

Sidharth Rajmohan

I build RL environments that stump frontier models

India | [LinkedIn](#)

TL;DR: I enjoy finding the limits of frontier models.

I build RL environments and scenarios that stump frontier models. Usually that means long-horizon coding tasks where agents have to preserve constraints, reason across files, recover from broken assumptions, and handle messy real-world systems. I do this at Polymath.

The same pattern runs through my open-source work. A few hours after Andrej Karpathy published autoresearch, I had a PR merged into it: a self-healing loop that lets the agent parse its own stack traces and recover from crashes automatically. Across Google Workspace CLI, NVIDIA NemoClaw, LangChain, LangGraph, and PyTorch, I've worked on the same class of problem : hidden failure modes in systems people are starting to trust: agent crashes, silent tool-calling failures, credential exposure, shell injection, and compiler correctness bugs.

I also like helping people understand how AI can make the things they want to do 10x easier

Interested in frontier AI evals, coding-agent reliability, agent red-teaming, AI infrastructure, and AI education.

EXPERIENCE

Software Engineer

Jun 2026 — Present

Polymath Fellows

- Designed and built multiple RL/SWE evaluation environments for frontier AI coding agents for long horizon tasks including task specs, Dockerized sandboxes, hidden/public tests, and scoring logic
- Validated tasks through oracle/no-op checks and model rollouts to ensure failures reflected genuine reasoning gaps, not flaky tests or verifier ambiguity

Founder

Apr 2025 — Sep 2025

Celer AI

India

At Celer AI, my goal is to solve the #1 cause of doctor burnout: the burden of documentation.

- Building a tool that uses AI to convert spoken voice notes into ready-to-sign medical reports, designed to work with any existing EMR.
- Currently in the pre-launch phase, actively connecting with and learning from medical professionals across India to build the solution they truly need.

Founder

Apr 2025 — Jun 2025

Valurize

India

Designed a secure document processing pipeline with ephemeral storage, ensuring GDPR and SOC II compliance with automatic data purging

Implemented OpenAI integration with custom prompt engineering to deliver comprehensive pitch deck evaluations with 50-point scoring system

Architected and deployed a serverless AI-powered pitch deck analysis system using Cloudflare Workers, achieving 99.9% uptime and sub-second response times

Contributor

Jan 2025 — Present

GitHub

Contributor to Karpathy/autoresearch, Google Workspace CLI, NVIDIA NemoClaw, LangChain, PyTorch, and Kubernetes: security hardening, compiler correctness, and agent reliability across codebases used by millions of developers.

Karpathy/autoresearch: self-healing loop that parses its own stack traces and recovers from OOM and syntax crashes.

Shopify's CEO ran it uninterrupted overnight and got 53% faster rendering from 93 automated commits (merged)

Google Workspace CLI: eliminated a TOCTOU race condition leaking OAuth tokens to ~21,000 users, fixed silent auth failure where credentials never persisted while auth appeared to succeed, added Google Meet integration with deterministic requestId hashing (3 merged PRs)

NVIDIA NemoClaw: enforced 0o600 on session config world-readable by default, fix propagated into 3 downstream forks within 13 hours of merge. Closed two distinct .env exposure windows on sandbox startup and remote deployment (3 merged PRs)

LangChain & LangGraph: fixed silent tool-calling failures across Mistral, Groq, and Ollama causing entire agent workflows to stop on provider switch, across ~500k developers monthly (4 merged PRs)

PyTorch: fixed silent data corruption in Inductor C++ kernels, any codebase mixing OpenMP with OpenCV silently produces wrong training outputs with no error thrown (approved)

Kubernetes: modernizing Device Resource Allocation API across ~3M production clusters (under review)

AI Developer | Software Engineer

Jan 2024 — Present

Freelance

India

Partnering with founders and executives to design and build custom technical solutions that solve expensive problems and create new opportunities.

My work focuses on delivering clear ROI by reducing manual overhead, creating new revenue streams.

REPRESENTATIVE PROJECTS:

Automated client onboarding for Meteoric Boost by wiring DocuSign, Notion, and Discord via webhooks — cut manual overhead by 93%.

Built Node.js scripts for Isotta Homes to parse fee structures from Google Sheets and auto-generate Xero invoices, eliminating a recurring 2-hour manual billing cycle.

Deployed production booking and lead generation pages for property clients (devasramam.live, anpustay.com).

Quantitative Analyst

Jun 2023 — Mar 2024

Self-Employed / Proprietary Trading

Actively managed and executed trading strategies across funded futures platforms, achieving consistent payouts via disciplined price action and risk-controlled discretionary strategies. Conducted advanced research and development into ML-driven quantitative strategies, serving as the foundational work for the subsequent Statistical Arbitrage System.

EDUCATION

Manipal Institute of Technology

Jul 2019 — Aug 2023

Bachelor of Technology in Electronics

CERTIFICATIONS

3rd Place Winner, Global Splunk Build-a-thon, Issued by Splunk

Jul 2025

August Season of #T20DSACHallenge ft. "Stacks & Queues", AlgoUniversity

NISM-Series-VIII: Equity Derivatives Certification, National Institute of Securities Markets (NISM)

Project management (ongoing), Google

Cybersecurity analyst (ongoing), IBM

Marketing in a digital world, Coursera

Digital marketing analytics in practice, Coursera

Digital marketing analytics in theory, Coursera

SKILLS

Analytical Skills · Decision-Making · Critical Thinking · Communication · Problem Solving · Software Infrastructure · Reinforcement Learning · Go (Programming Language) · TypeScript · Rust (Programming Language) · AI Agents · Programming Languages · PyTorch · Prompt Engineering · Software Development · Solution Architecture · IT Consulting · Custom Software Development · Technical Consulting · AI Strategy · Workflow Automation · System Architecture · Full-Stack Development · API Integration · Scalability · Product Management · Artificial Intelligence (AI) · Go-to-Market Strategy · Systems Thinking · Pandas (Software) · MATLAB · TensorFlow · Verilog · Python (Programming Language) · Microsoft Excel · Java · C++ · Visual Basic · Electrical & Instrumentation · Financial Data Analysis · Product Strategy · Equity Derivatives · Equity Trading · Equity Research · Agile Project Management · Cybersecurity · Information Security · Electronic Instrumentation · Product Development · FinTech · Electronics Technology · Entrepreneurship · Investment Advisory Services · Scikit-Learn · Signal Processing