Ritwik Verma

Senior Data Engineer – 5.10 YoE | GCP Certified Data Engineer

✉ ritwik2199@gmail.com 📱+91 8840459910 in linkedin.com/in/ritwik-verma/

TECHNICAL SKILLS

Programming Languages / Framework's: Python, Spark, Apache Beam, Hadoop, Java
Databases & Orchestration: Oracle, Snowflake, SQLServer, ElasticSearch, NoSQL, Airflow
Google Cloud Platform Services: BigQuery, Dataproc, Dataflow, Cloud Run, Cloud Function, Cloud Storage, Secret Manager, Kubernetes, Pub/Sub,Cloud SQL, IAM
Control Systems and Documentation: Git, BitBucket, Jira, Confluence, Visio, Draw.io
Build Tools & Containers: Maven, Jenkins, Docker

WORK EXPERIENCE

Persistent Systems | Senior Data Engineer | Apr 2022 – Present

Project 1: GCP Cloud Electrical asset location accuracy prediction & correction platform
- Architected and developed spark data processing solution that calculates location accuracy score using more than 5 billion photos captured by drones & ground teams.
- Migrated 30 million high res.photos from sources like Sharepoint, Fulcrum , NAS drives etc. to Google cloud storage and Elasticsearch saving SaaS subscription costs.
- Designed application using cost-optimized cloud technology stack and Implemented RBAC IAM provisioning.

Project 2: LIDAR Data Ingestion Platform: GCP Cloud Data Lake management & QC platform
- Migrated 50 TB unstructured data from Azure blob storage to GCS then processed these files at scale and categorically stored files into structured folder hierarchy defined as per business use case.
- Designed a file processing solution that uses ElasticSearch as downstream system for capturing ingested files metadata also created data reprocessing pipeline and reconciliation summary pipeline.
- Developed REST API's for web interface to facilitate users to perform bulk operation on ingested files at scale..

Project 3: YouTube Data Ingestion Pipeline: Batch data pipeline for YouTube videos metadata
- Developed batch data pipeline solution using Dataflow to ingest & transform data received from Youtube API and stored data in BigQuery for AI/ML application.
- Migrated large Snowflake tables to BigQuery using Dataflow jobs.
- Developed stored procedures and deployed on BigQuery scheduled queries for transformation.
- Reduced client's daily batch ingestion pipeline time by 50%.

Project 4: GCP Dataflow Ingestion Pipeline: ETL pipeline from Snowflake to GCP BigQuery
- Developed GCP Dataflow templates with dynamic value provides to configure query & other database related parameters to help the Data Scientist team to migrate required tables from RDBMS sources like Oracle, SAP to BigQuery.
- Created batch ELT pipeline from Snowflake to BigQuery which creates dynamic schema at downstream.
- Developed dataflow templates to migrate relational data from on-premise to BigQuery.

HCL Technologies | Software Engineer |Oct 2019 - Mar 2022

API Development: Menu & Cart API for US large Quick Service Restaurant client
- Developed Menu REST API's in an agile production environment for major pizza restaurant chains in the USA.
- Developed a variety of menu for multiple online food delivery aggregator partner's requirements.
- Worked on multiple business requirements to improve the menu which increased sales.
- Worked on Development of modules and functionalities, mainly involved in development related to aggregators partners.

ACHIEVEMENTS

Top Talent | Persistent Systems | 2025
Bravo Individual Award | Persistent Systems | 2024
Top Talent | Persistent Systems | 2023

EDUCATION

Bachelor Of Technology - Electronics & Communication | Jaypee Institute Of Information Technology | Noida | 2019

CERTIFICATIONS

Google Professional Data Engineer | 2025 | Link

Google Professional Database Engineer | 2025 | Link

Google Associate Cloud Engineer | 2025 | Link

Google Cloud Digital Leader | 2022 | Link