



Vijay Chhatbar

Data Engineer

✉ vijay@chhatbar.net

📍 Hamburg, Germany

🔗 <https://chhatbar.net>

🧠 SKILLS

- Python
- Airflow
- Informatica PowerCenter
- SQL
- Kafka
- GCP/AWS/Azure Cloud
- Spark/Databricks
- CI/CD
- Unix shell scripting

📁 WORK EXPERIENCE

Oct 2023 – present
Hamburg, Germany

Bauer Media Group

Senior Data Engineer

- I handle end-to-end Data Engineering, from ingesting data from heterogeneous sources to generating insights that help make sense of it.
- Our data stack runs on GCP, and we use mostly serverless services like Cloud Run and Cloud Functions for the EL part, along with Pub/Sub for real-time data use cases.
- For the transformation layer, we use dbt orchestrated through Airflow where we use Cosmos framework, where all our pipelines are scheduled. BigQuery serves as our data warehouse.
- I also manage our data platform, which includes provisioning and maintaining Kubernetes clusters, managing VPC connectivity, configuring ingresses, implementing IaC with Terraform, using GitOps practices, deploying applications with Helm, setting up CI/CD pipelines, and managing private package repositories.

May 2022 – Sep 2023
Hamburg, Germany

FREE NOW

Data Engineer

- Design, develop, and maintain data pipelines for various data sources using PySpark, Databricks and Airflow.
- Integrate various heterogeneous data sources, such as PostgreSQL, Kafka, MySQL, and 3rd party APIs, into the data warehouse using Airflow as the scheduler.
- Optimize and fine-tune ongoing data pipelines to reduce costs and improve performance, using techniques such as Delta Live, Spark UI profiling, cluster management, and optimization of Spark shuffling and joins.

Jan 2020 – Apr 2022
Mumbai, India

Reliance Jio

Deputy Manager

- Design, develop, and maintain efficient, scalable, and robust data pipelines for ingesting, transforming, and loading data from various heterogeneous sources into the Hadoop ecosystem.
- Create and maintain streaming jobs using technologies such as Kafka, PySpark streaming, and Hive, to process live data in real-time.
- Use workflow management tools such as Airflow or Azkaban to orchestrate and schedule PySpark jobs running on YARN, ensuring optimal resource utilization and failover.

Oct 2017 – Jan 2020
Mumbai, India

Teradata

Technical Consultant

- Migrate legacy systems, including data and ETL processes, from Oracle and Informatica 9.1 to Teradata and Informatica 10, ensuring data consistency and maintaining a history of data loads.
- Create logical and physical data models using industry best practices to ensure data quality and reduce redundancy.
- Design, develop, and implement data loading solutions using Informatica Intelligent Cloud Services (IICS) for semi-structured data sources, such as JSON files.

Feb 2016 – Sep 2017
Mumbai, India

Cognizant

Associate

- Design, develop, and maintain over 10 regulatory reports, comprising of more than 100 KPIs (Key performance indicators) to meet business and regulatory requirements.
- Perform data manipulation using various Informatica transformations such as Filter, Expression, Lookup (Connected and Unconnected), Aggregate, Update Strategy, Normalizer, Joiner, Router, Sorter, and Union.
- Implement partition techniques to load data into target tables in SybaseIQ, resulting in improved efficiency of 70-75%.

Dec 2013 – Jan 2016
Mumbai, India

IGATE Global Solutions

Software Engineer

- Design, develop, and maintain data mappings, sessions, and workflows in Informatica to support data integration and ETL processes.
- Write, execute, and maintain unit test cases to ensure the accuracy and completeness of the data being processed by the mappings and workflows.
- Perform data validation, cleansing, and transformation techniques to ensure data quality and consistency throughout the pipeline.

EDUCATION

2010 – 2013

Bachelor of Science in Information Technology

University of Mumbai

CERTIFICATES

- Astronomer Certification for Apache Airflow Fundamentals
- Microsoft Azure Fundamentals (AZ-900)
- Teradata Vantage Certified Associate 16.00