

EEGAPURI LOKESHWAR REDDY

Data Engineer

+919959242262 | elokesh4292@gmail.com | Telangana, India | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

SUMMARY

Data Engineer with 3 years of experience designing and supporting batch and streaming data solutions, ETL/ELT workflows, and enterprise data warehousing across production and disaster-recovery environments. Hands-on expertise in dimensional modeling, OLAP systems, and data marts, gained through extensive EDW experience. Proficient in relational database architectures, big-data frameworks, and cloud-based analytics, with a strong interest in data lake and lakehouse technologies for unified platform solutions. Combines deep operational discipline with a problem-solving mindset to build robust, real time and batch data solutions.

SKILLS

Databases & Warehousing: PostgreSQL, Snowflake, Teradata, Delta Lake, Data Warehousing, Data Modeling, Dimensional Modeling, ERD, Data Lakes, Data Lakehouse

ETL / Workflow Orchestration: Ab Initio, Apache Airflow, Shell Scripting

Big Data & Batch Processing: Apache Spark (SparkSQL, PySpark), Hadoop (HDFS, YARN)

Streaming & Messaging: Apache Kafka, Confluent Kafka, ksqldb

Cloud & DevOps Tools: AWS Glue, AWS S3, Databricks, Red Hat Linux, Git, Looker Studio

Programming Languages & Libraries: Python, SQL, Pandas, NumPy, Requests, SQLAlchemy, DB API

Collaboration & Issue Tracking: JIRA, Remedy

Soft Skills: Problem Solving, Communication, Analytical Thinking, Adaptability, Attention to Detail, Risk Management, Decision-Making, Emotional discipline, Patience, Reliability, Knowledge Sharing

Languages English, Hindi, Telugu

CERTIFICATIONS

IBM Data Warehouse Engineer, IBM	Mar '25
Snowflake Data Engineering, Snowflake	Jun '25
Confluent Fundamentals Accreditation, Confluent	Jan '25
Academy Accreditation - Databricks Fundamentals, Databricks	Dec '24

PERSONAL PROJECTS

RentLok: Real-Time Rental Management System | [Link](#)

- **Tech Stack:** Android, FastAPI, PostgreSQL, Apache Kafka, Kafka Connect, ksqldb, SQLAlchemy, Pydantic, JSON
 - Built an end-to-end **OLTP** data pipeline for a rental management system with a native **Android app** and backend powered by **FastAPI**, **PostgreSQL**, and **Apache Kafka**.
 - Developed secure **RESTful APIs** over **HTTPS** to process **JSON** data, enforced schema validation with **Pydantic**, and used **Kafka Connect** to stream PostgreSQL data into Kafka topics.
 - Implemented **ksqldb** for real-time transformations and aggregations, enabling live in-app metrics; optimized the **PostgreSQL** schema for low-latency access and modular microservice deployment.
- **Key Learning:**
Successfully transitioned from **OLAP** to **OLTP** system design, demonstrating the ability to build reliable, API-driven, real-time data pipelines and translate business requirements into scalable, end-to-end data solutions.

Telecom Data Pipeline | [Link](#)

- **Tech Stack:** Apache Airflow, Apache Spark (PySpark), PostgreSQL, Shell Scripting, Python
 - Designed a batch processing **ETL pipeline** to process millions of **CDR records** using **Airflow** for orchestration and **PySpark** for scalable, distributed transformations.
 - Developed **fraud detection logic** to identify suspicious patterns like high-frequency calls and short-duration international interactions.
 - Implemented **audit logging** and **error tracking**, improving pipeline observability and operational reliability.
- **Key Learning:**
Demonstrated the ability to design **scalable, distributed** data transformations using **PySpark**, while enhancing reliability through fraud detection logic and robust logging.

