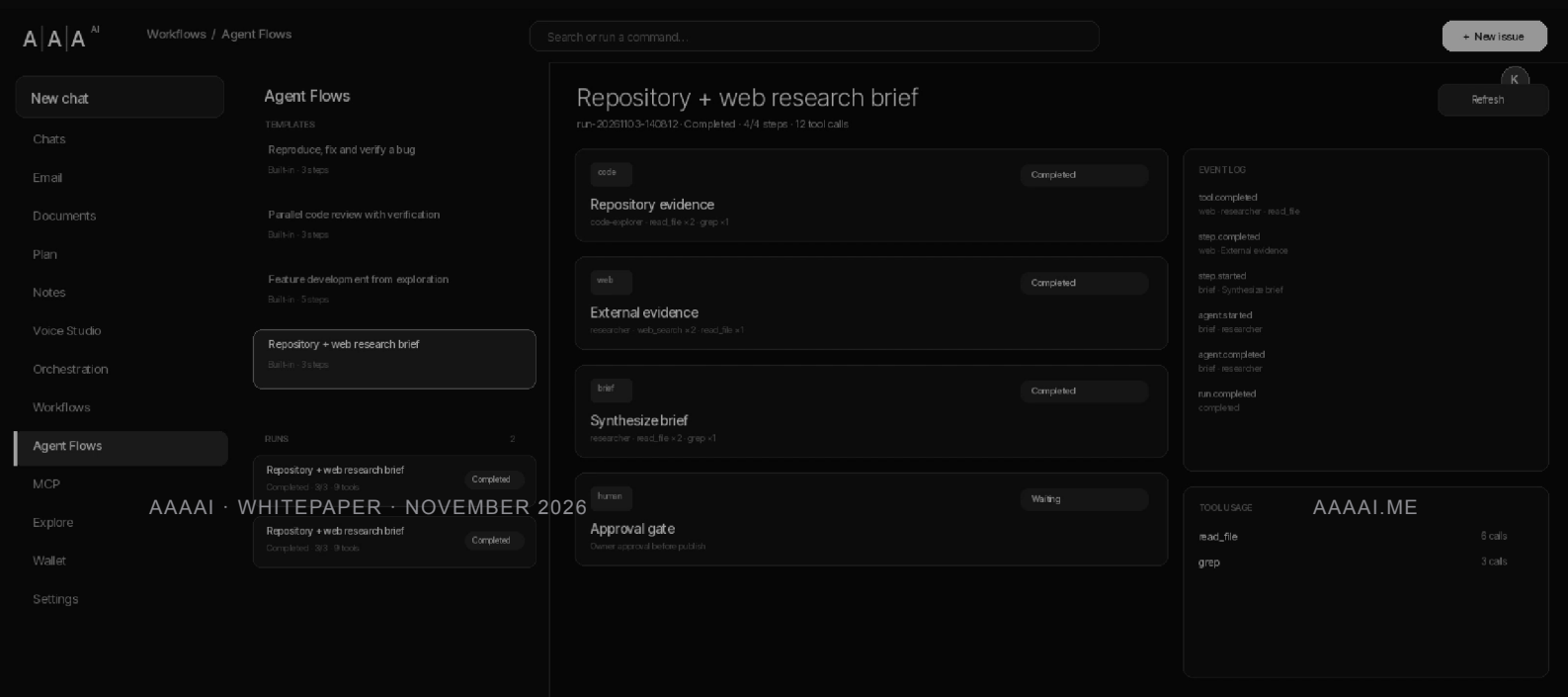
A|A|A AI

PRODUCT WHITEPAPER · AGENT FLOWS

Agent Flows

Durable, resumable multi-agent execution with reviewed YAML, explicit tools, approval gates and an append-only run log.



WHITEPAPER · 02

Reviewed execution, not an opaque session.

Agent Flows is the durable execution layer for multi-agent work. A flow is a YAML graph of steps. Each step runs one or more named agents with an explicit tool allowlist. Every run is persisted with an append-only event log, per-step tool usage, file changes and a run-scoped chat.

Flows are designed for repeatable work: code review, feature development, bug hunting, research and business operations. The spec can be reviewed before it runs and audited after it finishes.

3

surfaces: CLI, macOS IDE and web

1

run directory with the evidence

What is inside

- **YAML flow spec** with inputs, steps, agents, dependencies, prompts and tool allowlists.
- **Approval gates** for steps and risky tools, with ask / auto / strict modes.
- **Durable run state** — run.json, events.jsonl and a spec snapshot for deterministic resume.
- **Run-scoped chat** for evidence, edits and the next step.

> A flow should be inspectable before it runs and auditable after it finishes.

WHITEPAPER · 03

The run model.

Specification

FIELD	PURPOSE
agents	Named agents; sequential chaining or parallel fan-out.
tools	Exact allowlist for the step; empty means all tools or the agent's own list.
approval	Explicit approval before the step, via prompt or --yes.
retries / timeout	Bounded retries and per-step deadline with cancellation.
depends_on	Execution order and data availability for prompt templates.

Run evidence

- **events.jsonl**: run, step, approval, agent, tool and chat events, in order.
- **run.json**: state, steps, agents, tool usage, changes, chat and usage.
- **flow.yaml**: the exact spec snapshot used, so resume is deterministic.
- **Run chat**: a bounded context built from inputs, step outputs, tool usage and recent chat.

Resume and approvals

Completed steps are reused; pending, failed or interrupted steps are rerun. Approvals are explicit; non-interactive runs must pass --yes or the step fails with an actionable message.

WHITEPAPER · 04

Surfaces, scale and limits.

Surfaces

- **CLI** — aaaai flows list/show/run/status/logs/resume/ask; JSON and NDJSON for native clients.
- **macOS IDE** — Flows tab in the Chat rail, Agents window inspector, composer action and full timeline.
- **Web** — Workflows → Agent Flows with run history, live steps, tool usage and run chat.

Scale and persistence

Run/event state is Postgres-backed when configured; bounded execution, incremental reads, per-user retention and distributed cron claims support production use. Run directories are private (0700), and tool allowlists bound what each step can touch.

Known limits

- There is no in-UI flow editor; flows are reviewed YAML files.
- Flows are not yet a visual Workflow node or an orchestration/agent-team tool.
- In-app MLX is not supported for IDE flows; the panel explains why.

> Durable execution is a governance feature, not just an infrastructure one.