

Anurag Pappula

+91 9121767554 | anuragp2003b@gmail.com | github.com/anuragp22 | linkedin.com/in/anuragp2622

EDUCATION

Indian Institute of Technology Kharagpur

B.Tech, Industrial and Systems Engineering

Kharagpur, India

2022 – 2026

EXPERIENCE

Software Engineering Intern, Neosophical Labs

Feb'26 – May'26

- Built the **schema-matching engine** with alias-learning, fuzzy tiers, and type validation, cutting LLM-bound columns to **35%**
- Grounded LLM column mapping with a **hybrid RAG retriever** over 4.5k schema embeddings, fusing dense and **BM25** via RRF
- Developed a **transactional importer** that writes all migrated data atomically and reconciles row counts to catch silent data loss
- Shipped a **batch video-generation** pipeline behind a 4-provider adapter layer and crash-safe retries that prevent double-billing

Full Stack Developer Intern, Destini AI

Jun'25 – Aug'25

- Built **per-user TTL cache** for 9 prediction types via MongoDB compound indexes and request lock, cutting API calls by **70%**
- Cut Node-to-Python inter-service overhead via Azure Redis, swapping 250KB horoscope payloads for **key-reference lookups**
- Built GitHub Actions **CI/CD** shipping Dockerized microservices behind Nginx into the Azure App Service test and prod slots
- Coordinated with a team of 6 across **20+ PRs**, resolving cross-branch conflicts and promoting tested builds into production

Software Engineer (Freelance), Searchlook

May'25 – Jun'25

- Automated developer onboarding with **Terraform**, provisioning a per-dev **OpenVPN** gateway and scoped IAM CLI credentials
- Built **S3-backed PDF archival** with MongoDB indexing, enabling deduplicated report storage and reuse across audit workflows
- Integrated WhatsApp authorization via Twilio into legal verification workflows with server-side checks alongside SMS & email

PROJECTS

TraceScope – PR Blast-Radius Analyzer · github.com/Anuragp22/TraceScope

Mar'26 – May'26

- Built a **Go CLI** that maps any **pull-request diff** to its changed functions and ranks the downstream code most likely to break
- Derived the **dependency graph** by resolving **SCIP** symbol occurrences into typed cross-file call, import, and inheritance edges
- Implemented a downstream-impact traversal as a **reverse BFS** across the call graph, scoring risk by depth, fan-in, and exports
- Shipped it as a **CI gate** with risk-tiered exit codes and a GitHub PR comment that flags high-risk functions and their owners

SynthFlow – AI Music Generation Platform · synth-flow-ai.vercel.app

Apr'25 – May'25

- Engineered a **text-to-music pipeline** orchestrating audio, LLM, and image models on a GPU into a 3-min song with cover art
- Decoupled GPU inference from request lifecycle onto Inngest queues with **per-user concurrency** limits that prevent starvation
- Built **genre auto-tagging** that classifies every track with the LLM and dedup-persists categories that power the discovery feed
- Designed a **credit system** that reserves a credit with an atomic compare-and-decrement, refunding it whenever generation fails

JobVoice – Real-Time AI Interview Simulator · interview-assistant-nu.vercel.app

Feb'25 – Mar'25

- Built **3-agent voice panel** unified by **context-carrying handoff**, with in-session RAG fact-checking the CV and JD claims live
- Hardened against **prompt injection** with deterministic tool-call guards and prompt-leak detection, gated by a **50-prompt audit**
- Instrumented **end-to-end OpenTelemetry tracing**, propagating a **W3C traceparent** across Next.js, Firestore, and voice agent
- Engineered an **LLM eval harness** gating CI on quality drift, with **p95 latency budgets** per stage and per-session cost telemetry

ACHIEVEMENTS

Smart India Hackathon Grand Finale 2024 | Bharat Electronics Limited | doritos-cb.vercel.app

Oct'24 – Dec'24

PS: build a conversational chatbot that recognizes objects in a user-uploaded image and answers any questions about it in natural language.

- Secured **Top 1.11%** building a **multimodal RAG chatbot** (text + image) that grounds every answer in a cited source passage
- Engineered **hybrid retrieval** over pgvector vector search and Postgres full-text search, **RRF**-fused and **cross-encoder** reranked
- Built a **streaming** ingestion pipeline that extracts, chunks, and embeds large document uploads per page in **bounded memory**
- Integrated Meta **SAM2** on a serverless GPU for **zero-shot** object cutout via point or box prompts, fed to the vision model

OpenSoft General Championship 2024 | Gold Medal, IIT Kharagpur

Mar'24 – Apr'24

PS: build a movie-database powered by MongoDB Atlas autocomplete, fuzzy, and semantic search plus adaptive-resolution video playback.

- Built a hybrid search engine on **MongoDB Atlas** fusing Full-Text, **fuzzy**, and BERT Vector Search via RRF with autocomplete
- Engineered a video-streaming platform on **VidStack** with **HLS** adaptive playback, speed controls, and **Alan AI** voice navigation

TECHNICAL SKILLS

Languages: C++ | Python | Go | JavaScript **Frameworks:** Next | React | Node | Express | FastAPI | Django | Prisma | Tailwind

Databases: PostgreSQL | Neon | MongoDB | Redis | Firebase | Supabase **Cloud:** EC2 | ECR | IAM | S3 | ACR | App Service

DevOps: Docker | Nginx | CI/CD | Git | GitHub **Tools:** OpenTelemetry | Inngest | Twilio | LiveKit | LangChain | LangGraph

Coursework: Programming and Data Structures | Algorithm Design and Analysis | Software Engineering and Management

MOOCs: TUF plus DSA | GFG Core CS Subjects for Interview Prep | Coursera Supervised ML: Regression and Classification

ACTIVITIES

- Contributed bug fixes to popular open-source projects, including **Apache Airflow** (46k ☆), **Puter** (42k ☆), and **Kestra** (27k ☆)
- At E-Cell IIT Kharagpur, headed Local Startups Meet (LSM) Chennai 2023, connecting 90+ startups with 10+ active investors
- Built a Django-based web portal at E-Cell IIT KGP handling registrations and event data management for flagship initiatives