databricks

Competitive Positioning

Kevin Barlow
Senior Specialist Solutions Architect

Peter Sprague
Sales Performance Specialist, C&SI Partners

databricks

Competitive Positioning

1. Introduction and Market Landscape
2. Comparing Databricks to Snowflake
3. Using cloud native solutions and the DIY approach
4. Why Databricks?

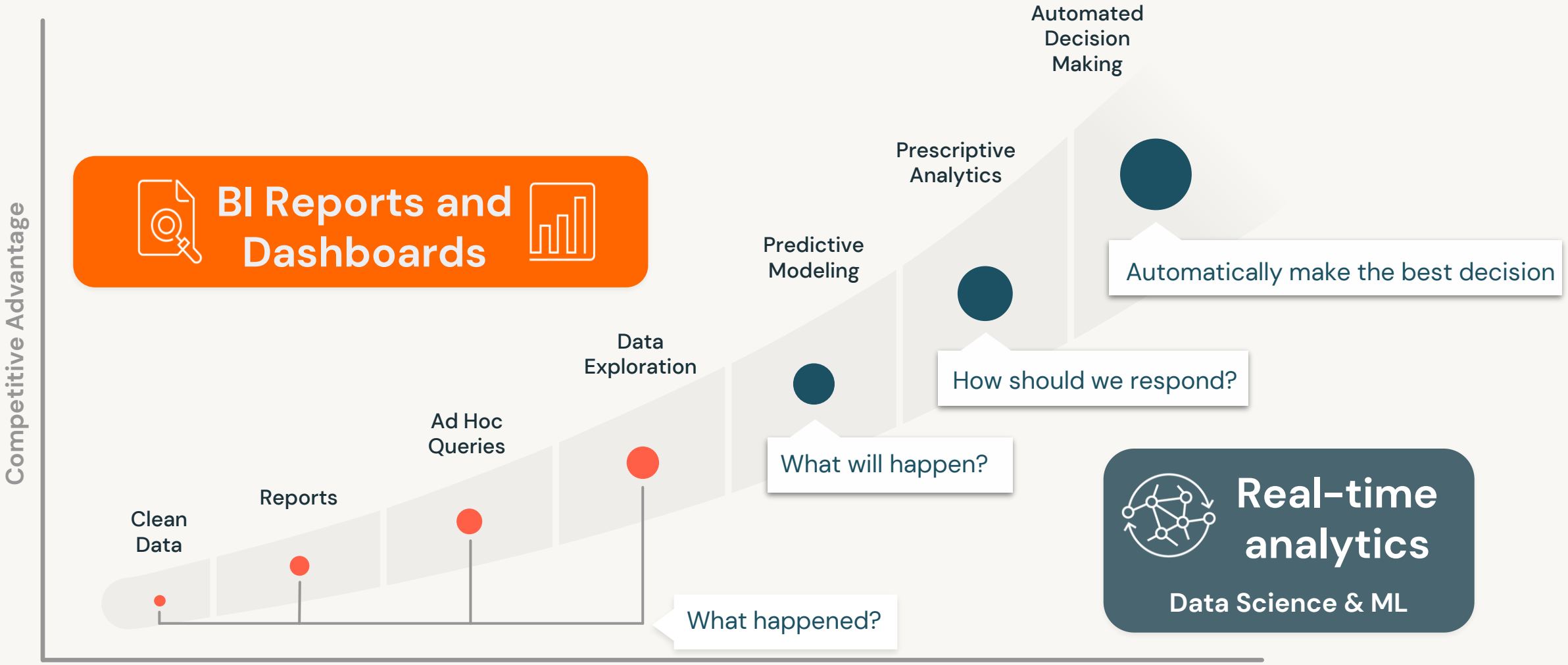
databricks

Competitive Positioning

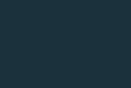
Introduction and Market Landscape

What are the categories of use cases that our customers are seeking?

