Tiju Raju Lukose

AI Engineer | LLM Agents, RAG and Evaluation | Python, FastAPI, AWS
Navi Mumbai, India (open to Pune or remote) | +91-9175301750 | tijulukose0402@gmail.com
linkedin.com/in/tiju-lukose | github.com/tiju07 | leetcode.com/tijulukose

SUMMARY

Software engineer with 2+ years of experience at Hexaware, now building agentic AI systems for financial services clients: LangGraph and MCP agents, tool calling, evaluation, guardrails and observability, backed by the FastAPI services, CI/CD and AWS infrastructure that run them.

EXPERIENCE

Hexaware Technologies, Navi Mumbai, India May 2024 - Present
Software Engineer-2 (Jul 2026 - Present), Software Engineer-1 (May 2025 - Jul 2026), Associate Software Engineer (Sep 2024 - May 2025), Premier Graduate Engineer Trainee (May 2024 - Sep 2024)

- Unified OpenAI, Azure OpenAI and Hugging Face models behind one service layer with retries, fallbacks and cost-aware routing, and added Pinecone and FAISS retrieval over 25,000 documents, cutting token spend per query by 40%.
- Built FastAPI microservices for RAG document search used by financial services clients: 50 endpoints across 5 services, typed end to end with Pydantic, serving 70,000 requests a day.
- Cut API latency by 13% and PostgreSQL (RDS) read load by 28% with async I/O, query tuning and a Redis cache layer.
- Shipped client-facing React and Next.js apps with server-side rendering and a typed API client, bringing first contentful paint down to 1.2 seconds.
- Ran containerised services on AWS ECS/Fargate and EC2 behind an Application Load Balancer with Auto Scaling, least-privilege IAM and CloudWatch alarms, holding 99.9% uptime through 3x traffic spikes.
- Automated lint, test, build and deploy in GitHub Actions, cutting release time from 3 days to 15 minutes with 95% test coverage and zero-downtime rolling releases behind API Gateway.
- Led code reviews and mentored 2 junior developers to independent feature ownership.

PROJECTS

- **Cashless Claims Copilot for Indian Hospitals** | *FastAPI, LangGraph, PostgreSQL, Next.js, AWS*
  Built an AI assistant for a hospital insurance desk that turns admission and discharge records into insurer-ready cashless claims, predicts deductions before submission and answers insurer queries with cited evidence. A person approves every action on an insurer portal.
- **Data Integrity Copilot for Pharma QC Labs** | *FastAPI, LangGraph, PostgreSQL, React, Kubernetes*
  Built an investigation assistant for pharmaceutical quality labs that flags data integrity red flags in audit trails, drafts deviation reports with cited evidence and helps prepare for FDA inspections. Designed to run on premises with open models.
- **Indic PII Shield and DPDP Rights Desk** | *FastAPI, MCP, PostgreSQL, Kubernetes, Azure*
  Built an API, SDK and MCP server that finds and redacts Indian personal data in code-mixed text before it reaches an LLM, plus an agent that handles DPDP access, correction and erasure requests.

SKILLS

Languages: Python, TypeScript, JavaScript, SQL
AI and LLMs: LLM APIs (OpenAI, Azure OpenAI, Anthropic Claude), Hugging Face, prompt and context engineering, structured outputs, tool and function calling, RAG, embeddings, vector search (pgvector, Pinecone, FAISS), LLM evaluation
Agentic AI: LangGraph, LangChain, multi-agent systems, Model Context Protocol (MCP), agent memory and state, human-in-the-loop, guardrails, agent testing, observability and tracing (LangSmith), CrewAI
Backend: FastAPI, Pydantic, REST APIs, microservices, async I/O, Node.js, PostgreSQL, Redis, MongoDB
Frontend: React, Next.js (server-side rendering)
Cloud and DevOps: AWS (ECS, Fargate, EC2, Lambda, S3, RDS, API Gateway, IAM, CloudWatch, SQS), Microsoft Azure, Docker, Kubernetes, GitHub Actions, CI/CD, Linux, Git

AWARDS

- Winner, Shark Tank for GenAI in Financial Services hackathon, Hexaware Technologies (2024)
- Winner, EcoFuelers hackathon, Hexaware Technologies (2025)
- Solved 400+ algorithmic problems on LeetCode, GeeksforGeeks and Codeforces

EDUCATION

Indira College of Engineering and Management, Pune CGPA: 8.83 / 10
Bachelor of Engineering (B.E.), Computer Science and Engineering