

Joseph Matan

San Francisco, CA · josephmatan@gmail.com · (408) 838-6960 · linkedin.com/in/josephmatan

CTO building systems that combine frontier models with durable context, tools, and execution authority to produce economically useful work. Previously built agentic investment-research workflows at a \$300B AUM manager. Stanford MCS graduate.

EXPERIENCE

Righthand AI (American Productivity Co.)

Sep 2025 – Present

Co-Founder & CTO

San Francisco, CA

Venture-backed startup (Activant Capital) building persistent autonomous AI employees that operate inside real teams with their own computers, memory, email, phone numbers, and software identities.

- Architected an environment for autonomous agents to carry responsibility over months: responding to events, operating computers and SaaS tools, communicating with people, and pursuing follow-ups. Preserved identity and project state across model changes, deploys, and replacement compute.
- Built persistent memory from self-edited files, entity records, semantic retrieval, hierarchical summaries, episodic ledgers, and learned procedures. A production agent resumed a records project **86 days later** honoring the customer's original communication constraints.
- Built persistent multi-agent organizations in production: an AI product manager delegated work, reviewed pull requests, required regression-test corrections, and merged under its own GitHub identity. The manager and one engineer exchanged **1,027 messages over five months**.
- Scaled agents to **411k+ tool invocations**, 53k+ messages/emails, 1.6k+ calls, 36k+ completed reasoning episodes, and 230M+ generated tokens across 108 production/development identities; roughly 98% of tool/model activity remains excluding benchmarks and tests. One assistant operated for **255 days**, handled 5,000+ messages/emails and 684 calls, and persisted through 326 software updates and 511 sandbox instances.
- Designed fault-tolerant execution for identities that outlive compute: generation-fenced execution authority, durable ingress, replay protection, recoverable sandboxes, and explicit "outcome unknown" states to guard against duplicate external actions.
- Built **Human-Bench** to detect behavioral regressions in long-lived agents. A 400-trial experiment found a proposed judgment framework reduced success from **79% to 69%** ($p = 0.0226$) as agents treated enumerated examples as loopholes; the result prevented the change from shipping.

Western Asset Management

Apr 2024 – Sep 2025

AI Engineer

Pasadena, CA

One of three engineers building the AI platform at a \$300B AUM fixed-income manager for investment and research teams.

- Designed parallel teams of research agents for the muni bond trading desk to extract potential alpha across large document datasets where false negatives were unacceptable.
- Built a tool capturing analysts' undocumented research methods through conversation and turning them into reusable multi-agent workflows executable against arbitrary documents.
- Expanded document chat into multi-agent research, scheduled tasks, daily portfolio briefs, and code generation.
- Built an LLM-judge evaluation suite benchmarking application quality against expert outputs; owned AI policy/compliance, internal APIs, and CI/CD for secure enterprise deployment.

Capgemini

Jan 2022 – Apr 2024

Consultant, AI & Analytics

San Francisco, CA

- Architected a GPT-powered RAG application demoed at AWS re:Invent, adapted for legal, compliance, and financial use cases; delivered production time-series forecasting pipelines (Python, Spark, Snowflake/Snowpark).

Independent AI Consulting — Pre-Seed Startups

2022 – 2024

- Built v1 of the visual annotation engine for **Remy** (architectural plan markup) and the scoring engine and AI outreach for **AdviceAnalytics** (wealth-advisor prospecting).

EDUCATION

Stanford University

2016 – 2021

B.S., Mathematical and Computational Science

Stanford, CA

- Four-year varsity athlete, Stanford Men's Rowing; three-time Pac-12 All-Academic; National Hispanic Scholar.
- Research with Prof. Scott Sagan on nuclear-policy survey design and statistical analysis; findings covered in *The Wall Street Journal*, *Foreign Affairs*, and *The New Yorker*.

SKILLS

Languages

Python, Rust, TypeScript, SQL, C/C++

AI Systems

Agent runtimes, Claude/Agent SDK, MCP, multi-agent orchestration, RAG, LLM evaluation, sandboxed execution (E2B)

Infrastructure

Postgres/Supabase, MongoDB, Stripe, Twilio, WorkOS, AWS, Azure, Render, Vercel, Cloudflare R2, CI/CD