

RUTURAJ JENA

Big Data | Python | Spark | Hadoop | AWS Certified Solution Architect | AWS Certified Data Analytics Specialty | Google Cloud Certified Professional Data Engineer | SME AWS Glue

Phone: +91 7873060130 | Email: raturajjena2020@outlook.com

LinkedIn: www.linkedin.com/in/raturajjena GitHub: <https://github.com/raturajjena> Address: Kondapur, Hyderabad, Telangana, 500084 | 30 days Notice Period.

PROFESSIONAL SUMMARY

Results-oriented Data Engineer with 8 years of experience architecting, developing, and optimizing high-scale, distributed data pipelines and cloud-native big data solutions across regulated insurance and enterprise data domains. Currently lead the design and architecture of a 145-source ingestion platform spanning RDBMS (Oracle SQL, MySQL, Postgres), NoSQL (DynamoDB, MongoDB), and internal API sources, built on a multi-zone lakehouse using AWS Glue, DMS, Step Functions, Lambda, SNS, and Snowflake. Strong software engineering skills in Python, with deep SQL and data-modeling expertise across Spark, ETL, and CDC, and a proven track record building high-scale distributed data systems. Hands-on with LLMs, prompt engineering, and agent frameworks (Amazon Bedrock), applying best practices in cloud, data engineering, and LLMOps, and using Claude and ChatGPT routinely to prototype, test, and ship reliable code faster. Proven at taking new functionality from design through development and testing to production deployment while collaborating across functional teams.

WORK EXPERIENCE

Deloitte USI Consulting – Senior Consultant – AI & Data – Nov 2025- Present

- Lead the design and architecture team building an enterprise data-ingestion platform for 145+ source systems distributed across Oracle SQL, MySQL, Postgres, DynamoDB, MongoDB, and internal API sources for a large insurance client (Liberty Mutual).
- Architected a three-zone lakehouse — raw landing (Zone 1) through to the consumption layer — storing Zone 1 output in Amazon S3 and cataloging it in the Snowflake Horizon catalog, serving multiple business pods including Claims Distribution, Underwriting, and Data Office.
- Built ETL and CDC pipelines using AWS Glue, AWS DMS, Step Functions, Lambda, and SNS, with Stonebranch as the external enterprise scheduler triggering the Step Function workflows.
- Designed an Audit-Balance-Control (ABC) framework on Postgres to track, balance, and govern every pipeline run across all pods.

- Designed a source-to-target reconciliation framework to guarantee data completeness and integrity across 145 heterogeneous sources.
- Developed full-load and CDC pipelines from DynamoDB to Amazon S3 using AWS Glue (Spark).
- Designed and developed Reconciliation framework between with all RDBMS sources and snowflake as target.

Amazon Web Services- Cloud Engineer – Big Data Oct 2021 – Oct 2025

- Designed and implemented scalable ETL pipelines using AWS Glue (Apache Spark-based) to process data from diverse sources such as Amazon S3, RDS, Redshift, and OpenSearch, enabling efficient data ingestion and transformation.
- Developed real-time data processing solutions using Apache Kafka and Apache Flink, ingesting data from Kafka and writing it to Amazon S3.
- Led team under for data replication projects with multiple JDBC source and targets.
- Created Cross cloud integration from GCP Big Query to Amazon S3 using AWS Glue (Spark).
- Managed AWS Lake Formation to build secure and organized data lakes with fine-grained access controls.
- Worked with open table formats like Apache Iceberg, Delta Lake, and Hudi to enable scalable and reliable ACID-compliant data lake architectures on Amazon S3.
- Orchestrated data workflows using Amazon MWAA (Managed Workflows for Apache Airflow), streamlining end-to-end pipeline execution.
- Recognized as a Subject Matter Expert (SME) in AWS Glue, leading initiatives to troubleshoot and optimize Spark-based ETL jobs for enterprise clients.
- Worked directly with various enterprise customers, delivering high-quality solutions to complex big data problems.
- Supported Amazon Bedrock projects by integrating GenAI use cases with secure and governed data access.
- Worked on building AI agents in BedRock using python.
- Authored Public Facing Articles for AWS

Key Accomplishment: Built multi-cloud data pipelines and introduced scalable, automated data ingestion from Big Query to S3 using Glue, improving cross-cloud compatibility and reducing manual overhead.

Deloitte USI Consulting – AI and Data Sept 2018- Sept 2021

Google Cloud Architect | Big Data Admin | Walmart (Client)

- Managed and maintained Cloudera cloud-break clusters using Shell and Python scripting.
- Designed finance data pipelines from Teradata to Google Cloud using GCP services.
- Built and managed Dataproc clusters, configuring autoscaling policies for optimized resource usage.
- Developed Python automation scripts for cluster health checks and node monitoring, enhancing platform stability.
- Provided technical support for Hive, Spark, and other big data technologies, troubleshooting critical production issues.
- Collaborated with Google and Cloudera teams to resolve infrastructure-related challenges.

Key Accomplishment: Led the implementation of Independent Edge Node architecture that significantly reduced cloud billing and enabled multi-cluster connectivity.

EDUCATION

Bachelor of Science, Computer Science (GOLD Medalist) July 2015 - May 2018 Ravenshaw University, Cuttack

Intermediate Senior Secondary School July 2008 - Mar 2015 Jawahar Navodaya Vidyalaya, Mundali, Cuttack

TECHNICAL SKILLS

Programming Languages: Python, SQL, Shell, PySpark

Big Data: Spark, AWS Glue, Apache Airflow, HDFS, Hadoop, YARN, MapReduce, Hive, Kafka, Flink, EMR, lake formation, distributed computing, Iceberg, Delta Lake, Hudi, Lambda, SQS, SNS, Crawler, Step Functions, AWS DMS, Snowflake, Stonebranch

Databases: Oracle SQL, PostgreSQL, MySQL, DynamoDB, MongoDB, Snowflake, Athena, Google BigQuery, Hive, Amazon RDS, Redshift, S3

Cloud Platforms: AWS, GCP, Azure

CI/CD- GitHub, CFN, Terraform

AI / GenAI: Amazon Bedrock, SageMaker, LLMs & agent frameworks, prompt engineering, LLMOps, Claude, ChatGPT, and other native AWS AI tools.

CERTIFICATIONS

- AWS Certified: AI Practitioner
- AWS Certified: Solution Architect

- AWS Certified : Data Analytics Specialty
- Google Cloud Certified: Professional Data Engineer
- Microsoft Certified: Azure Fundamentals

ACADEMIC PROJECTS

Customized Android OS Development

- Built customized Android OS for Snapdragon devices with kernel-level optimization for gaming performance.
- Deployed monthly updates with security patches and performance improvements.

AWARDS/HONORS

- Gold Medalist in Computer Science
- Subject Matter Expert in Glue ([Credly badge](#))
- Won All rounder Award at AWS ([Credly badge](#))
- Won Most Valuable Player (MVP) 2 times At AWS ([Credly badge](#))
- Won Knowledge Champ Award at AWS ([Credly badge](#))
- State-level Cyber and Mathematics Olympiad Winner