

Data Lake ingestion pipeline

With the strong support and guidance of b.telligent METRONOM has implemented a fully automated GCP data ingestion pipeline.

The challenge

With Metro's global footprint and the multitude of countries and verticals which have their own data silos, METRO follows the strategy to bring all this data together in one Digital Analytics solution. Based on METRO's strategic decision to move to GCP, the existing Hadoop cluster has been replaced with a Data Lake built with Big Query, incl. reengineering of the data ingestion pipeline.

The solution

b.telligent has advanced a concept for a group of REST-enabled microservices, an ingestion endpoint and custom Schema Registry for AVRO-based schemas (incl. schema evolution), both supporting separation of data for individual countries, verticals and stages. Data flows through HTTP(s) ingestion endpoint to Kafka, storing schema midway, and further via Cloud Dataflow to target dataset in BigQuery.

The results

Low learning curve of the ingestion pipeline has enabled its quick adoption by various countries and verticals. The pace at which the data ingestion into Data Lake has accelerated, opened new use cases of Data Lake on GCP across METRO. Avro-based schema evolutions and data consistency has proven to be crucial building blocks for future capability extensions.

“ Due to b.telligent's ability to tackle complex problems and deliver high quality concepts we were able to successfully kickoff the Digital Analytics solution. ”

Marko Schwob, Domain Owner Analytical Platform Engineering



About METRONOM GmbH

METRONOM GmbH is the tech unit of METRO, a leading international wholesale and food specialist company. METRONOM provides customized IT services and IT solutions for all METRO countries worldwide.

Industry: Retail & Wholesale

Primary project location: Germany

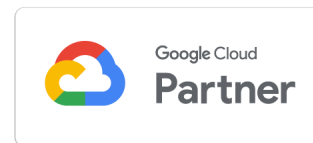


b.telligent

smart data. smart decisions.

About b.telligent GmbH & Co. KG

Technology-independent consultancy specialized in BI, big data, data science and cloud technologies, serving 300+ customers in Europe since 2004.



Products

Google Cloud Platform