

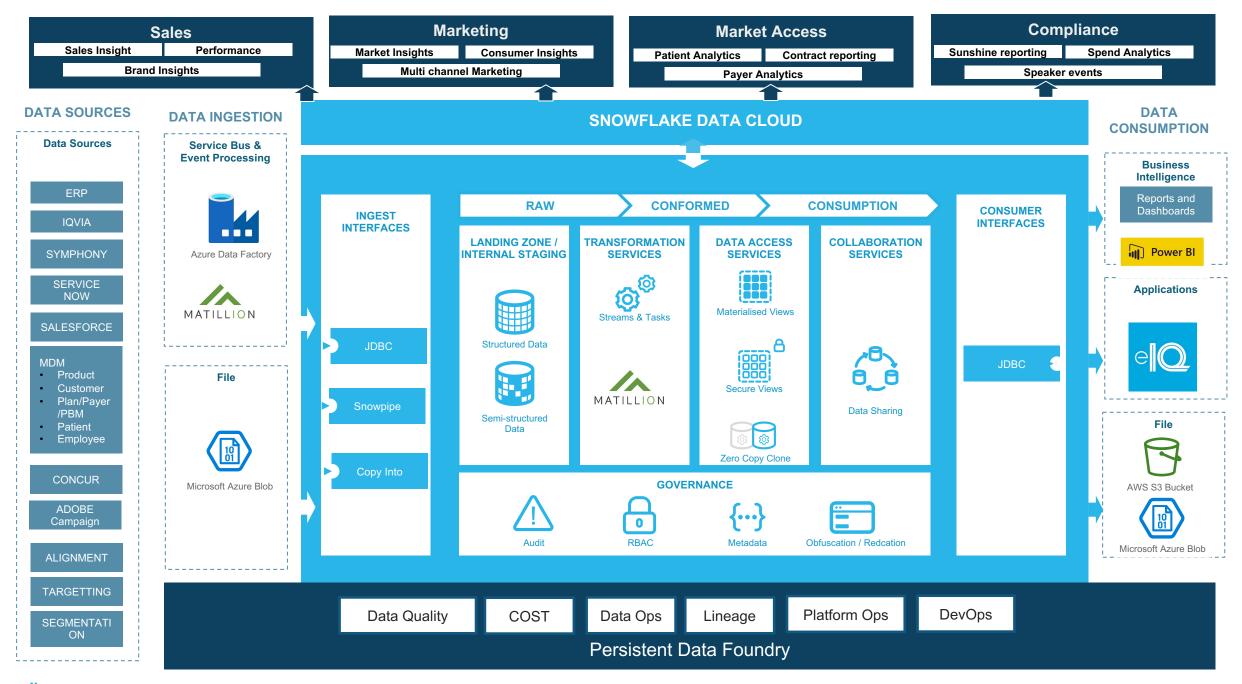
PERSISTENT

Life Sciences Commercial Data Platform

SNOWCASE

ARCHITECTURE TEMPLATE

Snowflake, Persistent Systems



Persistent Data Foundry	Cloud data platform built on top of Snowflake which provides 1. Tight integration of various platform tools (ETL, BI, Jupyter notebooks, ML ops, monitoring tools etc) with Snowflake 2. Inbuilt accelerators for Data Quality, Cost Monitoring, Operations Monitoring and Lineage Tracking 3. Deployment automation for workspace creation
Data Sources	The solution provides prebuilt capability to connect to various sources like IQVIA, Symphony, Service Now, Salesforce etc and ingest data into Snowflake. The solution uses API based connectivity for SaaS sources like Salesforce, and file-based connectivity for the systems which simply provide their extracts on FTP locations.
Azure Data Factory, Matillion	Used for data ingestion and transformations across various data zones in Snowflake. There is predefined transformation logic implemented in ADF and Matillion for formatting, cleansing, transforming and merging the datasets. Further, a metadata driven ELT model is built to ingest and transform data from file-based source.
Snowflake Data Cloud	Snowflake is used as the underlying data platform for the commercial data warehouse (CDW). The data is managed using multiple zones in Snowflake (Raw, Conformed and Consumption). The consumption zone is used as source to build the KPIs and reports in PowerBI. Snowflake is also used to efficiently execute the Data Quality rules, execute complex transformations in ELT framework, storing the history of data across the zones, and storing the audit logs. In nutshell, Snowflake is the central nervous system of the solution.
Power BI	PowerBI is used to create canned reports on top of the tables/views created in Snowflake. Various Metrics and KPIs related to direct and indirect sales, calls, activities, opportunities, events, inventory, rebates etc are visualized using PowerBI.