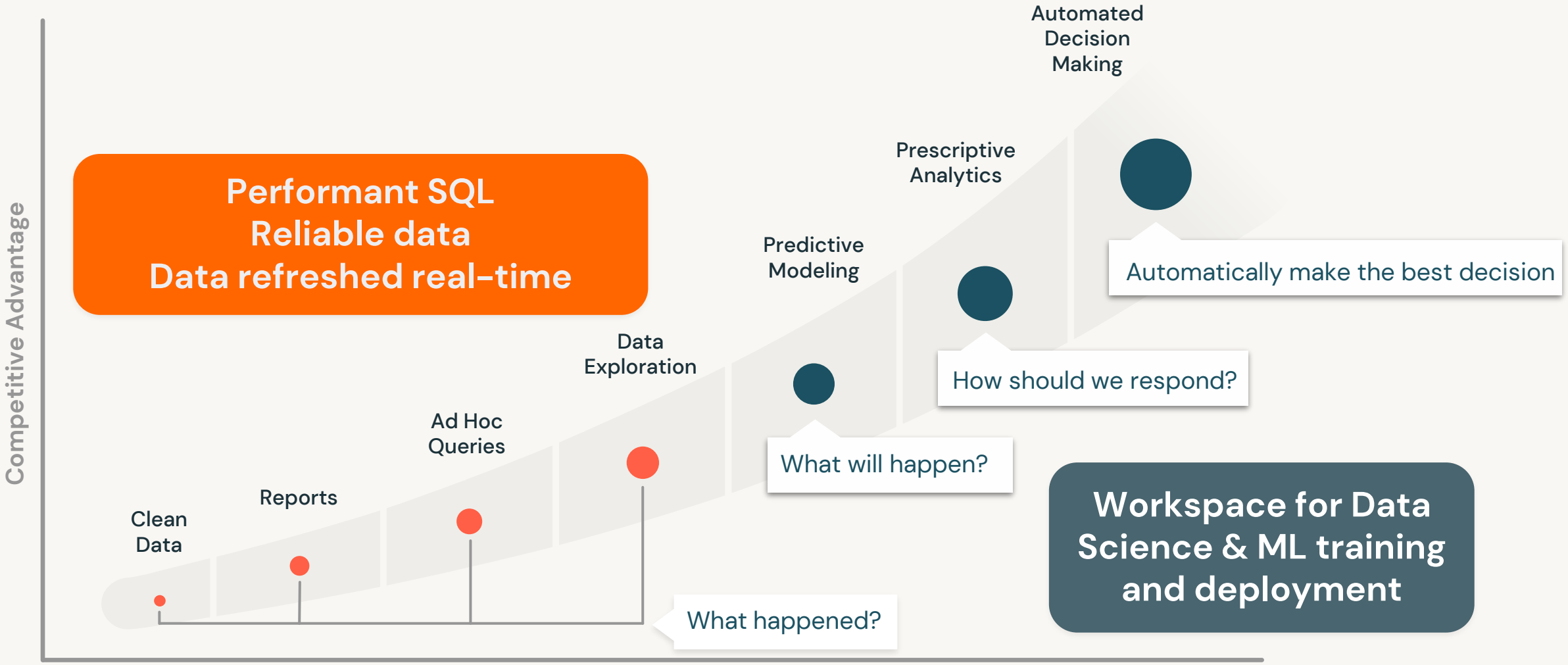
Data Maturity Curve: From Hindsight to Foresight



What capabilities are needed to deliver those use cases?



