Transforming Data Observability in Pharma

Client Success Story





Industry

Pharmaceutical (Human Pharma)

Client

One of the leading international pharmaceutical company

Solution

Observability – FinOps and monitoring platform

Challenges

- Data pipelines were not actively monitored which led to poor SLA and lack of notifications about failed ETL jobs
- Lack of audit of the most used datasets and frequency of their freshness and volume led in lack of prioriterization of the ETL jobs
- Lack of performance metrics about the platform led to unoptimal usage of the Databricks clusters and high compute costs
- Lack of possibility to send custom data metrics led to insufficient knowledge about the crucial parts of the ETL workloads
- There was no dashboards which could indicate issues with the ETL pipelines and usage of the data

Solution

- We implemented a monitoring platform which gathered metrics related to performance and freshness and volume of the data. This led to better understanding of the data workloads
- We introduced PowerBI dashboards as a single source of truth which could be used as an indicator showing potential issues with the data
- We were albe to obtain the data about most popular queries and tables which led to higher prioritization of respecitve ETL jobs
- We were able to identify not used datasources and eliminate not needed computer resources leading in reduce costs

Benefits

- 10% costs reduction of the compute in lower priority ETL workloads
- Raising awareness and introduction of active monitoring with higer SLA of the data pipelines led to reduce number of errors in the jobs and data was available faster and were more reliable
- Thanks to BI dashboards it was easy to check if my reports are using fresh data and if not it was possible to investigate what was the root cause of the problem
- Thanks to OpenTelemetry custom metrics being available the more sophisticated jobs could send advance metrics and check their specific parameters crucial for the ETL process