

Ria Jha

Los Angeles, CA | +1-213-849-4978 | riahja02@gmail.com | [LinkedIn](#) | [GitHub](#)

EDUCATION

University of Southern California (USC)

M.S. Computer Science

Los Angeles, CA
Aug 2025 - May 2027

– Coursework:

Agentic Information Retrieval Systems, Advanced Game Design & Development, Analysis of Algorithms, Web Technologies

Indira Gandhi Delhi Technical University for Women (IGDTUW)

B.Tech. Information Technology

Delhi, India
Aug 2017 - Aug 2021

– GPA: 9.0/10.0

TECHNICAL SKILLS

Languages: Python, Java, C++, SQL, JavaScript, TypeScript

Backend & Systems: Spring Boot, Django, Flask, Node.js/Express, Angular, REST APIs, Microservices, Event-Driven Architecture, Kafka Streams, RabbitMQ, OAuth2

Mobile: Android (Java, SDK, Retrofit 2, FusedLocationProviderClient, ViewPager2)

Cloud & DevOps: AWS, GCP (App Engine), Docker, Kubernetes, CI/CD, NGINX, Monitoring & Observability, Distributed Tracing

AI & ML: LangChain, RAG, FAISS, Agentic Pipelines, LLM Integration, Embeddings, PyTorch, NLP, Deep Learning (GRU)

Databases: PostgreSQL, MySQL, MongoDB Atlas, Redis

Testing & Tools: JUnit, Mockito, Swagger/OpenAPI, Slack API (Events, Block Kit), Git

WORK EXPERIENCE

Deutsche Bank - DWS Group

Software Engineer II

Pune, India
Mar 2024 - Jul 2025

- Owned end-to-end development of a client-facing trade processing dashboard (React + Spring Boot microservices) processing 10K+ daily trades with 80% reduction in data discrepancies; adopted as the primary reconciliation tool across trading desks.
- Built real-time event-driven data pipelines using Python and Kafka Streams to ingest external financial APIs with circuit breakers, retry logic, and fault-tolerant error handling, cutting production incident rate by 40%.
- Deployed containerized microservices on Kubernetes with Vault-managed secrets and distributed tracing, achieving 85% uptime SLA across 4 services.
- Authored technical documentation on Kafka partitioning, microservices scalability, and API design patterns; adopted as onboarding reference material across the team.

Deutsche Bank - DWS Group

Software Engineer I

Pune, India
Jul 2021 - Feb 2024

- Led migration of 3 legacy monoliths to containerized microservices with CI/CD pipelines, cutting deployment time from 2 weeks to 3 days across 4 production rollouts.
- Built Kafka-based event-driven pipelines to ingest, validate, and process 5K+ daily streaming transactions with automated compliance checks, eliminating 100% of manual audit workflows.
- Optimized PostgreSQL query plans and indexing for high-throughput data processing, reducing report generation time by 30%; owned production support including incident triage, root cause analysis, and hotfix deployments.

PROJECTS

Event Search Platform - Angular 20, TypeScript, Node.js, MongoDB Atlas, Java/Android, GCP

[Live](#)

- Dual-platform event discovery app - Angular SPA and native Android client (Java, Retrofit 2, ViewPager2) sharing a single Express/MongoDB backend on Google App Engine; integrated Ticketmaster, Spotify, Google Geocoding, and IPInfo APIs via a server-side proxy layer for credential security.
- Implemented geohash proximity search, adaptive radius geocoding, Spotify OAuth2 token caching, RxJS-debounced autocomplete, and reactive BehaviorSubject favorites state persisted in MongoDB Atlas.

PopRadio - Node.js, MongoDB, Redis, NGINX, Docker

- High-concurrency music streaming backend using Redis distributed locking to eliminate race conditions across 100+ concurrent users; NGINX load balancing for horizontal scalability and Redis-MongoDB write synchronisation for data consistency.

RAG Financial Q&A System — Python, FAISS, sentence-transformers, flan-t5, Flask, scikit-learn

- End-to-end RAG pipeline for natural-language Q&A over SEC filings and earnings reports: chunks and embeds documents into 384-dim vectors (sentence-transformers), indexes in FAISS for sub-millisecond similarity search, and generates grounded answers via flan-t5-base with prompt injection.
- Built an evaluation layer with retrieval metrics (precision@K, recall, MRR, MAP) and response quality metrics including faithfulness scoring (hallucination detection); added a prompt optimisation module using scikit-learn ParameterGrid to search over templates, temperature, and top-K, surfacing quality/latency tradeoffs across 18 configurations. Served via Flask REST API with 12 endpoints.

AWARDS, HONORS & RESEARCH

Global Runner-Up DWS Group AI Hackathon, Deutsche Bank

AI financial platform for stock portfolio recommendation, optimisation, and financial literacy education

3rd Place / 80+ Teams Microsoft for Startups AI Hackathon

Fraud detection backend - Django REST APIs, RabbitMQ, Random Forest classifier with custom feature engineering

Research Excellence Award Check-Worthy Sentence Detection, Published Research in Springer Journal - [Paper](#)

G2CW framework (GloVe + GRU) achieving 0.92 F1-score on 23K+ claims, surpassing prior best of 0.81; also curated IndianClaims dataset (953 Indian political sentences); fastest training among all CNN/LSTM baselines at 12.5s