Data Maturity Curve: From Hindsight to Foresight

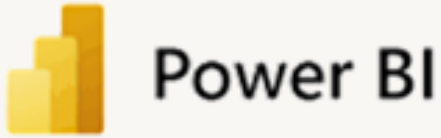


Market landscape

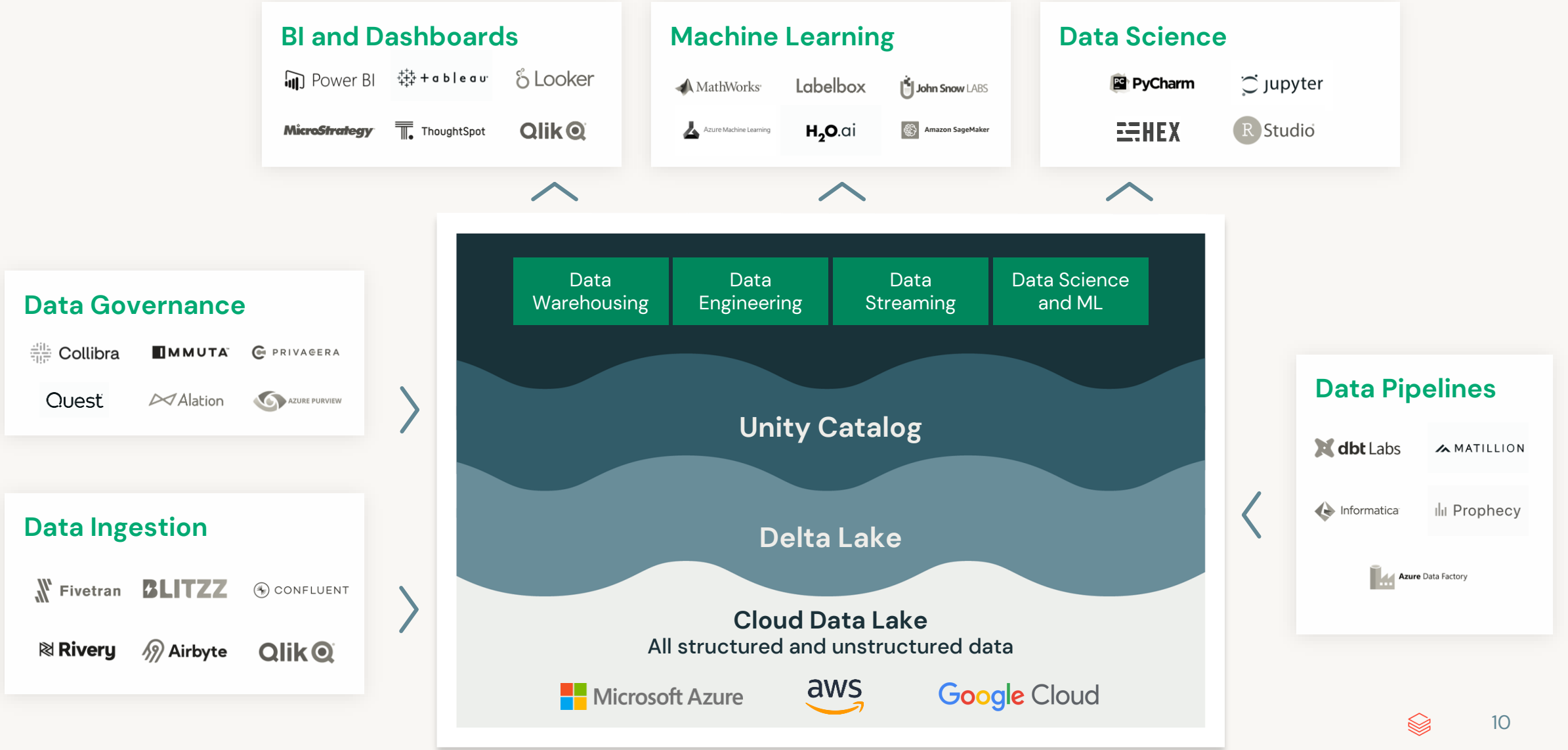
Legacy Vendors

Market landscape

Point Solutions



Databricks thrives within your modern data stack



Market landscape

Open Source

Market landscape

Cloud Provider Native Solutions

EMR

Redshift

SageMaker

Azure Synapse Analytics

Cloud Dataproc

Cloud Dataflow

BigQuery

Market landscape

Platform Providers

Market landscape

Point Solutions

DataRobot

dataiku

Starburst

Open Source—DIY

Apache Spark

Apache hudi

ICEBERG

DELTA LAKE

Cloud Provider Native Services

Azure Synapse Analytics

HDInsight

Amazon EMR

Google Dataflow

Amazon Redshift

Google BigQuery

Amazon SageMaker

Cloud Dataproc

Platform Providers

databricks

C3.ai

Palantir

snowflake

CLOUDERA

Legacy Vendors

teradata.

