

ASHUTOSH YADAV

Dublin, Ireland · ashutoshy115@gmail.com · linkedin.com/in/ashutosh-yadav-057181194 · github.com/Kronos279

PROFESSIONAL SUMMARY

AI Engineer with 2+ years of experience bridging applied artificial intelligence with enterprise-grade software development. Delivered production-ready autonomous RAG systems and LLM workflows at EXL Service using Python, LangGraph, and vector databases. Completing an MSc in Computer Science at UCD with strong expertise in building scalable, cloud-native backend services.

EDUCATION

MSc Computer Science — University College Dublin

Dublin, Ireland — Sept 2025 – Sept 2026

- GPA 3.40/4.20

B.Eng Information Technology — Mumbai University

Mumbai, India — May 2019 – Jun 2023

- CGPA 8.34/10

EXPERIENCE

AI Software Engineering Intern — EXL Service

Dublin, Ireland — Jun 2026 – Present

- Independently built a production-grade Agentic **RAG** platform within two weeks of joining, featuring **S3** knowledge-base sync via **LiteLLM**, **pgvector** storage with fine-tuned chunking, document de-duplication, a stateless backend, and a **LangFuse** LLM-as-judge evaluation pipeline.
- Migrated the RAG agent's knowledge base to **AWS Bedrock**, integrating it into the platform API with automated document sync and health checks, and deployed the containerised service to **Dev/QA** via Docker and Helm.
- Built a graph-based fraud detection agent over **Neo4j** (45M+ transactions), combining tool-calling, agent-skills, and **LangGraph** approaches with a supervised ML risk-scoring model tuned for high recall.
- Extended the Platform **MCP** server with tools enabling an internal chatbot to query and edit AI solution, agent, and model metadata, adding chat-history context and **Bedrock Guardrails** policies for reliable interactions.

Software Engineer — Capgemini Technology Services

Mumbai, India — Mar 2024 – Aug 2025

- Developed and maintained full-stack enterprise features using Java, Spring Boot, Maven, and **REST APIs** on the backend and JavaScript/jQuery on the frontend, shipping **5+** production features across bi-weekly Agile Scrum releases for Sony and Universal Studios.
- Built event-driven workflows using Java EJB and asynchronous processing, reliably handling **1,000+** weekly automated transactions with validated data integrity via **SQL** and JUnit.
- Deployed builds to remote Linux servers over SSH across DEV, staging, and UAT environments, reducing release cycle times by **30%**.

Data Analyst Associate — Morningstar, Inc.

Mumbai, India — Oct 2023 – Feb 2024

- Automated data processing and validation pipelines for large-scale private equity datasets, improving workflow efficiency by **30%** and delivering analytical insights that contributed to a **15%** increase in client retention.

TECHNICAL SKILLS

Languages & Frameworks: Java · Python · JavaScript · TypeScript · SQL · Spring Boot · Spring AI · Spring Security · Angular · React · FastAPI · Hibernate · Node.js · JWT

AI / GenAI & LLM Tooling: **RAG** · **LangGraph** · **LangChain** · **LangFuse** · **AWS Bedrock** · Bedrock Guardrails · **LiteLLM** · **MCP** · Agentic Systems · LLM-as-Judge Evaluation · PyTorch · Sentence Transformers · Florence-2 (VLM) · FAISS (HNSW) · pgvector · Vector Search

Databases & DevOps: PostgreSQL · MySQL · MongoDB · **Neo4j** · RabbitMQ · Docker · Kubernetes · Helm · AWS · GCP · GCP Cloud Run · Azure (AZ-900) · Jenkins · CI/CD · Git · Linux · SSH · JUnit · JMeter · WebSocket

OPEN SOURCE CONTRIBUTIONS

seed4j-cli [\[npm\]](#) — Java, CLI

- Enhanced the help command in seed4j-cli to display example values for module parameters by surfacing **defaultValue** definitions from the core library, improving developer experience and CLI discoverability; merged and **live on npm**, verified with **mvnw clean verify**.

db2rest [\[GitHub\]](#) — Java, PostgreSQL, RSQL, JUnit

- Diagnosed and fixed a PostgreSQL boolean type-casting bug in db2rest (**457+** GitHub stars) where RSQL filters (e.g., **active==true**) caused runtime type mismatch errors; patched **BaseDialect** to return native Java Boolean objects and added a regression test in **PgReadControllerTest** to prevent recurrence.

RELEVANT PROJECTS

SnapSearch: Semantic Image Retrieval System [\[GitHub\]](#) — Python, PyTorch, Florence-2, FAISS, Sentence Transformers, PySide6

- Built a privacy-first macOS desktop app for natural-language image search over large photo libraries, engineering a 3-stage on-device pipeline (Florence-2 captioning, Sentence Transformers embeddings, **FAISS** HNSW ANN retrieval) achieving **95%+** accuracy with Apple Silicon MPS acceleration.

SkillHub: AI-Enhanced Full-Stack LMS [\[GitHub\]](#) — Java, Spring Boot, Spring AI, Angular, JWT, MongoDB, MySQL, Chart.js

- Built a modular LMS with Java microservices, an API Gateway, and an Angular/Chart.js frontend, integrating Spring AI for a conversational training assistant and **JWT**-based RBAC across Admin, Trainer, and Learner roles.

Corporate Finance Copilot [\[Live URL\]](#) — FastAPI, React, TypeScript, Docker, GCP Cloud Run

- Deployed a containerised FastAPI/React app to **GCP Cloud Run** with automated **CI/CD**, designing a **RAG** pipeline with HNSW vector indexing and real-time WebSocket streaming for sub-second document retrieval.