Kuldeep Saini

mailtokuldeepsaini@gmail.com | Linkedin | GitHub | Portfolio

SUMMARY

Software Engineer & Data Engineer with **3+ years** of experience building distributed data pipelines (Apache Spark) and reactive microservices (Spring Boot WebFlux) for banking regulatory reporting platforms. Delivered **45+ production features** across **8 microservices**, reduced batch job failures to zero, cut customer delivery time by 57%, and improved pipeline throughput by 50%.

EXPERIENCE

Sopra Banking Software, India

Software Development Engineer 2 **Jan 2026 - Present**

- Designed a **pluggable multi-strategy data pipeline** on Apache Spark 4 by building interchangeable output strategies and a step-based architecture, automating report generation across 100+ regulatory form types - **eliminated manual processing entirely**.
- Built a **compliance report submission engine** from scratch by implementing cell-level filtering rules and structured output generation in Java, producing audit-ready financial reports - **zero compliance rejections post-release**.
- Implemented **distributed record locking** in a reactive REST API (WebFlux) by tying write access to background job lifecycle with lease-based auto-expiry - **eliminated all concurrent data-corruption bugs in production**.
- Resolved **Spark OOM failures** by profiling executor memory, applying dataset caching, and moving parsing to executor threads - **reduced compute cost by 50%** and **improved job speed by 20%**.
- Led pre-production upgrades and deployment automation, **reducing customer delivery time from 7 days to 3 days (57% improvement)**.
- Delivered cross-cutting features across **five services** (APIs, Spark jobs, shared libraries, and frontend) **in a single sprint** with zero deployment downtime.

Software Development Engineer 1 **Aug 2023 - Dec 2025**

- Built a **modular file ingestion pipeline** (Excel/CSV) from scratch - load, validate, transform, store - reused across **3 job types** without modification; cut new-format onboarding time by 60%.
- Delivered an **institution data import system** end-to-end (file parsing, multi-layer validation, RBAC) for banking clients - caught 100% of data quality issues at import time, eliminating downstream failures.
- Developed **4 distributed Spark batch jobs** processing millions of financial records with Spark SQL, MongoDB, and Apache Iceberg.
- Improved bulk ZIP export performance in a Spring Boot WebFlux API by correcting task scheduling in the reactive execution flow, reducing export execution time by 70% (600 to 180 seconds) and Kubernetes resource allocation by 50% (4 GB RAM/4 vCPU to 2 GB RAM/2 vCPU).
- Extended **shared libraries used by 8+ services** - contributed reusable models, error codes, and batch step definitions, reducing platform-wide code duplication.
- Strengthened CI/CD (GitLab CI, K8s, Helm): **improved deployment stability by 30%**, **reduced production defects by 25%**, achieved **90%+ test coverage**.
- Owned the discovery and implementation of complex product requirements, partnering with product managers to clarify scope and decompose requirements into actionable Jira tickets with clear acceptance criteria and dependencies. Conducted code reviews and mentored five junior engineers.

KEY PROJECT

PURE - Regulatory Reporting Platform *Java 17, Spring Boot WebFlux, Spark 4, MongoDB, Iceberg, Trino, S3, Kafka, K8s*

- Cloud-native platform automating ingestion, validation, and submission of financial regulatory reports (FICOBA, FICOM, COREP, RUBA) for central bank compliance across multiple jurisdictions.
- Owned the **dimensional output pipeline end-to-end**: CSV/Excel load → validation → rule engine → report generation → storage (MongoDB + Iceberg) - handling 100+ form types per institution.
- Refactored a monolithic submission engine into a **modular step-based architecture**, enabling independent testing and reuse - adopted by 2 additional job types post-refactor.

TECHNICAL SKILLS

Backend:	Java, Python, Spring Boot, Spring WebFlux/Reactor, Spring Security, Spring Cloud, Hibernate, REST APIs, Microservices, OpenAPI, Kafka, JUnit, Gradle, Git
AI/LLM:	Prompt Engineering, RAG, AI Agents, MCP (Model Context Protocol), LangChain, Ollama, FAISS
Data & DB:	Apache Spark, Spark SQL, Apache Iceberg, Trino, MongoDB, MonetDB
Cloud/DevOps:	AWS (S3, EC2, EKS), Docker, Kubernetes, Helm, GitLab CI/CD
Frontend:	TypeScript, Angular
Fundamentals:	DSA, OOP, Concurrency, Multithreading, LLD, HLD, System Design, Agile, Scrum

EDUCATION

B.Tech - Computer Science **Aug 2019 - Jul 2023**

IMS Engineering College, Ghaziabad CGPA: 8.3/10

AWARDS

Kalam Award (3x) and Aryabhata Award for high-impact contributions to core product development.