NETEZZA

sas

CLOUDERA

informatica powercenter

IBM DataStage

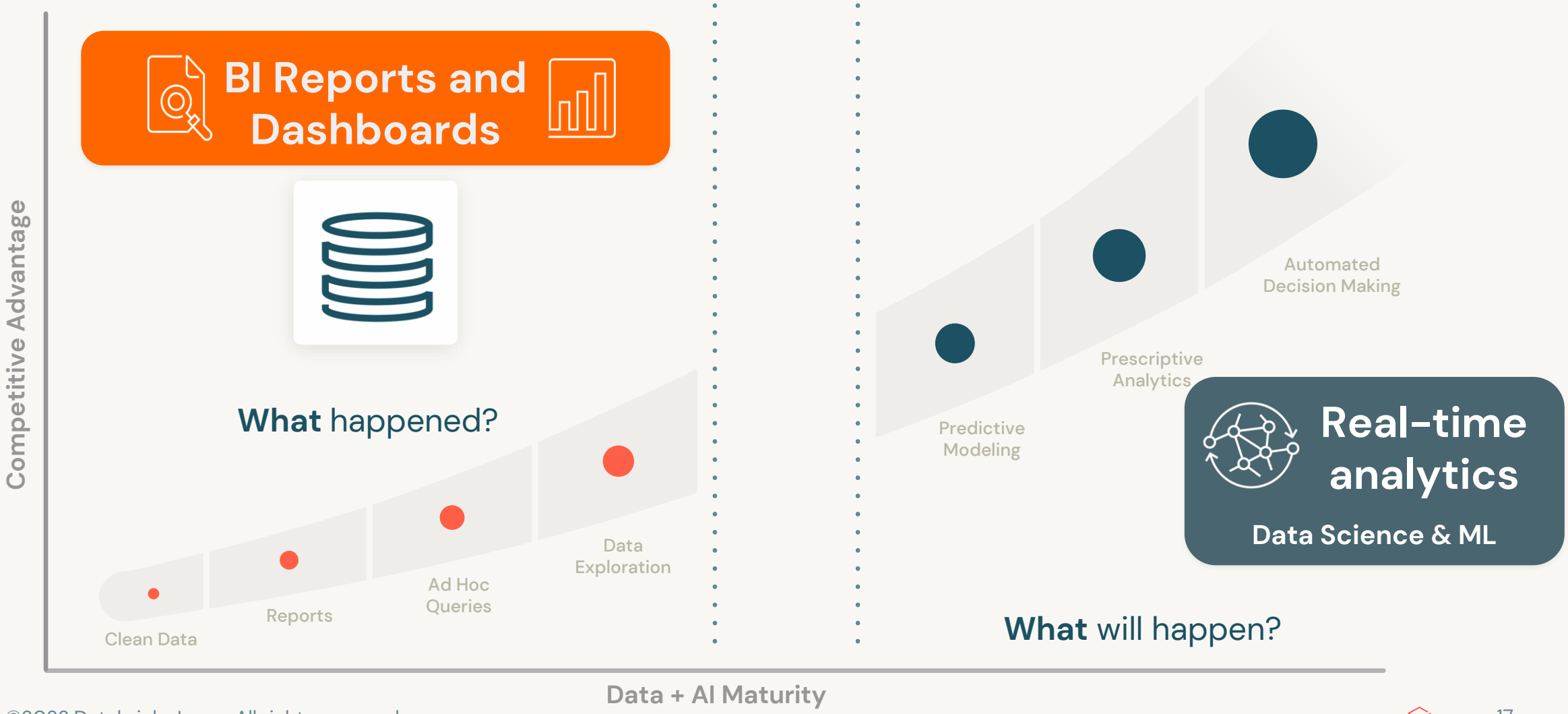
databricks

Competitive Positioning

Comparing Databricks to Snowflake

DATA CLOUD

Snowflake is a cloud data warehouse



Issues Snowflake Customers Will Face

High costs

