JAYANT MOHITE

Certified Cloud Data Platform Architect | Big Data Consultant | Technical Trainer

CONTACT

- 📞 +91-9545-831-218
- ✉️ hello@jayantmohite.com
- 📍 Baner, Pune, Maharashtra, India
- 🌐 www.jayantmohite.com

EDUCATION

Pursuing
Mumbai university

- Master of Business Management

2008 - 2012
pune University

- Bachelor of Engineering
- Computer Engineering

CERTIFICATIONS

- Microsoft Certified Azure Solutions Architect Expert
- Databricks Certified Data Engineer Professional
- Microsoft Certified Power BI Data Analyst Associate
- Microsoft Certified Azure Administrator
- Salesforce Certified Administrator (SCA)
- Salesforce Certified Advanced Administrator (SCAA)
- Salesforce Certified Sales Cloud Consultant

PROFILE SUMMARY

Results-driven **Azure Solutions Architect & Technical Manager** with 13+ years of experience designing and delivering enterprise-scale data platforms across **Azure, Databricks, Microsoft Fabric, and Power BI.** Proven expertise in building end-to-end cloud data ecosystems unifying ingestion, processing, governance, machine learning, & analytics. Strong background in **Hadoop, Spark, Kafka, MongoDB, Cassandra, and Neo4j,** evolved into modern lakehouse architectures leveraging **Delta Lake, Synapse, and Fabric OneLake.** Skilled in **FinOps, CI/CD, data governance,** & real-time streaming, with a track record of driving cost optimization, scalability, and actionable insights for global organizations.

WORK EXPERIENCE

SKAN AI LABS
Data Platform Manager / Solutions Architect

2020 - Present

Transformed traditional data ecosystem into a modern, cloud-native analytics platform built on **Azure, Databricks, Microsoft Fabric,** and **Power BI,** unifying data, AI, and governance to deliver real-time business intelligence and measurable cost efficiency.

- Partnered with CXOs, product owners, and client stakeholders to define modernization roadmaps, cloud strategies, and reference architectures aligned with business goals.
- Designed end-to-end architectures integrating **Data Factory, Synapse, Event Hub, Kafka, Airflow,** Blob Storage, **ADLS, OneLake,** and **Delta Lake** for unified batch and streaming workloads.
- Engineered optimized **PySpark** and Structured Streaming pipelines via Databricks Workflows, **Auto Loader,** and **REST APIs,** achieving a 40% increase in throughput and reduced latency.
- Led Delta Lake infrastructure design using Delta Engine, **Z-Ordering,** and Data Skipping, improving query and storage performance; orchestrated the Hive Metastore → Unity Catalog migration with ABAC and ACLs for robust data governance.
- Led architectural assessments for external client engagements, recommending Hadoop-to-Databricks, SQL-to-Fabric, and on-prem to cloud landing zone migration strategies
- Managed and modernized **MongoDB** infrastructure, overseeing migration to MongoDB Atlas with Global Clusters, Multi-Region Writes, and Automated **Sharding;** optimized schema design and indexing for a 30% improvement in query execution times.

TRAINING & PUBLICATIONS

- Best-Seller Author – MongoDB Administration course on Udemy (4.3+ stars, thousands of learners)
- Published Author – Technical content for HP Micro Focus Big Data, O'Reilly, Packt Publications
- Featured in Higher Education Magazine for contributions to skill-based technical education
- Conducted 300+ corporate, academic and retail trainings on Databricks, Big Data, Spark, Power BI, NoSQL Databases and Salesforce
- Certified Salesforce Trainer for 12+ courses
- Founded Wisdom Sprouts in 2016 - An Academic and Retail Training Platform

BIG DATA PROCESSING

- Apache Airflow
- Apache Hadoop
- Apache Nifi
- Apache Pig
- Apache Spark
- Apache Sqoop
- Cloudera Hadoop
- Databricks
- Data Factory
- Data Marts
- MongoDB
- MS Fabric
- One Lake
- Snowflake
- StarRocks
- Synapse Analytics
- Unity Catalog

WORK EXPERIENCE

- Built secure multi-region infrastructure with VNets, private endpoints, and managed identities; executed FinOps and performance-tuning strategies (auto-scaling, caching, parallelization, telemetry) via **Azure Monitor** and **Log Analytics**, reducing compute costs by 25%.
- Established data governance, lineage, and monitoring frameworks using **Unity Catalog**, **Purview**, Fabric Workspaces, and Log Analytics, ensuring 99.9% uptime for critical operations.
- Architected semantic and visualization layers with Power BI, Fabric DirectLake, and Synapse SQL, delivering governed dashboards and KPI reporting for leadership teams.
- Integrated **Azure Machine Learning**, **Snowflake**, **PostgreSQL**, MongoDB, and **Scikit-Learn** for predictive analytics and cross-platform data exchange.
- Promoted automated **CI/CD** and release management practices via **Azure DevOps**, enabling faster, scalable, and cost-effective pipeline deployments.
- Led the feature engineering team using **Agile** methodologies, managing sprint planning, reviews, and delivery of high-impact data and analytics features aligned with business objectives.

