

CAREER PROFILE

Experienced Data/DataOps/MLOps Engineer with thorough hands on experience in all levels of software development life cycle. I am a hybrid Data Engineer, Software Developer, Automation Engineer. Supportive and enthusiastic team player dedicated to streamlining processes and efficiently resolving project issues. Willing to take ownership of core components.

SKILLS

- **Languages & Frameworks**  
Python, Java, SQL, Fast API, Spring-boot, Shell/Bash script, Scala(Intermediate-level)
- **ELT & Data-Pipeline Tools**  
Apache Airflow, Apache Beam, Apache Flink, DBT, Xplenty, AWS DMS, Great expectations, Apache Spark, Kafka
- **Cloud Services**  
GCP(Dataflow, Composer, BigQuery, Vertex AI, Cloud Functions, Pub/Sub, DataCatalog, GCS, Cloud SQL, Cloud Run, Data Catalog), AWS(DMS, Redshift, RDS, Glue, Lambda, S3, Rekognition)
- **Databases**  
PostgreSQL, Oracle, MySQL, Elasticsearch, MongoDB, Redis, Memcache, RocksDB, Sybase
- **Technologies & Tools**  
Terraform, Kafka, Elasticsearch, Docker, Kubernetes, New Relic, Solr, RabbitMQ, Prometheus, Grafana, Graylog, Postman, Wireshark
- **Reporting Tools**  
PowerBI and Redash (Beginner to intermediate level of exp)
- **CI/CD & Automation Tools**  
Terraform, Github Actions, Jenkins, Buildkite, Selenium-Webdriver, RestAssured (API), Jacoco for code coverage
- **Load Testing Tools**  
Locust, Vegeta, Gatling and Wrk2
- **Version Management Tools & PMS**  
Git, JIRA

CERTIFICATIONS

- Professional Data Engineer** 2023 - 2025  
Google Cloud
- Professional Machine Learning Engineer** 2024 - 2026  
Google Cloud

EXPERIENCES

**Software Engineer III** 10/2024 - Present  
Wayfair, Bangalore/India

**Senior Data & MLOps Engineer** 05/2021 - 09/2024  
Wayfair, Berlin/Germany

- Collaborating with stakeholders across the business to understand their data requirements.
- Implemented the event driven ETL pipeline for processing the Google Analytics data using Apache Beam, Cloud Functions and other GCP managed services.
- Implemented the scalable realtime ETL pipelines using Apache Flink for processing Kafka events.
- Developed data dictionary using Google Data Catalog
- Implemented Quality Checks framework for the pipelines running on Airflow.
- POC on different ETL tools and technologies.
- Implemented CI/CD plugin to automate deployment to GCP managed services.
- Infrastructure provisioning thorough Terraform.
- Lead the productionization efforts for machine learning models using GCP Vertex Batch and Realtime services, ensuring seamless deployment and scalability in a production environment.
- Collaborate closely with Data Scientists to provide curated datasets for model training, leveraging GCP's managed services for data preprocessing and post-processing tasks.
- Serve as Subject Matter Expert (SME) for infrastructure architecture, implementing robust solutions for deployment, orchestration, and monitoring of ML pipelines and systems.
- Establish and maintain robust monitoring and alerting systems, ensuring the reliability and performance of machine learning models and associated infrastructure.

**Senior Data Engineer** 04/2020 - 04/2021  
Wefox, Berlin/Germany

- Migrated the legacy data infrastructure to the GCP utilising the GCP's serverless and managed services.
- Migrated the VBA based ETL tool to Airflow.
- Extract and Load setup using Xplenty.
- Data migration and CDC using AWS DMS.
- Built utilities in Scala & Python for processing structured and unstructured data from multiple data sources like API's, S3, CSV, SFTP.
- Gather and process raw data at scale (including writing scripts, web scraping, calling APIs, writing SQL queries and applications)
- Built API's using FastAPI.
- Setting up of CI/CD pipeline using Github actions and Bash Scripts.
- On-boarding the new MGA's with the necessary changes in the DWH and ETL.
- Implemented realtime alerting for the data pipeline and services.
- Developed the data dictionary.
- Effective documentation about the applications, tools and utilities developed.
- Creating high level design documents & source to target data mapping sheets.
- POC on different ETL tools and technologies.