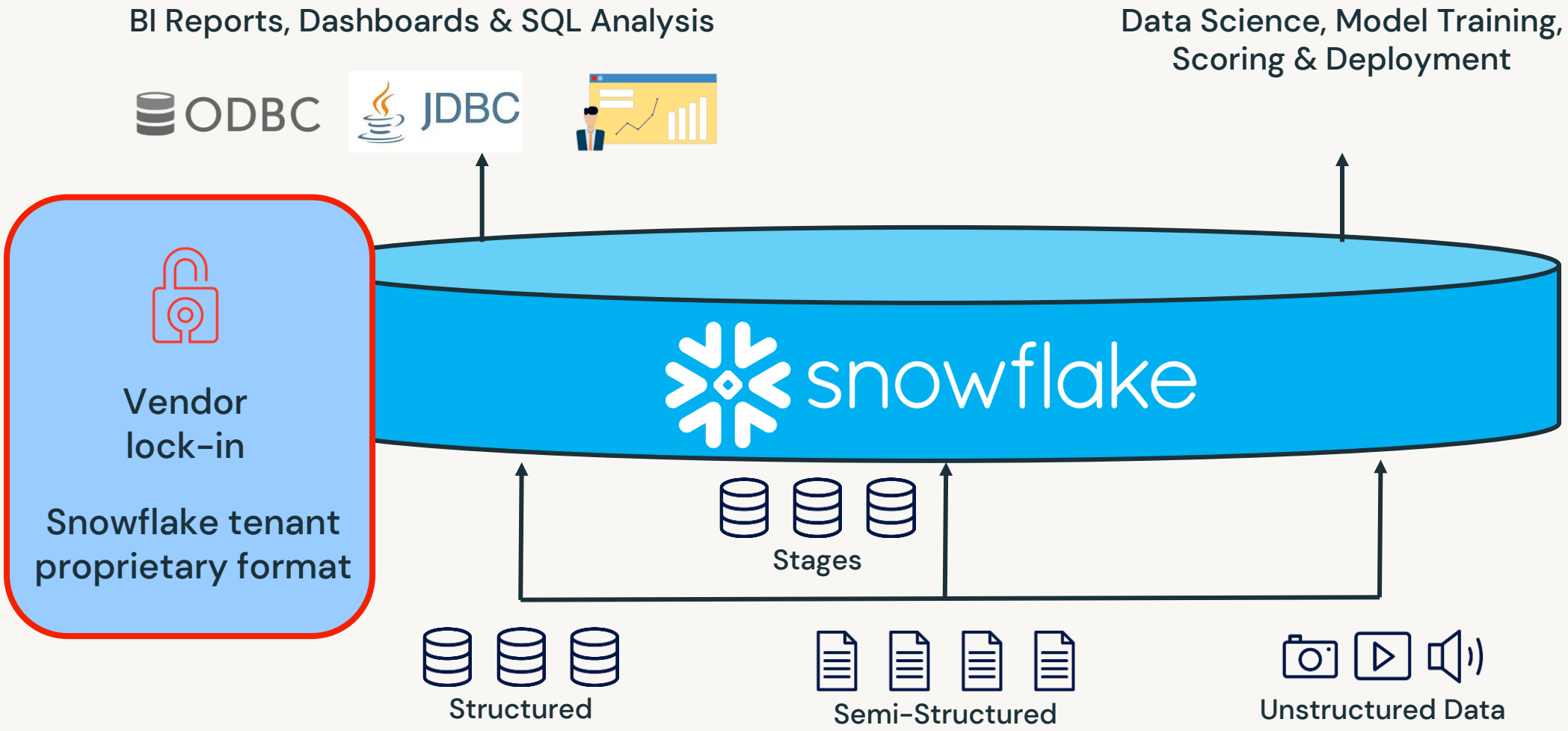
Inefficient data engineering

Incomplete data support

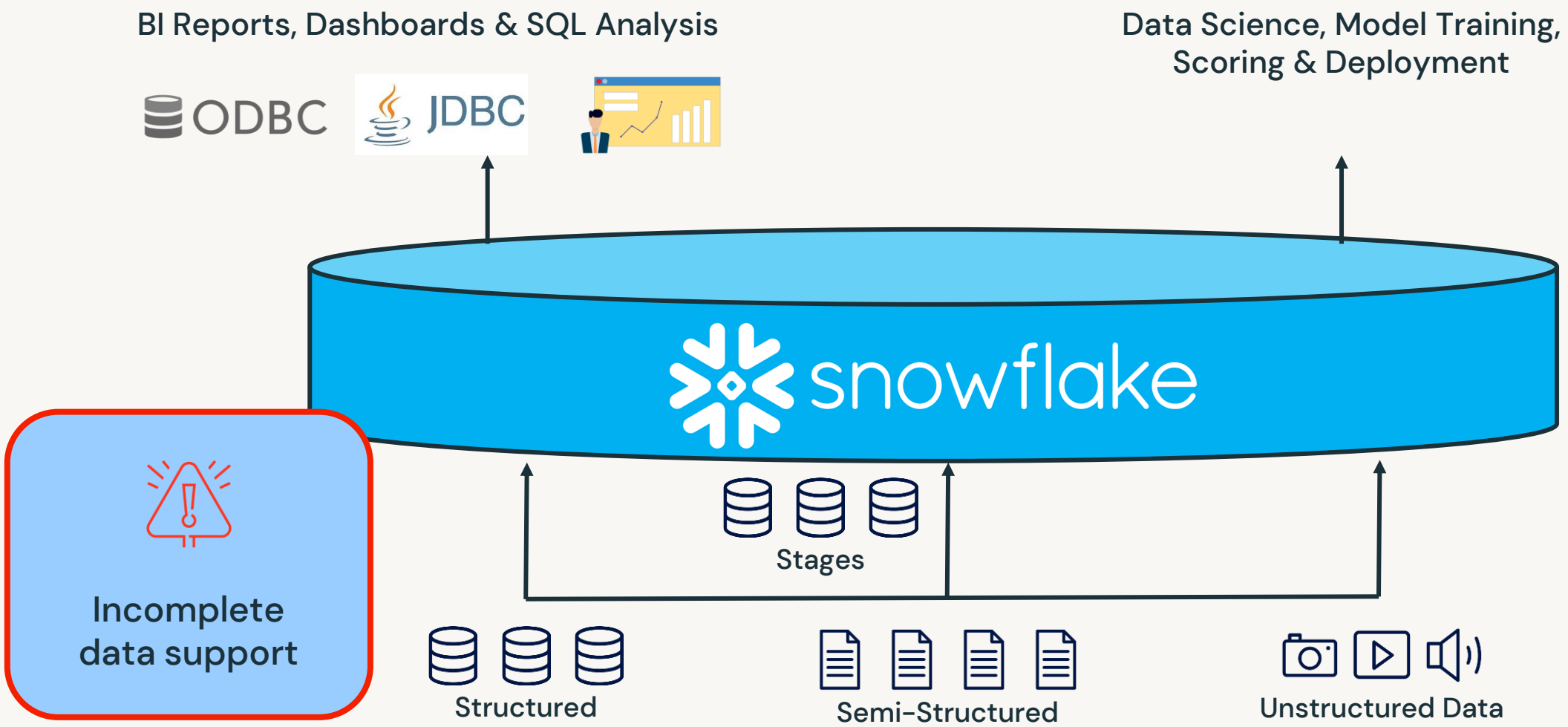
Limited DS/ML capabilities

Vendor lock-in

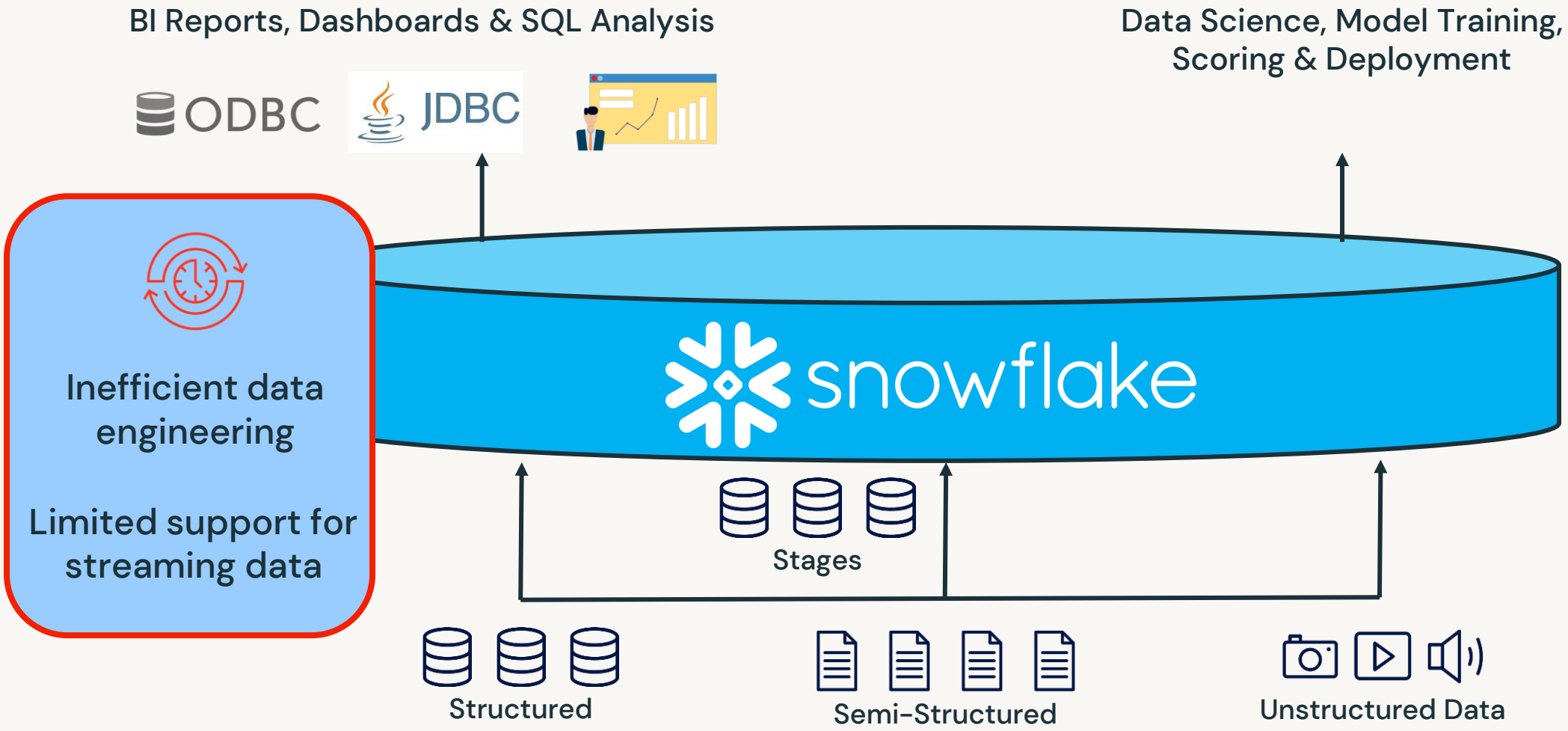
Issues with the Snowflake Data Cloud



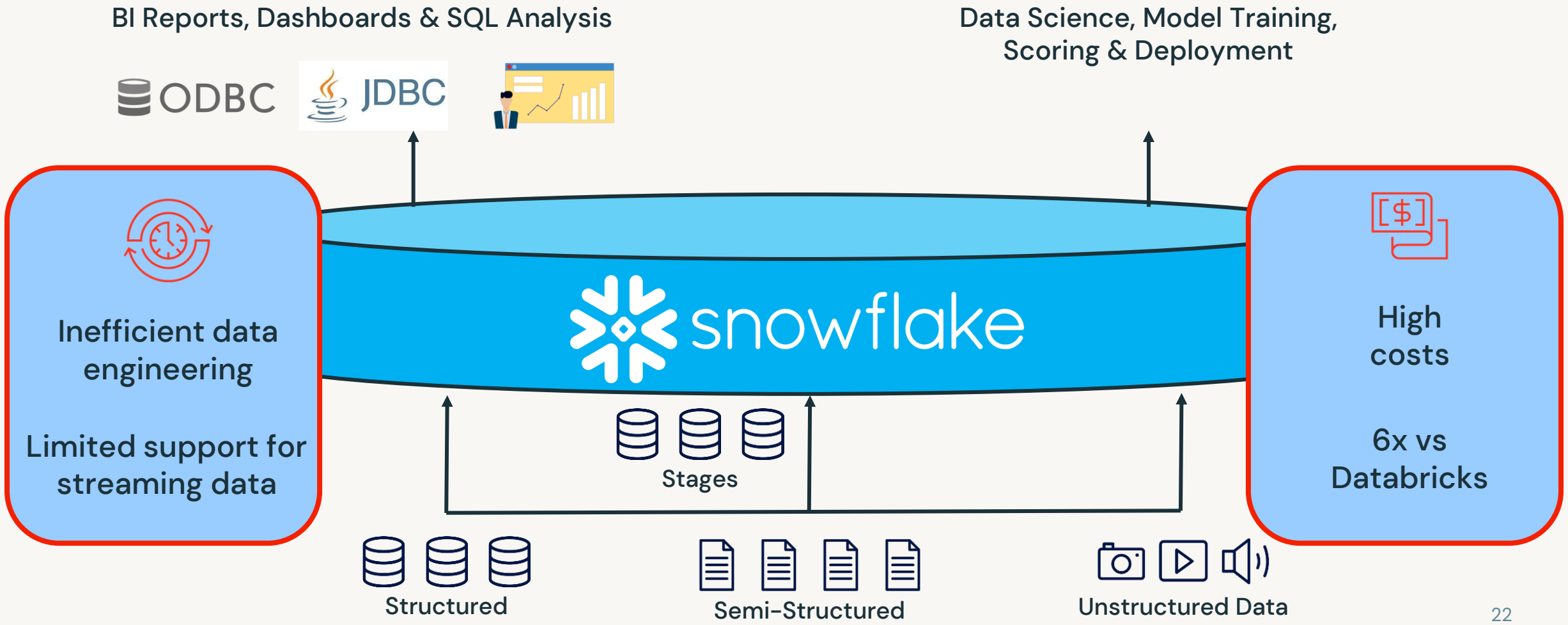
