Vicky Kumar

+918383848219 | npdimagine@gmail.com | https://www.linkedin.com/in/algsoch/ | https://github.com/algsoch | https://github.com/fiscalmindset

Work Experience

Outlier, Mercor, Soul AI, SME Careers · Remote | ***Freelance AI Trainer & LLM Evaluation*** **Sep 2024 - Feb 2026**

- Evaluated and ranked thousands of AI-generated responses across math, programming, and logic tasks against structured rubrics - scoring correctness, coherence, instruction-following, and reasoning quality.
- Surfaced model failure patterns and authored structured prompts feeding RLHF and supervised fine-tuning pipelines for enterprise LLM clients.
- Validated factual accuracy and intent alignment of prompt/response pairs, applying consistent criteria used to build model preference and training data.

Education

Indian Institute of Technology Madras **May 2024 - Dec 2024**
Foundational Certificate in Programming and Data Science · Dec 2024 (GPA: 8) *Chennai, Tamil Nadu*

Skills

- **Languages & Databases:** Python, TypeScript, JavaScript, Kotlin, Rust, SQL, PostgreSQL
- **LLM & Agents:** LangChain, LangGraph, function-calling agents, multi-agent systems, prompt engineering, OpenAI / Anthropic / Groq / Gemini / Ollama APIs
- **RAG & Retrieval:** ChromaDB, vector search, hybrid (dense + keyword) retrieval, nomic-embed embeddings, chunking
- **On-Device & Multimodal AI:** RunAnywhere SDK, WebAssembly, ONNX Runtime, llama.cpp, Whisper (STT), TTS, vision models
- **Evaluation & Interpretability:** LLM evaluation, RLHF, rubric-based scoring, hallucination analysis, PyTorch activation tracing
- **Backend, Frontend & Mobile:** FastAPI, REST APIs, React, Next.js, Vite, Tailwind CSS, Android, Jetpack Compose
- **Infra & Tools:** Docker, Git, GitHub Actions, Render, Coral SQL, Intel TDX / Terminal 3, Playwright, FFmpeg, Vitest, pytest

Projects

Blindfold - Confidential-Compute Security Layer for AI Agents | github.com/FiscalMindset/Blindfold

- Makes an AI agent's API keys unleakable to prompt injection by sealing them in an Intel TDX enclave and substituting them into outbound requests *inside* the enclave - the agent never holds the real secret. Authored the Rust→WASM enclave contract + a TypeScript SDK/proxy for one-line adoption (zero code change) across OpenAI, Anthropic, and LangChain clients.
- Shipped end-to-end on Terminal 3 testnet; concept validated directly by Terminal 3's Product Manager.
- Rust, WebAssembly, Intel TDX, Terminal 3, TypeScript

Personal RAG Assistant - Agentic Q&A Over Your Own Codebase | https://github.com/FiscalMindset/vickykumar

- Deep-indexes a user's GitHub repos (code + READMEs), blogs, and profile into a vector store, answering grounded questions via a function-calling agent loop; hybrid retrieval (dense + keyword) plus a webhook that auto-reindexes on every push.
- Multi-tier fallback - Groq → Ollama → keyless on-device RunAnywhere WASM in-browser - runs with zero credentials. Deployed on Render.
- FastAPI, React, Groq, Ollama, RunAnywhere SDK, RAG

CareOps - Coral-Powered Care Coordination Agent | github.com/FiscalMindset/careops

- Unifies 9 disparate record sources (lab PDFs, prescriptions, doctor chats, receipts, symptom logs) through Coral's SQL layer into a doctor-ready visit packet; authored 9 custom Coral source specs with cross-source JOINs and safety guardrails.
- **Top 50 showcase**, Coral × WeMakeDevs Hackathon (Track 2). - Coral SQL, Next.js, TypeScript

Algsoch News - Multi-Agent AI Newsroom | https://algsochnews-1.onrender.com/

- 5-agent pipeline converting a news URL into a broadcast screenplay and rendered MP4 (extraction → editing → QA retry routing → video). - FastAPI, LangGraph, React, Gemini, FFmpeg, TTS

Synapse-Graph - AI Interpretability & Hallucination Analysis | https://github.com/FiscalMindset/Synapse-Graph

- Engineered an AI autopsy engine focused on mechanistic interpretability, utilizing activation tracing, token-output analysis, and OpenMetadata for governance tracking.
- Implemented custom PyTorch hooks to capture and analyze activations at the attention-head level, enabling granular inspection of internal model response patterns.
- Executed targeted hallucination analysis via runtime control and intervention techniques, establishing and documenting the practical engineering limits of neural traceability.
- Python, PyTorch, FastAPI, OpenMetadata

Algsoch - Offline On-Device AI Assistant for Students | https://github.com/FiscalMindset/algsoch

- 1. Built an offline-first Android AI assistant that runs open-source models directly on-device using the RunAnywhere SDK.
- 1. Implemented multiple AI modes including Explain, Direct, Answer, Code, Directive, Theory, and Creative-style responses using mode-specific prompting.
- * Added vision-based question answering, custom AI personalities, companion chat, model status tracking, chat history, and local interaction analysis.
- * Kotlin, Jetpack Compose, RunAnywhere SDK

SpeakAI - Local-First English Speaking Practice Platform | https://speakai-af1l.onrender.com/

- Built a browser-based English speaking practice platform with voice and chat-based AI interaction.
- Designed mode-based speaking practice where users can select different AI personas and conversation styles.
- Integrated speech-to-text, text-to-speech, local-first AI workflows, and optional cloud fallback for improved reliability.
- React, TypeScript, WebAssembly, Whisper ONNX, TTS, Groq API

ACHIEVEMENTS & OPEN SOURCE

- **Open Source:**20 source specs contributed to Coral () - 12 merged, rest approved; also authored , an open-source evidence-grounded technical-writing skill for AI coding agents
- **Coral × WeMakeDevs Hackathon:**1. Top 50 showcase (Track 2)
- **Terminal 3 Hackathon:**built Blindfold; concept validated directly by T3's product manager
- **GitHub:**1. 100+ repositories, 350+ contributions, Pull Shark (22+ merged PRs).