

SUMMARY

Forward-deployed applied AI engineer turning ambiguous operating workflows into production-ready LLM systems from discovery and technical scoping through full-stack build, integrations, evaluation, and production deployment. Builds and ships TypeScript/Python systems, then codifies delivery patterns into reusable platform components.

EDUCATION

Northeastern University, San Jose, CA Sep 2022 - Aug 2025

Master of Science in Computer Science (**GPA: 4.0/4.0**)

University of California, Irvine, Irvine, CA

Sep 2015 - Mar 2019

Bachelor of Arts in Quantitative Economics (**GPA: 3.5/4.0**)

TECHNICAL SKILLS

Programming Languages: TypeScript, Python, JavaScript, SQL, HTML/CSS

Applied AI: Production LLM applications, agent orchestration, tool use/function calling, retrieval/RAG, evals, guardrails, human-in-the-loop, structured outputs, observability

Delivery & Integration: Discovery, technical scoping, solution design, rapid prototyping, production rollout, React/Next.js, REST/WebSocket APIs, OAuth, webhooks

Cloud & Data: AWS, GCP, Docker, Firebase, PostgreSQL, SQLite, Redis, Elasticsearch, DynamoDB

PROFESSIONAL EXPERIENCE

SiriusMindshare LLC, San Jose, CA

Oct 2025 - Present

Machine Learning Engineer | *Next.js, TypeScript, AWS, Vercel AI SDK*

- Scoped ambiguous operating workflows into **5 role-specific agents** and **10 registered tools** across marketing, design, bookkeeping, support, and sales; defined role, context, turn, and permission boundaries.
- Owned end-to-end delivery across solution design, streaming UX, API/tool integrations, human approvals, monitoring, and production debugging in Next.js, TypeScript, AWS, and LLM APIs.
- Codified a reusable context-engineering pattern that separated **approximately 4K static tokens** from dynamic state and compressed sessions through **4 stages at 80% capacity**, bounding repeated processing and long-session growth.
- Co-authored a paper on a generative AI/LVLM framework for retail attention-quality assessment, accepted for publication at **IEEE CCET 2026**; extended the platform's evaluation approach to multimodal workflows (forthcoming proceedings).

1 Thing Against Racism Organization, Monroe, NJ (Remote)

Jan 2025 - Apr 2025

Full-Stack Software Developer Intern | *Python, Flask, Firebase, GCP*

- Delivered Flask/GCP incident-management APIs, real-time notifications, and Twitter/LinkedIn publishing integrations; reduced frequent-query latency by **40%** through server-side caching for recurring dashboard requests.

1 Thing Against Racism Organization, Monroe, NJ (Remote)

Jan 2024 - Apr 2024

Full-Stack Software Developer Intern | *TypeScript, React, React Native, Firebase*

- Delivered React/React Native clients with Redux/Firestore synchronization, OAuth, encryption, and AI-assisted moderation across **2 mobile platforms** (iOS and Android), supporting synchronized content and moderation workflows.

PROJECTS

Agent Hook | *Python, Webhooks, SQLite, Claude/Codex CLIs*

2026

- Built a local-first event-to-agent integration layer across **6 trigger sources** - cron, webhook, GitHub, filesystem, failsafe, and beacon - with HMAC-signed webhooks and asynchronous 202 dispatch.
- Added capability whitelists, human approvals, debounce/rate/concurrency controls, retries, graceful shutdown, replay, SQLite lineage, and **5 prompt/context/token metrics**; validated with **279 tests**.

In-Context Search | *Chrome Extension, JavaScript, LLM APIs*

2026

- Delivered select-to-explain and in-page search across **6 AI/search providers** and **3 browser surfaces**; kept API keys in Chrome sync storage and validated critical flows with **32 Playwright tests**.

AgentFlow | *TypeScript, React, Hono, WebSocket*

2026

- Delivered a full-stack agent workflow product with React Flow visual authoring, multi-agent execution, human approvals, REST/WebSocket APIs, provider-neutral LLM integrations, and **1,670+ unit/e2e tests**.

Agent Bridge | *Rust, Python, TypeScript, Docker, AWS*

2026

- Standardized **5 sandbox lifecycle operations** across **8 provider backends** through a Rust CLI, Python/TypeScript SDKs, and **5 MCP tools**; added capability discovery, structured errors, and fail-closed security/cost controls.

AI Diary Companion | *Next.js, TypeScript, Fastify, Prisma, SQLite*

2026

- Delivered a full-stack AI journaling product with **6 persona companions**, SSE-streamed chat, structured diary generation, and importance/recency-ranked memory with automated decay and consolidation.
- Designed **6 Fastify/Prisma API domains** spanning auth, chat, diaries, memory, reminders, and custom characters; validated backend APIs and Next.js workflows with **170+ unit/e2e tests** across **5 test suites**.

Additional open source: *jsfetch (Python Playwright CLI/MCP for JS-rendered web extraction)* | *Kaggle Team (3-agent, 4-phase loop)*