

# Niranjan Sankeshwari *Applied AI Engineer*

niranjansankeshwari9@gmail.com • +91 9606694256 • Bengaluru, India • linkedin.com/in/niranjans8 • github.com/NiranjanS8

## SUMMARY

---

AI Engineer focused on RAG systems, LLM applications, and agentic workflows using LangChain, LangGraph, FastAPI, and vector databases. Built multi-component AI systems with retrieval pipelines, evaluation workflows, semantic search, and safety guardrails. Strong foundation in backend engineering, API development, and production-oriented AI architecture.

## PROJECTS

---

### CureX Healthcare AI [🔗](#)

- Built an agentic **healthcare RAG assistant** using FastAPI, LangGraph, LangChain, and pgvector for evidence-grounded medical question answering over uploaded clinical documents.
- Engineered a retrieval pipeline with **document ingestion, chunking, metadata enrichment, vector search, and citation-backed response generation** using PostgreSQL + pgvector.
- Implemented **medical safety guardrails, faithfulness scoring, and human-in-the-loop review workflows** to flag unsupported or unsafe healthcare responses.
- Developed secure multi-user architecture with **JWT authentication, user-scoped chat memory, evaluation dashboards, and MCP healthcare tools** for ICD-10 lookup, drug interaction checks, and retrieval search.

### Debugging Intelligence System [🔗](#)

- Implemented **RAG-based debugging explanations** using retrieved incident evidence and contextual knowledge expansion for grounded AI-assisted troubleshooting.
- Developed **semantic deduplication and similarity detection** workflows using embedding-based retrieval to identify recurring incidents and reduce redundant debugging entries.
- Built REST APIs with **FastAPI + Pydantic v2** supporting ingestion, analytics, clustering, retrieval, and graph traversal endpoints.
- Integrated **Neo4j knowledge graph modeling** to map relationships between bugs, root causes, technologies, and recurring failure patterns.

### Text-to-SQL Agent [🔗](#)

- Built a **Text-to-SQL agent** using FastAPI, LangChain, and Mistral to convert natural-language questions into validated SQL and MongoDB aggregation queries.
- Engineered a **schema-aware RAG pipeline** with database introspection, business-rule retrieval, and live domain-value grounding to improve query-generation accuracy and reduce hallucinations.
- Implemented **semantic safety guardrails** and strict read-only query validation to block unsafe execution, ambiguous analytics queries, and unintended sensitive-data exposure.
- Added support for **SQLite, PostgreSQL, and MongoDB** with dialect-aware query generation, approval-based execution workflows, and Redis-backed semantic query caching.

## SKILLS

---

**Languages:** Python, Java, JavaScript

**AI/ML:** RAG Pipelines, AI Agents, Prompt Engineering, LLM Evaluation, Semantic Search, NLP

**Frameworks:** FastAPI, LangChain, LangGraph, PyTorch, Scikit-learn, Pydantic

**Databases & Retrieval:** PostgreSQL, ChromaDB, FAISS, pgvector, Neo4j, Redis

**Backend:** REST APIs, FastAPI, Spring Boot, JWT Authentication, Docker, Git

**Tools:** Git, Docker, FastAPI, REST APIs, LangSmith, Streamlit

## EDUCATION

---

### BE in AI and Data Science

East West Institute of Technology

GPA: 8.3

2022 – 2026 | Bengaluru

## PROFESSIONAL EXPERIENCE

---

### Java Developer

Dyashin TechnoSoft

02/2026 – 06/2026 | Bengaluru

- Built and deployed a Java-based backend application using Spring Boot, implementing REST APIs for user authentication and data management.
- Designed and integrated PostgreSQL database schemas, optimizing queries to improve data retrieval performance.
- Implemented core features using OOP principles, ensuring modular, reusable, and maintainable code structure.
- Used Git for version control, managing code changes and collaborating through branches and pull requests.
- Tested and debugged applications using tools like Postman and logging techniques, reducing errors and improving stability.