Issues with the Snowflake Data Cloud



Issues with the Snowflake Data Cloud



Issues with the Snowflake Data Cloud



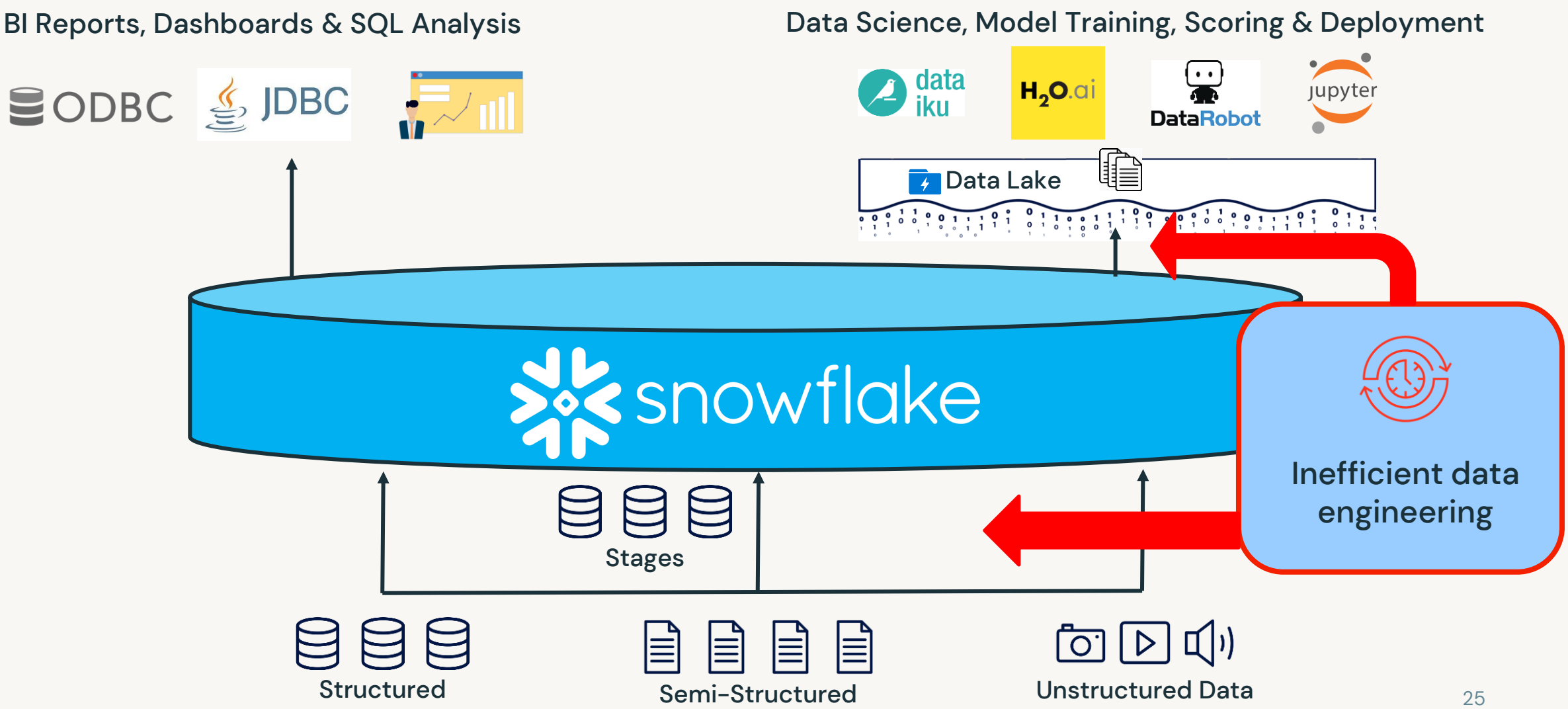
snowflake

DATA CLOUD

Move date to another system >
Increases total cost of the architecture
AND….lower time to value for customer

