Rahul Jain

Last update on November 12, 2022

Head of Business Intelligence, Analytics Engineering and Data Engineering at Beat

rahulj51@gmail.com • Berlin • Germany

https://www.linkedin.com/in/rahulj51/ • twitter.com/rahulj51 • rahulj51.github.io/

Summary

*I am a technical leader in data with over **20 years of experience in building and running software products** for various domains with a focus on data and analytics engineering.*

My key strengths lie in **leading large scale evolution of data architectures** by building a strong data foundation rooted in core engineering principles, aimed to meet the needs of both Business and Operational intelligence at growing organizations. I firmly believe in getting things done.

Over the last few years, I have led various initiatives such as **large scale data platform migrations**, establishing a data modeling strategy and modernizing the data stack towards governance-first architectures.

In my current role as Head of BI at Beat Mobility, I lead the Analytics Engineering and Data Engineering functions, responsible for building a robust data and analytics engineering foundation for our ride-sharing business in Europe and Latin America.

Some of my recent accomplishments include

a. Identified the need for a **tech strategy to change Beat's central data architecture** from a service-based practice to a systems based architecture aligned with the principles of data modeling, governance and data-as-a-product.

b. Revamped our data platform from a data-lake/raw-data based system to a modern data stack built on Snowflake, dbt and Airflow. This would eventually **save Beat over USD 500,000 in cloud cost** but more importantly bring higher efficiency and productivity for analytics and data science.

c. Utilizing the new data platform, I **established a data governance strategy** that covers access control, data privacy, data cataloging as well as isolation of compute resources in order to make resource utilization more measurable, independently scalable and cost effective.

d. **Led the Data leadership chapter** at Beat, mentoring and supporting other data leaders with a recent focus on data contracts and a shift towards a data mesh culture.

e. Attracted, retained and grew high performing teams of Analytics Engineers and Data Engineers in a self-driven, inclusive culture with a focus on autonomous decision making and continuous improvement.

f. Built the next level of leadership within the team with continuous coaching and delegation; focusing on autonomy, decision making and empathetic leadership.

g. As **part of the extended leadership team at Beat**, I represented the Data Engineering function within the larger Beat organization; closely working with data producing teams, data platform teams as well as multiple consumers of data across Europe and Latin America.

Although my current role does not require me to write code, I am fairly hands-on with technology, software engineering and coding and actively contribute to our code at Beat (**over 1000 code commits in 2022**). I also occasionally contribute to open source communities in the data space.

Outside work, I am passionate about giving back to the community by mentoring engineers, analysts and engineering managers across the world. I offer free mentoring over hangouts as a member of a mentoring club. Occasionally, I provide consulting to clients in Europe and Asia on their data/analytics engineering setups.

I write frequently on data strategy, engineering and leadership on my blog, Twitter and LinkedIn.

Recent writings

Headless-BI 101 Ghost wrote an article on Headless BI architecture and an in-depth evaluation of some existing tools in this space.

Modern Data Stack 101 Ghost wrote an article on Modern Data stack and a framework for choosing the right tools.

Data Build Tool - The Beat story An article on introducing dbt as a data transformation tool at Beat

16 fundamental principles for transforming data in a warehouse Some basic principles for data transformation and data modeling.

Migrating a data warehouse across Cloud providers My experience and approach as I led my team to migrate our entire data warehouse from Redshift to BigQuery.

Work Experience

Beat
Athens, Amsterdam and Remote

Head of BI, Analytics Engineering and Data Engineering *July 2021 – Dec 2022 (1 year 6 months)*

Leading the Analytics Engineering and Data Engineering functions, responsible for building a robust data and analytics engineering foundation for our ride-sharing business in Europe and Latin America.

See **summary section** for current roles and responsibilities.

Omio (formerly GoEuro)
Berlin, Germany

Principal Engineering Manager, BI *Nov 2017 – Jun 2021 (3 years 8 months)*

Responsibilities included:

- Providing a vision to the team for the evolution of our data platform and data strategy.
- Providing technical leadership to my team with respect to system design and architecture, managing tech debt and future planning and prototyping to keep our engineering systems up-to-date.
- Working with multiple stakeholders including BI analysts, data scientists and other business functions to provide data for their rapidly growing needs.
- Working with the engineering leadership to establish technical best practices in the area of data engineering
- Growing a motivated team with regular feedback cycles and participation in their career development.
- Managing data infrastructure cost
- Managing data governance around privacy (GDPR), meta-data management and data quality management.

Key accomplishments:

- Evolved the data engineering architecture from an early stage batch heavy system to a state of the art data platform based on **real time streaming systems**.
- Stabilized the data pipelines from an MTBF of 10 days to 90 days.
- Significantly improved timeliness, quality and reliability of analytics data.
- Led the development of self serve, **real-time data ingestion platform based on Kafka and Apache Spark** that enabled any team at Omio to setup their own ETL pipeline to send data to warehouse.
- Grew the BI engineering team from 4 members to 10.
- Executed a **complex data warehouse migration project to move from Redshift to BigQuery**. This migration resulted in an annual saving of $200K in cloud cost.

Thoughtworks
Bangalore & Boston

Lead Consultant and Technical Lead *Oct 2012 – Nov 2017 (5 years 3 months)*

Key responsibilities included:

- Technical Project Lead on Thoughtworks projects.
- Responsible for end to end delivery of product features.
- Architecture and Design of software systems.
- Implementing agile technical practices and processes
- Agile Consulting

Credit Suisse
Singapore

Senior Analyst and Java Development Lead *Jul 2007 - Sep 2012 (5 years 3 months)*

Responsible for developing multiple java/j2ee based ETL jobs used for ingesting large volumes of financial market data (balance sheets, stock prices, earnings estimates etc.) into Credit Suisse's data products.

Infosys — Bangalore, Hyderabad and Boston

Software Developer & Project Lead — *Jan 2000 - Jun 2007 (7 years 6 months)*

Worked on various projects as a senior Java programmer and later as a project lead.

Education

altMBA — Online

3 months Workshop on Leadership and Decision making — *2021*

Indian Institute of Technology (ISM) — India

Bachelor of Technology, Mining and Mineral Engineering — *1995 – 1999*

Main project on creating a simulation model of the grinding efficiency of a ball-mill based on feed properties.

Skills

Technical expertise:

- Modern Data platforms and tools
- Data collection and ingestion systems - including both batch and real-time
- Data transformation and modeling - Tehcniques, practices and tools (e.g., dbt)
- Data governance architecture and tooling. Metadata management systems
- Analytical Data storage formats and systems, including data lakes (Parquet.Orc on S3) and cloud data warehouses - Snowflake, Bigquery and Redshift
- Advanced data modeling techniques including domain driven design of data
- Setting up and operating Reporting and Dashboarding tools such as Tableau and Looker
- Reverse ETL techniques and processes for operationalizing data
- Engineering Leadership, strategy, planning and execution
- Agile software development processes
- Software design principles and enterprise integration patterns
- Microservices and service-oriented architecture
- Data engineering principles and DataOps
- CI/CD pipelines and modern software release processes
- Programming languages - Python, Scala, Java
- Distributed processing and storage systems - Kafka, ElasticSearch, Apache Spark
- Cloud platforms - AWS and GCP
- ETL orchestration frameworks like Airflow and Argo
- DevOps Toolchain - Virtualization, Containers (Docker) and Kubernetes