

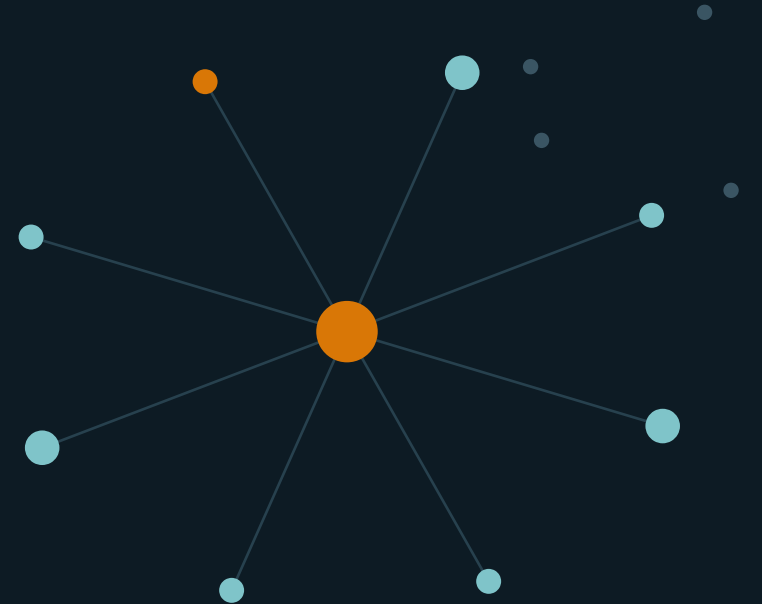
COLONY AI · v24.0.0

Multi-Agent Cognition Runtime

Mentor as a Service — AI agents that remember, govern themselves, and prove what they did.

44 containers running right now · orchestrator :9090 · 467K memory vectors · Society Protocol with receipts

RECEIPT · LIVE FLEET · 2026-10-21



Valentin Kalinin, founder · valentin@mind.glass · github.com/tdesc/atma · Saint Petersburg

THE PROBLEM

LLM agents wake up with no past and no rules.

Anyone can spawn a hundred agents in 2026. Almost nobody can govern them — or trust what they claim.

No memory.

Every session starts from zero. No long-term recall, no continuity, no relationship with the user.

No governance.

No admission gates, no rules of engagement, no separation of powers between agents that act.

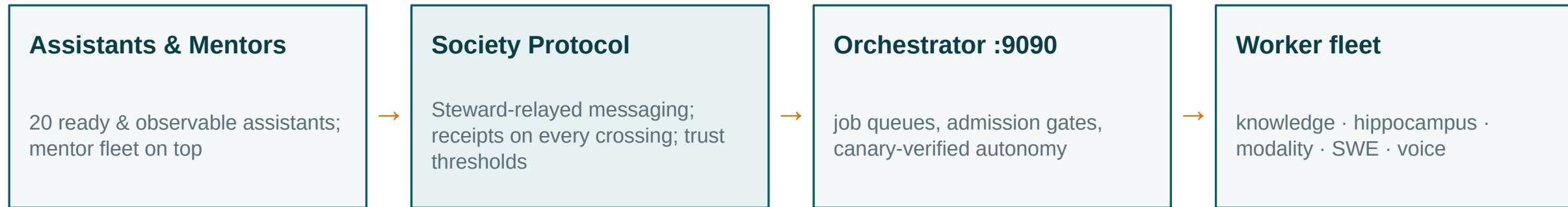
No receipts.

Claims can't be verified. “Trust me” does not scale to fleets — or to customers, or to investors.

The missing layer is not a smarter model. It is a runtime that gives agents memory, government, and a verifiable record.

THE SOLUTION

Colony: one runtime — memory, government, receipts.



Qdrant memory — 467K vectors

· Langfuse traces

· Temporal workflows

· OpenTelemetry

Receipt-based honesty. Every message, arrival, and crossing carries a verifiable receipt — fabrication is a protocol violation, not a bug.

Government, not vibes. A Steward relays all cross-agent traffic; privileged actions need accrued trust; sponsor approval gates the rest.

Memory that persists. Causal hippocampus records episodes; Qdrant recall brings them back in the next session.

Mentor as a Service.

Persistent mentors. Every mentor remembers its students across sessions — vector recall of goals, progress, and history, not chat logs.

Live sessions. Goal-setting, homework, follow-up. Voice-ready: 170 ms turn-taking on the local speech bench.

Honest by construction. Each session closes with a receipt: what was covered, what was decided, what comes next.

A business on the runtime. Subscription mentors running on governed agent infrastructure — memory, governance and receipts included.

GABI

Flagship mentor — live sessions today

REMEMBERS EVERY STUDENT

THE MENTOR FLEET

Tim	Maya
Stas	Dreamweaver
Storyteller	Nathan Gold

Mentor personas are registered assistants in the live fleet — not mockups.

VERTICAL CASE

Dota 2 Copilot: realtime AI inside a live game.

Realtime, not chat. Reads match state and advises while the game runs — decisions in seconds, on a ticking clock.

Same runtime underneath. Per-player match memory, governed tool use, traceable decisions — Colony in a hostile, realtime domain.

A paid lane from day one. Premium tier served over a dedicated MCP server — subscription access to the copilot.

Why it matters. Proof the runtime survives realtime, high-stakes load. If it holds here, it holds anywhere.

LIVE IN THE FLEET

3 services, running now

dota-copilot-portal

frontend + control + backend

dota-premium-mcp

premium access over MCP

Container names from live docker ps — 2026-10-21.

PROOF, NOT PROMISES

The fleet is running right now.

44

containers live

docker ps, without stopping

467K

memory vectors

Qdrant — plus 161K code chunks

20

assistants ready

lawful, kernel-ready, observable

:9090

orchestrator ready

ready_for_programs — canary passed

Workers on duty knowledge (arc · swe · colony · biology) · hippocampus · modality (vision · audio · motor · text) · research · court · admission · SWE · voice · metabolic

Platform Langfuse traces · Temporal workflows · OpenTelemetry · Centrifugo realtime · local MLX model gateway

Source: live docker ps and GET :9090/readyz on the founder's machine, 2026-10-21. Numbers are observed, not aspirational.

Grow the fleet from one box to a cloud.

1

Managed memory & traces

Qdrant and ClickHouse in the cloud — vector memory and Langfuse traces that outlive any single machine.

2

GPU nodes

Voice mentors and model work beyond the local MLX bench — realtime speech needs real GPUs.

3

Public Mentor API

Take Mentor as a Service multi-tenant: pilot customers, billing, SLAs. The monetizing step.

4

Always-on fleet

Multi-region Temporal and 24/7 uptime — live mentor sessions that never sleep.

ROADMAP

From working prototype to hosted runtime.

NOW



Self-funded working prototype —
44-container fleet, memory,
receipts, Society Protocol.

+3 MO



Mentor as a Service pilot —
public API, first external
students, paid sessions.

+6 MO



Cloud fleet — managed memory
& traces, GPU voice mentors,
24/7 uptime.

+12 MO



Governed copilots beyond
gaming — agents you can audit,
as a category.

Honest baseline: no users and no revenue yet. This deck sells a working prototype with live infrastructure — the roadmap is a plan, not a projection.

TEAM & CONTACT

Valentin Kalinin

Founder — built the whole stack solo: cognition runtime, Society Protocol, mentor fleet, Dota 2 Copilot.

valentin@mind.glass

github.com/tdesc/atma

mind.glass · Saint Petersburg, Russia

RECEIPT · SELF-FUNDED WORKING PROTOTYPE · OCT 2026

Colony AI — Mentor as a Service

