

Thomas Crockett

eumemic.ai · GitHub: [eumemic](#) / [pelotom](#)

tom@eumemic.ai

+1-626-864-3289

*I build the infrastructure AI agents run on: runtimes, sandboxes, durable workflows, and the evaluations that keep them honest. Eleven years at NASA JPL building Mars-rover operations software (**NASA Software of the Year Award**); senior engineer at Stripe; author of the open-source agent platforms **AI Legion** (1.4k★) and **aios**.*

SELECTED SYSTEMS

- **AI Legion** — early autonomous multi-agent platform github.com/eumemic/ai-legion
TypeScript · 1,431 stars · 170 forks · open-sourced April 2023
 - **First days of the agent wave:** Released alongside AutoGPT and BabyAGI in the same April-2023 week. From its first commit it had agents messaging each other, plug-in tools, a typed action language, and memory that summarized older conversation to fit a budget. Its ideas grew into aios (below).
- **RagZoom** — agent-memory engine & benchmark study [eumemic/dynamic-summary](#)
Python · repo: [dynamic-summary](#) · negative results published in full
 - **Rigor over wins:** Built a memory engine for long-lived agents (hierarchical summarization) and the evaluation harness to test it fairly. [Published the full results](#), including where a simpler retrieval baseline won (LongMemEval-S). Traced that gap to a single retrieval-policy setting and recovered part of it with one change; retracted an early apparent win when a larger sample reversed it. The takeaway: the bottleneck wasn't finding the evidence — it was getting the model to reason over it.
- **aios** — durable multi-agent runtime github.com/eumemic/aios
MIT-licensed · Python · in production as the substrate of Eumemic LLC (below)
 - **What reinforcement-learning (RL) environments need:** One hardened production runtime provides isolated per-agent sandboxes, complete recording and exact replay of every agent run, repeatable task setup, and many isolated tenants — and is partly maintained by the agents that run on it (the dev pipeline, below).
 - **Event-sourced core:** Runs on Postgres. Instead of a central loop that stalls while a tool runs, every tool call runs in the background, so an agent keeps receiving and acting on new input mid-task — where mainstream agent frameworks block until the tool returns. Every event is kept forever: sessions replay exactly, and agents recall the past by querying the log instead of losing it to summarization.
 - **Durable sandboxes, tenant isolation:** Each agent's container keeps its entire filesystem for months across stops and restarts. The agent runs as root inside its sandbox yet can never read the credentials it authenticates with — secrets are injected outside it. Tenants are isolated by per-account encryption keys, and network egress fails closed.
 - **Durable execution, tool use & MCP:** Durable execution in the style of Temporal: workflows journal every step and replay after a crash or redeploy, spending no model tokens on coordination. One schema generates the REST API, the MCP server (the standard for exposing tools to models), and the client SDK.
- **Eumemic company stack** — a company operated by agents built on aios
Eumemic LLC — my agent-services company — runs near-autonomously on these
 - **Dev pipeline:** A durable aios workflow takes a GitHub issue, implements the change, fixes CI failures, reviews the code, and merges it. Low-risk changes merge autonomously; high-risk ones wait for my approval. In June 2026 it carried an issue to a merged pull request on aios itself with no human in the loop — agents shipping the runtime they run on.
 - **Watchdogs & CEO:** A fleet of scheduled watchdog workflows (stalls, drift, resource pressure) audits the system around the clock, and a conversational CEO agent on Telegram keeps me current on the company.

EMPLOYMENT

- **Eumemic LLC** — *Founder & Principal Engineer* *Remote · 2025 – Present*
 - Built aios and the production stack in Selected Systems above; company operations run near-autonomously on those agents.
- **Stripe** — *Senior Software Engineer* *San Francisco, CA · Jul 2022 – Jan 2025*
 - Shipped an LLM feature that auto-generates credit-card statement descriptors for new merchants during onboarding, cutting data-entry errors 77%. Built the service that shields known-good users from being wrongly flagged by ML fraud models.
- **Tuple Health** — *Senior Software Engineer* *Washington, DC · Oct 2016 – Mar 2021*
 - Owned the company monorepo, developer platform, and CI; built healthcare web interfaces (TypeScript, React).
- **NASA Jet Propulsion Laboratory** — *Software Engineer* *Pasadena, CA · Jun 2005 – Oct 2016*
 - Built Maestro and MSLICE, science-planning tools used in daily operations of the Mars rovers Spirit, Opportunity, and Curiosity. Core contributor to OnSight, the Microsoft HoloLens interface for planning Curiosity activities on a virtual Mars; the project won **NASA's Software of the Year Award**.

OPEN SOURCE — TYPE SYSTEMS

- **runtypes** (2,696 stars, >1.3M npm downloads per month): Runtime validators that derive static TypeScript types — plus [six more TypeScript type-system libraries](#) totaling ~2,200 stars.

SKILLS

agent runtimes & harnesses · sandboxing & trajectory capture · LLM evaluation & graders · multi-agent orchestration · distributed systems · tool use & MCP · Postgres · Python · TypeScript/React

EDUCATION

University of Arizona — B.S. Computer Science, Aug 1999 – Dec 2004