

Eegapuri Lokeshwar Reddy

Data Engineer

9959242262 ◇ elokesh4292@gmail.com ◇ Telangana, India ◇ [LinkedIn](#) ◇ [GitHub](#) ◇ [Coursera](#)

SUMMARY

Data Engineer with 3 years of experience in ETL operations, data pipeline automation, and enterprise data warehousing. Proven ability to monitor, troubleshoot, and optimize large-scale data workflows in production and disaster recovery environments. Hands-on expertise in dimensional modeling, OLAP, datamarts, and workflow orchestration, with skills honed through EDW experience and a strong foundation in scalable data infrastructure. Passionate about development and automation, leveraging modern Big Data tools to build efficient data pipelines. Combines operational discipline with a problem-solving mindset to design robust, high-performance data solutions. Eager to apply deep operational knowledge and engineering skills to develop innovative, data-driven systems.

WORK EXPERIENCE

System Engineer

Tata Consultancy Service (TCS)

Apr '19 — Feb '22

Hyderabad, India

Project 1: STC (Saudi Telecom Company) – Enterprise Data Warehouse (EDW) Operations

Environment: Production & Disaster Recovery(DR)

Tools: Ab Initio (ETL), Teradata (EDW), Red Hat Linux, SQL, Shell Scripting, Jira, Remedy

- Monitored **2000+** enterprise-scale **ETL jobs** across OLAP and data mart environments; responded to daily email requests and escalations from source and reporting teams to ensure uninterrupted data flow.
- Oversaw data ingestion from multiple upstream systems (including CRM, CDR, and IRB platforms) using both pull (SFTP, SQL extraction) and push (file transfers) methods; implemented automation to reduce manual interventions.
- Handled **200+ Change Requests (CRs)** via Jira, including **45+ new ETL Jobs** deployment, optimizations, data reloads, and job decommissions; gained strong understanding of pipeline and SQL-level logic within **distributed, MPP-based systems**, maintained a **flawless deployment track record**, and helped prevent multiple potential SLA breaches through careful planning and execution.
- Developed and optimized SQL queries to monitor daily business-critical KPIs, including postpaid, prepaid, landline, and sales metrics; automated discrepancy reporting and resolution tracking.
- Investigated and resolved critical system issues such as database locks, CPU spikes, schema mismatches, connectivity errors, and data rejections; coordinated with developers for patching and long-term fixes.
- Developed **shell scripts** to monitor system resources, detect transformation failures, and alert teams. **Proactively** identified and cleaned unwanted files, recovering **30% storage space** and improving operational efficiency by **35%**.
- Led real-time communication during escalations, tracked incidents via Jira and Remedy, and ensured internal and client teams received timely updates.
- Reprocessed rejected data records and validated fixes to ensure **100% data completeness** post-resolution.
- Participated in **4–5 disaster recovery (DR) drills** as part of a cross-functional team, replicating live production environments to validate failover processes.
- Tested and verified critical components — including job workflows, source connections, scheduling configurations, and stakeholder access mappings — to ensure seamless business continuity during disruptions.
- Created and maintained documentation of all ETL workflows, data dependencies, run schedules, and recovery steps to support smooth DR execution.
- Managed pre- and post-DR communications, coordinated system validations, and confirmed data consistency and readiness across environments.

Project 2: STC (Saudi Telecom Company) –Kafka-Based Event Streaming Operations

Environment: Production

Tools: Confluent Kafka, Python, Kafka Client Libraries, Avro, JSON.

- Managed operations and maintenance of **250+ Kafka topics**, supporting real-time event streaming alongside EDW batch workflows.
- Handled **Avro and JSON** message formats to ensure schema compatibility between producers and consumers, maintaining downstream data integrity.
- Developed Python scripts using Kafka client libraries to automate health checks, monitor consumer lag, and send real-time alerts — significantly reducing manual monitoring.
- Resolved connector issues related to **JDBC/ODBC** and **HTTP** integrations, including connectivity failures, SSL/certificate errors, schema mismatches, and data type casting challenges that impacted reliable data delivery.
- Addressed core Kafka issues including RBAC permission errors, client-side driver mismatches, producer/consumer configuration tuning for performance, and **schema evolution conflicts** on topic fields.

EDUCATION

B-Tech in Mechanical Engineering, CVR College of Engineering, Hyderabad (GPA: 9.56)

2015 — 2019
Hyderabad, India

Intermediate in MPC, Narayana Junior College (GPA: 9.7)

2013 — 2015
Hyderabad, India

CERTIFICATIONS

Confluent Fundamentals Accreditation, Confluent

Academy Accreditation - Databricks Fundamentals, Databricks

IBM Data Engineering Professional Certificate, IBM (In Progress)

ACHIEVEMENTS & AWARDS

- 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.

PERSONAL PROJECTS

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

- **Tech Stack:** Apache Airflow, MySQL, PostgreSQL, Foreign Data Wrapper (FDW), Python, Pandas, SQL, Ubuntu, LucidChart.
 - Built a batch ETL pipeline using Apache Airflow to automate CSV ingestion and transformations in MySQL and PostgreSQL.
 - Enabled cross-database queries with PostgreSQL Foreign Data Wrapper (FDW), optimized performance, and secured credentials using .my.cnf and .pgpass.
 - Exported transformed data for analysis, showcasing a reusable and scalable ETL workflow aligned with industry standards.

Telecome Data Pipeline | Data Engineering Project [Link](#)

- **Tech Stack:** Apache Airflow, Apache Spark (PySpark), PostgreSQL, GitHub, Shell Scripting, Python, Ubuntu, LucidChart.
 - Designed a batch processing ETL pipeline to process millions of CDR records using Apache Airflow for orchestration and PySpark for scalable, distributed data transformations in a single-node setup.
 - Developed fraud detection logic in PySpark to identify suspicious patterns such as high-frequency calls, short-duration international calls, and excessive customer service interactions.
 - Implemented audit logging and error tracking, enhancing pipeline observability, debugging, and operational reliability.
 - Delivered a scalable and modular fraud detection system, enabling future integration with Kafka for advanced and real-time processing.

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

- **Tech Stack:** Unity, C#, Blender
 - Developed a PC game ("**Cubon**") using Unity and C#, implementing player controls, interactive mechanics, and UI components.
 - Created and animated 3D models in **Blender**, applying texturing and rendering techniques, and integrated them into the game engine.
 - Showcased a blend of programming, problem-solving, and creative design skills through self-driven projects during college.

SKILLS

Tools/Platforms: Apache Airflow, Ab Initio, Confluent Kafka(Client Library Integration), Apache Kafka, Hadoop(HDFS, YARN), Apache Spark(SparkSQL, PySpark), Dimensional Modeling, Data Modeling, Entity-Relationship Diagram(ERD), Data Warehousing, JIRA, Remedy, Red Hat Linux, Ubuntu, Databricks, Git, CI/CD, Looker Studio

Programming Languages: Python, SQL, Shell/Bash Scripting

Libraries/Frameworks: Pandas, NumPy, Requests, BeautifulSoup, SQLAlchemy, DB API

Data Repositories: Teradata, PostgreSQL

Cloud Tools & Services: AWS S3, AWS Glue(Learning)

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

Interests Big Data and Cloud Technologies, Distributed Systems, Devops

Languages English, Hindi, Telugu