

Venkatesh Guduru

Lead Data Engineer

✉ venki24365@gmail.com

☎ 8977682268

🌐 venkatesh-guduru

SUMMARY

Senior Data Engineer with strong technical, project and people skills. **9.7+ years'** of experience in designing, developing, and maintaining data warehouses and Data pipelines with vast experience in Relations Databases and tools like AWS Glue with PySpark, Informatica PowerCenter, and other ETL Tools.

4+ years of people management experience and managing engineers. My current role involves team management, design reviews, project planning, estimations, development, oversight of agile ceremonies and deliverables, hiring, and building product delivery roadmaps along with business and other stakeholders.

WORK EXPERIENCE

Innova Solutions (Formerly GGK Technologies)

Data Engineer

May, 2018 - Present

Tata Consultancy Services Ltd.

ETL and SQL Developer

Oct, 2015 - May, 2018

COMPETENCIES

Problem Solving	Data Quality	System Design (HLD/LLD)
Data Modeling	Data Integration/Migration	Cloud & Big Data Technologies
Agile Project Management	Resource Planning	Team Leadership
People Management	Performance Management	Hiring & Talent Development

SKILLS

- **Cloud Services** : AWS - Glue, S3, Athena, RDS, EC2, CloudWatch, RedShift, RDS
- **ETL Tools** : IBM DataStage, Informatica PowerCenter
- **Programming** : PySpark, Python, SQL, Spark
- **Database Management** : Snowflake, Oracle, SQL Server, PostgreSQL, Sybase, Greenplum
- **Streaming** : Kafka, AWS Kinesis, Pub-Sub
- **Scheduling Tools** : Airflow, Automic and crontab
- **Other Tools** : Git and TFS (Version Control)
Jira (Project work and issue tracking)
ServiceNow (Service management)
Wiki and Confluence (Knowledge meet)
- **Project Management** : Scrum, Retrospectives, Estimations, Project Planning, Design reviews, Requirement analysis
- **Team Management** : Hiring, People management, Performance management

PROJECTS

Enterprise Data Warehouse Modernization- US and EU Based HealthCare Provider Client

2022 - Present

Tech Stack: AWS(Glue, PySpark, Athena, S3), PySpark, SnowFlake, Oracle, Airflow

- Designed and implemented data pipelines using AWS Glue and PySpark to extract data from various sources.
- Implemented Glue ETL pipelines using spark, RDS, Athena and S3.
- Improved the performance of existing processes with changing the file types used and optimized techniques.
- Designed and implemented a hybrid approach to maintain business continuity during staged migrations,

ensuring seamless integration of on-premises and cloud systems.

- Implemented CDC tracking in RDS for run logs and recorded timestamps for successful job completions for audit and data quality purposes.
- Error Handling and Logging: Implemented robust mechanisms for monitoring API requests, logging errors, and retrying failed tasks.
- Leveraged materialized views and Tasks in Snowflake to enhance application performance
- Processed various file types, including JSON, delimited files, and Parquet.

Customer Data Integration Project - US Based HealthCare Insurance Client

2018 - 2022

Tech Stack: IBM Datastage, Oracle, Greenplum, Linux, Automic

- Worked on Developing, supporting, and maintaining ETL processes using IBM DataStage, including custom SQL processes.
- From design to implementation of critical Sybase to Oracle – data migration and a few other processes.
- Developed end to end incremental data loads from OLTP database tables to data lake using ETL pipeline.
- Implemented CDC on all OLTP tables using DB procedures, triggers, magic tables and audit tables.
- Worked on Performance related issues and fine-tuned system processes which improved the operational efficiency.
- Worked on revamp of CMS extracts process to improve performance that impacts evaluation and Federal Ratings.

ODS and DWH Enhancements - EU Based Telecom Client

2015 - 2018

Tech Stack: BnCorLatlca Powerbenter, Sybage, Linux, Crontab

- Created ETL mappings and workflows to extract data from various sources and load into the data warehouse.
- Conducted process analysis using data modelling techniques to identify and implement optimizations.
- Implemented automated failure tracking mechanisms by leveraging Informatica metadata tables for improved monitoring and maintenance
- Worked on SQL tools to perform ETL tasks and prepare reports using stored procedures.
- Coordinate with all stakeholders until the completion of the SDLC cycle of every change request.
- Dealt with client issues related to system design as well as to business and suggested process changes.

EDUCATION

JNTU, Kakinada, Andhra Pradesh, India

Bachelor of Technology (ECE) - 71.9

2011 - 2015

Board of Intermediate Education, Andhra Pradesh, India

Senior Secondary School Certification - 85

2009 - 2011

CERTIFICATIONS

AWS Certified Developer – Associate

Snowflake Certified Associate - Data Warehouse Essentials

Oracle PL/SQL Developer Certified - Associate

AWARDS

Employee of the Year 2022 - Innova Solutions

Employee of the Year 2021 - Innova Solutions

Customer Champion, July 2021 - Innova Solutions

Star of the Month - May 2020 and Nov 2018 - Innova Solutions

Best Performer 2021 - Tata Consultancy Services Ltd.