PALANNTECH INDIA   2017 - 2020
Senior Big Data Consultant

Directed large-scale Big Data platform design and delivery initiatives for enterprise clients, leading cross-functional teams of data engineers and DevOps specialists to modernize data ecosystems across Hadoop, NoSQL, and cloud environments (Azure, AWS).

Key Achievements & Contributions
- Managed agile data-engineering and DevOps teams delivering new analytics capabilities and advanced feature-engineering pipelines under aggressive timelines.
- Designed and deployed end-to-end **Hadoop** clusters using **HDFS**, **YARN**, **Hive**, **Pig**, **Sqoop**, and **Spark**, enabling distributed analytics on multi-terabyte workloads.
- Built **ETL** and **ELT** workflows with **Apache Airflow, Oozie**, and **Kafka**, orchestrating ingestion from **RDBMS**, **APIs**, and file systems at enterprise scale.
- Architected and administered **Cassandra**, **MongoDB**, and **Neo4j** environments for high-volume, schema-flexible, and relationship-intensive use cases.
- Achieved a 20% reduction in cloud infrastructure spend through improved resource utilization and cross-cloud cost governance (**Azure + AWS**).
- Implemented partitioning and schema-design strategies in Hive and **HBase** for optimized query performance and scalability.
- Developed **Power BI, Tableau**, and **Superset** dashboards to provide real-time insights and visual analytics to business stakeholders.
- Configured cluster security, capacity planning, and performance tuning, ensuring high availability and efficient workload distribution.
- Provided architecture consulting and migration roadmaps for transitioning on-prem Hadoop workloads to cloud-native data platforms on Azure and AWS.

CLOUD INFRASTRUCTURE

- Azure Infrastructure Management
- Azure Cost Management
- Azure DevOps
- Azure Monitor
- AWS
- Docker
- Elastic Search
- FinOps
- Grafana
- Kibana
- Kubernetes
- Log Analytics
- Prometheus
- Terraform

EVENT STREAMING

- Apache Flink
- Apache Flume
- Apache Kafka
- Event Hub
- RabbitMQ
- Sqoop

DATA VISUALISATION

- Chart.js
- Power BI
- SuperSet
- Tableau

MANAGEMENT

- Agile Delivery Management
- Change Management
- Customer Engagement
- Product Management
- Risk Management
- Strategic Planning & Execution

WORK EXPERIENCE

FREELANCER   2014 - 2017
Big Data Consultant

Delivered end-to-end Big Data and NoSQL solutions for clients across finance, e-commerce, and analytics domains, leveraging the Apache Hadoop ecosystem and modern distributed data technologies.

Key Achievements & Responsibilities
- Built and deployed data platforms using **Apache Hadoop (Hortonworks, Cloudera)**, **Spark, Kafka, Cassandra, MongoDB**, and **Neo4j**, enabling high-volume analytics and real-time processing.
- Designed custom ETL frameworks with Apache NiFi, Airflow, and Pig to automate ingestion, transformation, and loading workflows, improving reliability and reducing manual intervention.
- Implemented Kafka + Spark Streaming pipelines for financial and e-commerce clients, delivering low-latency, high-throughput event-driven analytics.
- Developed scalable data models and **replication/sharding** strategies in MongoDB, Cassandra, and Neo4j, optimizing query performance and fault tolerance.
- Built Hadoop data lakes supporting batch and near real-time workloads with **Hive**, **Pig**, **Sqoop**, and Spark, consolidating structured and semi-structured data sources.
- Conducted hands-on client training sessions on Big Data stacks, empowering internal teams to manage and scale their own data infrastructure.
- Created shell/Python-based monitoring and data-flow scripts using **REST APIs** for pipeline observability and operational control.

DB XENTO SYSTEMS   2012 - 2014
Software Engineer

- Administered and optimized **PostgreSQL** databases, improving performance through query tuning and indexing.
- Automated data validation and reconciliation processes using **Python** scripts, reducing manual QA effort and improving data reliability.
- Developed **Power BI** prototypes for internal performance monitoring, combining **SQL** and Python outputs into visual insights for engineering managers.
- Monitored and fine-tuned application and database servers for high availability, fault tolerance, and scalability.
- Built reusable ETL utilities in Python to extract and transform data from **REST APIs**, CSVs, and transactional databases.
- Wrote stored procedures, functions, and triggers to support business workflows and reporting use cases.