S Aditya Reddy

+91-85229 89043 aditya.singireddy.21@gmail.com
github.com/Atrix21 atrixhq.dev

Professional Summary

Software Engineer focused on architecting and building production-grade software systems, distributed backends, and AI-powered applications. Experienced in multi-agent orchestration, retrieval systems, concurrent processing, and native product development. Independently designs and delivers end-to-end systems, with an emphasis on reliability, scalability, and practical AI integration.

Experience

AidenAI May 2025 – Present

Product Engineer

- Architected and delivered Hyrra, a production multi-agent recruitment platform for TASC using AutoGen, FastMCP, and Azure OpenAI. Designed agent orchestration and role-based workflows for six user types. The platform was handed over to TASC and is used in its everyday recruitment process.
- Designed asynchronous agent orchestration and parallel tool execution, reducing multi-step workflow latency by approximately 60–70% while supporting 50+ concurrent users.
- Refactored a 6.8K-line monolithic tool implementation into a modular architecture comprising 30+ modules and 80+ tools, improving maintainability and development iteration speed.
- Implemented Server-Sent Events (SSE) for incremental token delivery, reducing perceived response latency from 5–15 seconds to under 500ms.
- Optimized LLM context construction and retrieval workflows, reducing token consumption by approximately 50%. Improved external API reliability through retries, validation, and backoff strategies.

Selected Products & Engineering Research

Sojourner – Native Health & Fitness Platform *Swift, SwiftUI, Kotlin, Jetpack Compose*

- Independently architecting and developing a cross-platform health and fitness application for iOS and Android, combining workout planning, progression tracking, nutrition workflows, food recognition, and barcode-based food logging.
- Built native integrations including HealthKit, Apple Watch connectivity, Live Activities, and on-device Apple Foundation Models. Designed privacy-first and offline-oriented application workflows.
- Developed a supporting growth and analytics platform using Cloudflare Workers and D1, with campaign attribution, conversion tracking, search measurement, and content publishing workflows. iOS is in open beta; Android is undergoing closed testing.

TypeTrace – Typed Code Graph & Search for AI Coding Agents *TypeScript, Python, MCP*

- Independently designed and built a code intelligence tool that indexes TypeScript codebases with the compiler (typed calls, dependency injection, GraphQL resolvers, event dispatch), combines the graph with BM25 and local embeddings, and serves coding agents a single MCP tool capped at 800 tokens per answer.
- Benchmarked it head-to-head against graphify and code-review-graph on 64 audited tasks across three repositories (6K–763K lines), with pre-registered tests and frozen, tool-independent ground truth. Achieved Recall@10 of 0.33–0.69 against 0.10–0.31 for the strongest competing setup.
- In 216 automatically scored agent runs on a 763K-line codebase, agents used 23% fewer tokens per task than with graphify ($p = 0.003$) and half as many as with code-review-graph. Accuracy differences were not statistically significant and are reported as such.

Technical Skills

Languages: Python, Go, TypeScript, Swift, Kotlin, SQL

Backend & Architecture: Distributed Systems, Async Programming, Concurrency, Event-Driven Architecture, REST APIs, SSE, System Design

AI & ML Systems: LLM Orchestration, Multi-Agent Systems, AutoGen, FastMCP, RAG, Vector Search, Embeddings, Qdrant, FAISS, SentenceTransformers, Azure OpenAI

Infrastructure & Data: PostgreSQL, Redis, Docker, Azure, Cloudflare Workers, D1, CI/CD, Linux

Certifications

CCAR-P – Anthropic Verify Credential

Education

Keshav Memorial Institute of Technology 2021 – 2025

B.Tech in Computer Science and Engineering | CGPA: 7.9/10 *Hyderabad, India*