Renish Aryan

AI Engineer | Backend Software Engineer
+91-7575865081 renisharyan2000@gmail.com
linkedin.com/in/renish-aryan-84600a191 github.com/renish04

EDUCATION

Vellore Institute of Technology **2023 – 2027**
B.Tech in Computer Science and Engineering (AI & ML) — CGPA: 8.75/10 *Bhopal, India*

EXPERIENCE

AI Engineer Intern **May 2025 – January 2026**
Ywork.ai *[Letter of Recommendation]* *Remote, London, UK*

- Built **document-intelligence pipelines** on **AWS Lambda** combining OCR with LLM-based extraction, converting unstructured construction PDFs (quotes, invoices, spec sheets) into schema-validated JSON for downstream services.
- Built a **tool-calling folder-management AI agent** that reorganizes uploads into a clean hierarchy — embedding files into **Weaviate**, retrieving related files by similarity, and having an **LLM** group them by content, filenames and structure; new synced files are placed automatically, and chat corrections are applied via **tool calls**, retaining context.
- Engineered the **RAG** layer over a multi-tenant **Weaviate** corpus — chunking, embedding and semantic retrieval — powering a real-time AI chat over client project data.
- Integrated the **Google Maps API** to surface nearby suppliers for a location, letting users shortlist local vendors and request quotations in-platform.
- Built **event-driven** file-ingestion **REST APIs in Django** (**webhook** and polling pipelines over **AWS S3**) keeping cloud files synchronized, and architected **OAuth 2.0 authorization** in **Django REST Framework** across **3 cloud providers** (Google Drive, OneDrive, Dropbox), handling token refresh and per-provider scope differences.

PROJECTS

Clarivo — Autonomous Accounts Payable | *LLM Agents, Hybrid RAG, Document AI, AWS* *2026*

- Built an **agentic accounts-payable system** for **three-way matching** (PO–GRN–invoice) and **exception resolution** — catching **error classes like** (rate, quantity, tax, duplicate), and **autonomously closing** explainable exceptions.
- Engineered **real-time email ingestion** via **Gmail API** and **Google Cloud Pub/Sub** push notifications over **OAuth 2.0** — a mailbox watch triggers a webhook that streams invoice attachments into **S3** as mail arrives, replacing polling.
- Architected an **event-driven serverless pipeline**: S3 events trigger a containerized **AWS Lambda deployed** as a **Docker** image on **ECR** (to fit OCR binaries past zip limits) routing PDFs to native parsing and scans to **Tesseract OCR**, stored in **DynamoDB**.
- Replaced brittle field-matching with **multi-document LLM reasoning**: role classification, evidence via **hybrid BM25 + dense search** on **Weaviate** under a token budget, and a **grounding layer** checking every cited figure against its source.
- Modelled documents as a **state machine** with **event-driven re-evaluation**: new deliveries re-open closed invoices so the agent auto-resolves split-shipment overbills; via **DRF** with persistent **case memory**, audit trails and a **RAG copilot**.

HindSite | *FastAPI, PostgreSQL, pgvector, Cohere, Chrome Extension* *2026*

- Built **AI-powered semantic search** over personal browser history, retrieving pages from natural-language queries such as *"that ML article from last week"*; **engagement-based capture** (reading time, scroll depth) indexes only pages actually read, not every visited URL.
- Implemented **two-stage retrieval**: **pgvector** similarity search over 1024-dimensional Cohere embeddings, followed by **cross-encoder reranking** of the shortlisted candidates.
- Designed a **query router** classifying intent between semantic search and live tab-switching, enabling instant navigation across open tabs by voice or text.

TECHNICAL SKILLS

Languages & Frameworks: Python, SQL, Django, Django REST Framework, FastAPI
Backend & System Design: REST API Design, Event-Driven Architecture, Serverless Architecture, Multi-Tenant Systems, WebSockets, Webhooks, Pub/Sub Messaging, OAuth 2.0 Authentication & Authorization, LLD, HLD
Core CS: Data Structures & Algorithms, Object-Oriented Programming, DBMS, Operating Systems, Computer Networks
Databases: PostgreSQL, DynamoDB, Weaviate, pgvector, ChromaDB
Cloud & Tools: AWS (Lambda, S3, DynamoDB, ECR), GCP Pub/Sub, Docker, Git, GitHub, Postman (API testing)
AI/ML: Machine Learning, LLM Agents, Tool Calling, Structured Outputs, LLM Evals, RAG, Hybrid Search, Semantic Search, Vector Embeddings, Document Intelligence, OCR, LangChain, Prompt Engineering, Gemini, OpenAI, Cohere