

Rijul Sahu, Senior Process Manager | +91-8879019216 | rijulsahu@gmail.com

Skype: rijul.sahu | <https://rijul.xyz> | AWS Certified Solutions Architect



PROFILE

Seasoned process re-engineering and operational transformation **Big Data analytics** professional with 9 years of cross domain and latest technology stack experience in project advisory and IT project management for data and compliance functions with a proven track record in digital transformation using Hadoop & AWS cloud services

SUMMARY

Stakeholder & Project Management

- Managed multiple projects as a Manager - requirements gathering, defining project scope and charter, solution definition, documentation and risk and control framework
- Built and transformed teams - providing stabilization throughout rebuilding, including key quality new hires, training, and coaching
- Collaborated with onsite/offshore development resources from delivery partners to develop & test integrated solutions
- Analyzed client requirements to define business strategy & determined system architecture requirements to achieve business goals

Data Engineering & Integration

- Hands on experience in architecting, designing and building big data pipelines using Horton works data platform, Cloudera stacks in on-premise and **AWS** stack in cloud
- Utilized Spark ML Libraries to design and develop Topic modeling and text classification machine learning solution https://rijul.xyz/projects/Topic_Modeling.html
- Designed and deployed **ETL** data pipelines in **cloud infrastructure** and on premise
- Led and executed multiple data integration and migration projects involving enterprise **data lakes**
- Enhanced performance of applications for generating quick actionable insights <https://rijul.xyz/projects/Exasol.html>
- Expert on multiple tech tools for new implementations, migrations and upgrades
- Strong exposure in product development, enterprise **data management & data governance** compliance

TECHNICAL CERTIFICATIONS

Cloud Services

- AWS Certified Solutions Architect – SAA-C02

Big Data

- IBM & LinkedIn certified Big Data Analytics

Database

- Oracle certified SQL expert developer & administrator

ANALYTICAL & TECHNOLOGY SKILLS

Key skills

- Big data analytics, data warehouse design, digital transformation, data integration, solutions architect

Platforms, Languages, Framework & Tools

- Python, Unix shell, Hadoop – Hive, Pig, Sqoop, HDFS, SparkSQL, Spark ML & PySpark - Intermediate, SQL, ETL, Alteryx

Cloud Computing

- Amazon web services – VPC, EC2, Auto scaling, ELB, RDS, EMR, Redshift, S3, Route53, Lambda - Basics, IaaS – Terraform & Cloud formation - basics

MAJOR CLIENTELE & PROJECTS HIGHLIGHT

Client I: “European multinational online food-delivery service based in Berlin, Germany” – AWS Setup & Architecture

- Evaluated various proposals & recommended the best technical solution to develop the application in AWS
- Designed the architecture and authored the Terraform IaaS template to facilitate deployment

- Provided complete Infrastructure solution to client including end to end configuration and BOM of Server & Storage components
- Solutioning in progress using – Aurora Serverless RDS, EC2, S3, VPC, Terraform, GitHub

Client II: “Largest sportswear manufacturer in Europe, and the 2nd largest in the world” – Platform Migration, Automation
<https://rijul.xyz/projects/DI.html>

- Technically leaded onsite team of 10 member including managers and senior leads – I was deployed in Zaragoza, Spain for client engagement and delivery
- Developed in-house capability using existing tool sets for fast delivery & enabled DevOps model
- Automated and implemented the solution with high data quality and scalable architecture
- Eliminated delivery cost of 3rd party vendor and license cost (saved cost of **\$0.25 million**)
- Designed data model for multiple business unit's data consumption for multiple stakeholders enabling data analysis
- Solution implemented using – Exasol, Python, AWS EMR cluster, CI/CD – Jenkins, Bitbucket for code versioning

Client III: “Global Leader in Architecture, Engineering & Entertainment Software, HQ in USA” – Consumer Analytics

- Designed and developed ETL data pipelines from enterprise Data Lake
- Performed ad-hoc analysis and data pulls related to e-commerce, Marketing and/or customer insights data to support cross-functional partners in Marketing and Merchandising
- Authored business cases and project management support for key strategic initiatives and/or task forces based on current areas of focus and growth for the business
- Understand the nature and format of data to decide on the structure for the data warehouse & identify what operations to be performed in order to have consistency in data imported from heterogeneous sources/systems
- Solution implemented using – Shell Scripting, AWS EMR cluster, S3 & Redshift for storage, CI/CD – Jenkins, Hive, PySpark

Client IV: “Trading leaders of Electro components in Europe & Asia Pacific” – Competitive intelligence, Retail pricing
<https://rijul.xyz/projects/Hadoop.html>

- With the aim of reducing the end-to-end time spent on manual and time-consuming activities in the Operations team, I designed and deployed 6 nodes on premise Hadoop cluster for data ingestion and data processing with 20 cores horsepower, 150GB RAM, and 2TB storage.
- Created runtime transformation packages [ETL] and data structures for 12 million+ inventory items of 43+ competitors and 30 K product categories
- Identified Price change patterns and monthly trends along with multiple KPIs like Stock status, Article matches, Percent difference
- Saved cost of around **\$100K** which was previously spent in resourcing and ongoing support
- Solution implemented using – Hortonworks data platform, Hive, Pig, Sqoop, Shell, UNIX, Oozie for scheduling, SQL server

Client V: “eClerx in house capability building” – Product development, Center of Excellence

- Leded center-of-excellence team for providing end-to-end data integration expertise on new as well on-going projects
- Headed data engineering & governance related projects from end-to-end implementations
- Strategizing the data journey with customer's IT stakeholders
- Infiltrating technical understanding to developers and monitoring the deliverables

Client VI: “third-largest private-sector banks in India” – Customer Segmentation, Fraud Analytics

- Authored Sqoop jobs for data loading from ODS to Hive tables
- Wrote multiple data processing scripts for transaction data mining/ Text mining
- Authored Shell script for running programs on the go and to automate data loading and batch jobs
- Housekeeping activities of Hadoop Production and Development Cluster

PROFESSIONAL DESIGNATIONS & ROLES

Senior Process Manager [Reporting to VP – Data Analytics & Consulting, Digital]

Aug 2016 – Till present

eClerx Limited

Consultant

May 2015 – August 2016

Saksoft Limited

Consultant*ServPro Technologies Pvt Ltd*

March 2014 – February 2015

Software Engineer*Geinfosys Technology Pvt Ltd & Pearl Stone Technology Pvt Ltd*

August 2010 – Sept 2013

ACADEMIC SUMMARY

ITM Business School, Mumbai - Perusing Executive MBA in Marketing & General Mgmt.

2018-2020 Batch

CSVТУ, Bhilai – Completed Bachelors of Engineering in Electronics & Telecommunications

2010 Batch