

Raghav Ramasamy

San Francisco, CA · 425-505-1657 · raghavcollege@gmail.com · linkedin.com/in/raghav-ramasamy · github.com/RaghavR24

SUMMARY

Founding engineer at Casca (YC S23), leading the platform AI team while forward-deployed at our partner banks. That mix has let me iterate quickly with customers, then turn those learnings into core platform features that stem from real problems and account for real edge cases — an approach I've applied across our underwriting and closing workflows.

EXPERIENCE

Casca (Cascading AI, YC S23 – \$33M raised)

Oct 2024 – Present

Founding AI Engineer · Employee #5

San Francisco, CA

- Co-led the product demo and RFP process, alongside our CEO, in the winning pitch to **50+ executives and lenders at the nation's #2 SBA lender, a \$200B+ national bank** — which chose Casca as its loan origination system and became our largest enterprise customer.
- End-to-end tech lead for Casca's second (and fastest) bank implementation: took a \$10B+ bank and top-5 SBA lender **from first pitch through production go-live**, owning discovery, technical scoping, core-banking integration, and rollout.
- Grew with the company from **5 to 50+ people**, pre-seed through our current funding, helping set engineering practices, onboarding, and a culture of engineers who sit with customers.
- Led core platform AI — from document intake and analysis to agentic orchestration for underwriting and decisioning — serving **real production AI traffic** for live bank customers.
- Spent **5 months embedded on-site** with bank SMEs on the East Coast, sitting with bankers as they worked real loans; became Casca's **internal subject-matter expert** for loan closing and underwriting workflows, and turned what I watched into Casca's roadmap.

Amazon Web Services – S3

May 2024 – Aug 2024

Software Development Engineer Intern

Seattle, WA

- Cut latency on S3's cross-region copy API by integrating regionalized bucket metadata, eliminating the need for replicated bucket state; owned the change from design through stress testing and gamma deployment.
- Finished the primary project early and joined an S3 tiger team, building continuous bucket provisioning that cut testing-iteration time by **70%**, improving DevEx across the team.

Spoony.ai

Jan 2024 – Oct 2024

Bootstrapped side project

Atlanta, GA

- Built an AI POS and menu platform with a friend and sold it to local restaurants door to door. Gave me the **Smiling Lady Test** (the product test I'll always trust) and taught me what it takes to make software people will really use.

Qdeck.ai · Rotella Capital Management

Summers 2022 & 2023

AI / Quant Research Engineer Intern

Bellevue, WA

- Built the LLM pipelines behind Qdeck's AI financial advisor, and designed an eval harness over live market data that raised answer quality **40%** on our internal benchmark.
- Contributed to a patent filing on Interactive Visual Search; supported applied research for systematic trading strategies.

EDUCATION

Georgia Institute of Technology

May 2025

B.S. Computer Science — Threads: Intelligence & Modeling & Simulation

Atlanta, GA

- **4.00 GPA**, degree completed in three years · 4× Faculty Honors.

SKILLS

TypeScript, Python, Java, SQL · Next.js, React, tRPC, Node.js, Prisma, PostgreSQL · LLM engineering: agents & MCP, RAG, embeddings & vector search, self-correcting agent loops, eval harnesses, fine-tuning · PyTorch & model training · Claude Code & Codex as daily production tools · AWS, Docker, Cloudflare Workers, Vercel, Playwright.