**Data Engineer** 11/2017 - 03/2020  
Spotcap Global, Berlin/Germany

- Built data-pipeline in Scala & Python for processing structured and unstructured data from multiple data sources like API's, Audit Tables, S3, which was then consumed by the data science team.
- Gather and process raw data at scale (including writing scripts, web scraping, calling APIs, write SQL queries, writing applications)
- Built API's using Scala Play and Spring boot which acts as simulator for user journey and it's actively used by the product, QA and data science team.
- Developed a utility by implementing API's and integrating with the Flock for pro- actively debugging the credit decision on user's loan application.
- Setting up of CI/CD pipeline (Jenkins integration) for the automation test suite.
- Setting up the mocked data, which was critical during client demos and internal testing, and acted as POC for the mocked data services.
- Built automation framework for Spotcap's backend API's from scratch.
- Built load testing framework using Locust and Python.
- Implemented real time alarming for the quality check failures to all the stake holders.
- Effective documentation about the applications, tools and utilities developed.

**Tech Lead** 09/2015 - 11/2017  
OLA, Bangalore/India

- Worked on the application which served as a single source of truth for position/status of all cabs at any given point of time.
- Design and architecture of tracking service.
- Updating the Chef Cookbooks and Recipes and deploying the build on AWS- Cloud Environment.
- Developed utilities to minimize the environment setup, for production sanity test and for debugging the issues in staging.
- Application maintenance and troubleshooting.
- Giving demos to the product team and Knowledge sharing sessions to the team.
- Mentoring the junior team members.

**Senior Software Engineer** 11/2012 - 09/2015  
Altisource Labs, Bangalore/India

- Migration of existing code bases to AWS cloud.
- Migration of SQL based searches to Apache SOLR based search.
- Built a highly scalable and concurrent design by using multithreading principles and Rabbit MQ.
- Implemented build and deployment process for the product using maven, jenkins etc.
- Involved in the development of in house frame work and reusable components
- Interacted with the business team to understand and analyze the requirements.
- Designed the Creating Invoice, Creating Workflows Invoice Approval modules.
- Worked on data integration, payment file generation and interfacing with 3rd party.
- Worked with Support to resolve the production issues across the application.
- Optimizing code for increased scalability.
- Perform integration testing, system testing, regression testing.
- Monitoring and scheduling Daemons.
- Automation of test cases using Selenium Web driver, Sahi, Robot Framework and Shell script.
- Setup Jasper Reports server
- Wrote liquibase scripts, queries, stored procedures.
- Wrote Shell scripts to automate payment files, deployment, server health monitoring.
- Gave Demos, Knowledge sharing sessions to the team.

**Software Engineer** 09/2009 - 10/2012  
ZTE R&D Center, Kuala lumpur/Malaysia, Bangalore/India

- Preparing the system test cases, Test List and System Test Report.
- Preparation of test environment and deployment.
- Integration of all the third-party interfaces with RBT.
- Execution of IOT test cases and resolving the inter communication issues with 3rd party modules.
- Implementation and maintenance of matrices such as Test Matrix and Requirement Tractability Matrix.
- Executing Sanity/Smoke test on new builds.
- Preparing test reports for the system test performed.
- Study the failures and checking server logs and Wireshark logs.
- Installation of Linux, Database, tomcat, Jboss.
- Preparing Test bed.
- Upgrade/Migration testing.
- Execution of UAT test cases.
- Preparation of dual site environment for GR test and execution of test cases and preparation of document for GR to use at disaster condition.



Ashok Ballolli  
Senior Data & MLOps Engineer

- ✉ ballolliashok@gmail.com
- ☎ +91-9739929985
- 🌐 DataTe.ch
- 🌐 ashokballolli
- 🌐 ashokballolli
- 📧 ashokballolli
- 🐦 @ashok\_ballolli

EDUCATION

B.E in Computer Science  
VTU  
2005 - 2009

LANGUAGES

English (Professional)