

Ganesh Kumar Mangla

AI Full Stack Engineer

ganeshmangla2003@gmail.com | +91-7042317518 | [Linkedin](#) | [Portfolio](#) | Gurugram, India

Professional Summary

AI engineer who ships Agentic LLM systems to production. Delivered four Generative AI systems in a single year at Programming.com - enterprise, healthcare and ecommerce agents plus internal QA tooling - including a custom MCP (Model Context Protocol) server built from scratch over Epicor ERP. Full stack background in Python, FastAPI, React and Next.js, and I own the path to production: Docker, CI/CD across AWS and Azure.

Technical Skills

AI/GenAI Engineering: Agentic AI, Multi agent orchestration, Retrieval Augmented Generation (RAG), Hybrid search, Chunking, Reranking, Embeddings, MCP (Model Context Protocol) development, Tool calling, Agent Memory, Prompt Engineering, Guardrails

Frameworks & Libraries: LangChain, LangGraph, FastAPI, WebSockets, React.js, Next.js, Node.js, Express.js

Models & Platforms: Amazon Bedrock, Azure OpenAI, Anthropic Claude, OpenAI, Gemini, Groq, Ollama

Cloud & Devops: AWS, Azure, Docker, GitHub Actions, Nginx, SSL, CI/CD, Git

Data & Retrieval: Amazon OpenSearch, Azure AI Search, Amazon S3, PostgreSQL, MongoDB, MySQL, Supabase

Work Experience

Associate AI Engineer | Programming.com Aug 2025 – Present

- Shipped four production GenAI systems: Enterprise, Healthcare and Ecommerce agents plus an agentic LangChain QA workflow generating test cases from BRDs and codebases, cutting manual authoring ~ 70%..
- Owned CI/CD across AWS and Azure - Docker, GitHub Actions, Nginx, SSL with zero-downtime deploys.

Full Stack Developer | LetsAI Solutions Pvt. Ltd. Jan 2025 – Jul 2025

- Designed an AI research agent for an EB1A visa portal that gathered applicant achievements via web search into automated petition drafting, anchored to retrieved evidence rather than model recall.
- Implemented RBAC over sensitive immigration records and PII; modeled petition data in Strapi with REST APIs.

Software Developer Intern | Torit Innovations Pvt. Ltd. Aug 2024 – Dec 2024

- Built LLM orchestration for a Medical Science Liaison platform, coordinating workflows that synthesized scientific literature into briefings field teams took into physician conversations.
- Developed a Frappe/ERPNext and Next.js integration, building custom DocTypes, scripts, and REST APIs to bridge ERP data with a modern frontend.
- Managed full server infrastructure on DigitalOcean including CI/CD pipelines via GitHub Actions, Nginx reverse proxy configuration, SSL setup.

Technical Projects

LOBO – Enterprise Data Agent | Epicor ERP, MCP, LangGraph, Langchain, RAG, Azure OpenAI & AI Search

- Architected a custom **MCP server** from scratch over Epicor ERP with role scoped tool access and schema validated tool contracts, replacing multi screen ERP navigation with one governed conversational interface.
- Orchestrated multi agent workflows with LangGraph using state graphs.
- Built hybrid-search retrieval on Azure AI Search such that users could verify the ERP data behind an answer.

RAY – Healthcare Agent | AWS, Amazon S3, Amazon OpenSearch, Athena health, FastAPI, WebSockets, Guardrails

- Built the retrieval layer over patient documents in Amazon S3 ingestion, chunking and embedding into Amazon OpenSearch so clinical questions ran against source records rather than manual file store searching.
- Unified document retrieval with athenahealth EHR tools, enabling a single agent to answer across live patient records.
- Applied Guardrails and RBAC to FastAPI/WebSocket outputs, with chat memory persisting context across sessions.

KRIS – Ecommerce Support Chatbot | LangChain, LangGraph, RAG, Amazon Bedrock, FastAPI, Guardrails

- Built a conversational product discovery agent over client catalogs with LangGraph in an orchestrator worker architecture, carrying shoppers from natural language intent through to checkout redirect.
- Implemented a hybrid search RAG pipeline on Amazon Bedrock and OpenSearch for FAQs.
- Grounded availability and price responses in live catalog data, preventing unsupported stock or pricing claims.

Education

B.Tech in Computer Engineering 2021–2025
J.C. Bose University of Science and Technology, YMCA, Faridabad CGPA: 8.98/10

Class XII (CBSE) 2021
Lions Public School 96.6%

Class X (CBSE) 2019
Ryan International School 95.8%