Sr. Data Architect

**Location:** Bengaluru

# Job Description

Sr. Data Architect

San Francisco Bay Area, CA

Job Summary

We are seeking an experienced Sr. Data Architect with deep expertise in Databricks and other modern data platforms to design, optimize, and govern our data infrastructure. This role will be responsible for architecting scalable data solutions, integrating disparate data sources, and enabling advanced analytics and AI-driven insights. The ideal candidate will have extensive experience in cloud-based big data platforms, data modeling, and implementing ETL/ELT pipelines on Databricks and related technologies.

Key Responsibilities

· Architect & Implement Scalable Data Solutions: Design and implement high-performance data architectures leveraging Databricks, Apache Spark, and Delta Lake.

· Data Integration & Pipelines: Develop optimized ETL/ELT workflows using Databricks Workflows, Delta Live Tables, and integration with cloud storage solutions.

· Cloud Data Management: Design data lakehouse architectures in Azure/AWS/GCP using Databricks and similar platforms, ensuring high availability, security, and compliance.

· Performance Optimization: Optimize Spark jobs, query performance, and cluster configurations to ensure efficient data processing.

· Data Governance & Security: Implement data quality frameworks, access controls, and compliance measures using Unity Catalog, Delta Sharing, and equivalent tools.

· Collaboration & Leadership: Work closely with engineering teams, data scientists, and business analysts to enable data-driven decision-making.

· Machine Learning, Analytics & Agents Enablement: Support the development and operationalization of ML models, AI applications, and agent-based systems using Databricks, MLflow, Feature Store, and other ML/AI platforms such as SageMaker, Vertex AI, or Azure ML. Familiarity with agentic frameworks, AI agents, and orchestration tools (e.g., LangChain, AutoGen, CrewAI) is highly desirable.

Technical Documentation: Maintain detailed architecture diagrams, process flows, and best practices documentation.

Qualifications & Skills

· 10+ years of experience in Data Architecture, with strong expertise in Databricks, Apache Spark, and other distributed data platforms.

· Proficiency in SQL, Python, Scala, and Cloud platforms (Azure, AWS, GCP).

· Deep understanding of Delta Lake, Medallion Architecture, and Data Engineering best practices.

· Experience in data governance, cataloging, and security (Unity Catalog, RBAC, IAM).

· Strong knowledge of distributed computing, data pipeline orchestration (Airflow, Azure Data Factory), and cloud storage (S3, ADLS, GCS).

· Experience integrating BI tools (Power BI, Tableau) with data platforms for analytics.

· Strong understanding of machine learning workflows, AI infrastructure, and emerging agentic technologies.

· Excellent problem-solving skills and ability to communicate complex technical concepts to non-technical stakeholders.

Preferred Qualifications

· Databricks certifications (Databricks Certified Data Engineer/Architect).

· Experience with data streaming technologies (Kafka, Spark Streaming).

· Familiarity with agentic frameworks, AI orchestration tools, and multi-agent systems.

Hands-on experience with ML & AI platforms (e.g., SageMaker, Vertex AI, Azure ML).