E-Commerce Data Pipeline | [Link](#)

- **Tech Stack:** Apache Airflow, MySQL, PostgreSQL, Foreign Data Wrapper (FDW), Python, Pandas, SQL
 - Built a **modular batch ETL pipeline** using **Airflow** to automate **CSV ingestion** and transformations across **MySQL** and **PostgreSQL**.
 - Enabled **cross-database querying** using **PostgreSQL FDW**, secured credentials with **.my.cnf** and **.pgpass**, and optimized performance through **indexing**.
 - Exported **cleaned data** for downstream analysis, showcasing a **reusable and scalable** data integration workflow.
- **Key Learning:**

- Strengthened skills in Airflow orchestration, relational database ingestion, and **secure, cross-database batch processing**, aligned with real-world data engineering practices.

Game Development & 3D Modeling | [Link](#)

- **Tech Stack:** Unity, C#, Blender
 - Developed a **PC game ("Cubon")** using **Unity** and **C#**, implementing player controls, interactive mechanics, and UI features.
 - Created and animated **3D** models in Blender, applied **texturing** and **rendering**, and integrated assets into the **Unity engine**.
- **Key Learning:**
Gained hands-on experience in **software development**, event-driven programming, and **3D** rendering pipelines — combining creative and technical skills foundational to full-cycle product development.

EXPERIENCE

System Engineer/Data Operations	Apr '19 — Feb '22
Tata Consultancy Service (TCS)	Hyderabad
• Project 1: Saudi Telecom Company (STC) – Enterprise Data Warehouse (EDW) Operations (Apr '19 - Feb '22)	
Environment: Production & Disaster Recovery (DR)	
Tools: Ab Initio, Teradata, Red Hat Linux, SQL, Shell Scripting, Jira, Remedy	
• Monitored 2000+ enterprise-scale ETL jobs across OLAP and data mart environments, responding to daily requests and escalations to ensure uninterrupted data flow.	
• Oversaw ingestion from CRM, CDR, and IRB systems via SFTP and SQL Extraction, automating workflows to reduce manual interventions.	
• Handled 200+ Change Requests, including deployment of 45+ new ETL jobs, optimizations, data reloads, and decommissions—preventing multiple potential SLA breaches.	
• Developed and optimized SQL queries to track business-critical KPIs (postpaid, prepaid, landline, sales) and automated discrepancy reporting for proactive resolution.	
• Created Shell Scripts to monitor system resources, detect transformation failures, and clean unwanted files—recovering 30% storage space, improving operational efficiency by 35% , and ensuring 100% data completeness.	
• Collaborated in 4–5 Disaster Recovery (DR) drills to validate failover processes: tested job workflows, source connections, scheduling configurations, and maintained documentation for seamless execution.	
• Project 2: Saudi Telecom Company (STC) – Kafka-Based Event Streaming Operations (Mar '21 – Feb '22)	
Environment: Production	
Tools: Confluent Kafka, Python, Kafka Client Libraries, ksqlDB, Avro, JSON	
• Managed operations and maintenance of 250+ Kafka topics, supporting real-time event streaming and ksqlDB stream processing alongside EDW batch workflows.	
• Ensured schema compatibility by handling Avro and JSON message formats and resolving connector issues (JDBC/ODBC, HTTP integrations, SSL errors, schema mismatches).	
• Developed Python scripts using Kafka client libraries to automate health checks, monitor consumer lag, and send real-time alerts—significantly reducing manual monitoring effort.	
• Addressed core Kafka issues including RBAC permission errors, driver mismatches, performance tuning, and schema evolution conflicts to maintain reliable data delivery.	

EDUCATION

B-Tech in Mechanical Engineering , CVR College of Engineering (GPA: 9.56)	Mar '15 — Mar '19 Hyderabad
Intermediate in MPC , Narayana Junior College (GPA: 9.7)	Apr '13 — Mar '15 Hyderabad

AWARDS & ACHIEVEMENTS

- **On the Spot Award** for outstanding performance and exceptional contributions to **ETL operations** and **system optimization**.
- Received **client appreciation** for ensuring seamless system operations and maintaining **100% KPI availability** with **zero SLA breaches**.
- Cleared the first round of **TCS CodeVita Season 7 (2018)**, a global coding competition, demonstrating strong **programming** and **problem-solving skills**.