Databricks is built for optimized data engineering. Snowflake is not. Why does this matter?

Issues with the Snowflake Data Cloud



Databricks thrives within your modern data stack

BI and Dashboards

Power BI · tableau · Looker · MicroStrategy · ThoughtSpot · Qlik

Machine Learning

MathWorks · Labelbox · John Snow LABS · Azure Machine Learning · H2O.ai · Amazon SageMaker

Data Science

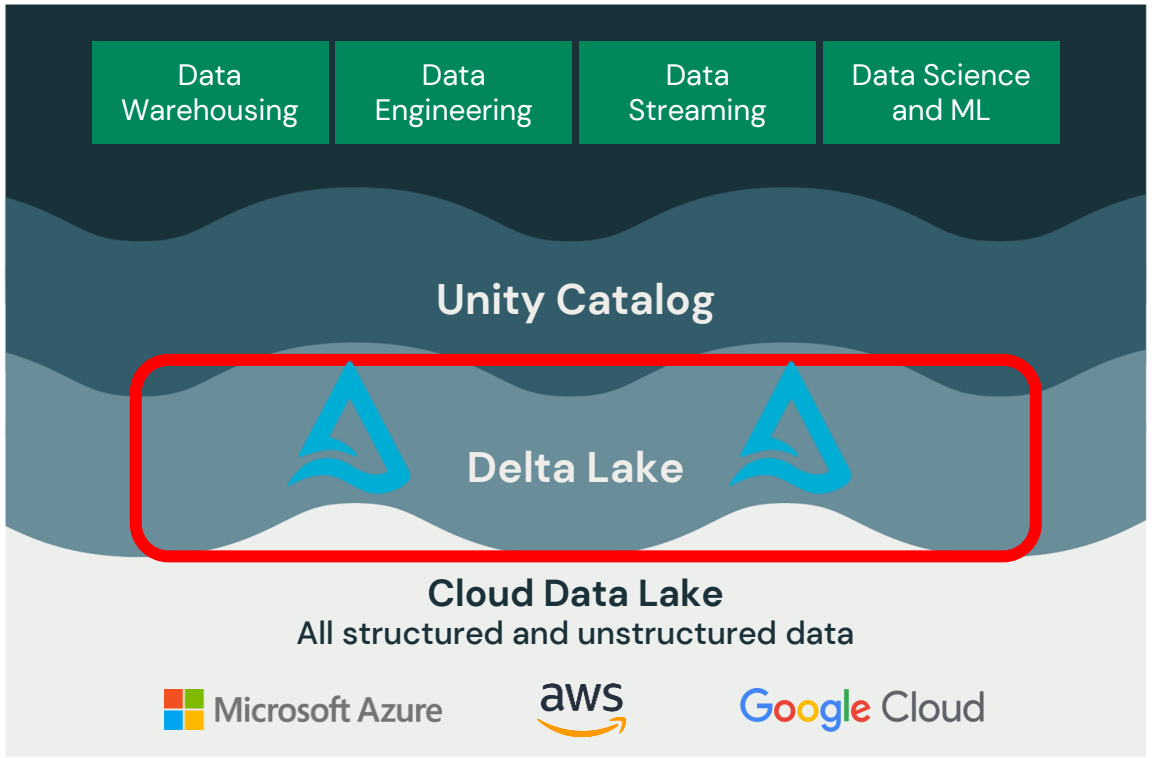
PyCharm · jupyter · HEX · RStudio

Data Governance

Collibra · IMMUTA · PRIVAGERA · Quest · Alation · AZURE PURVIEW

Data Ingestion

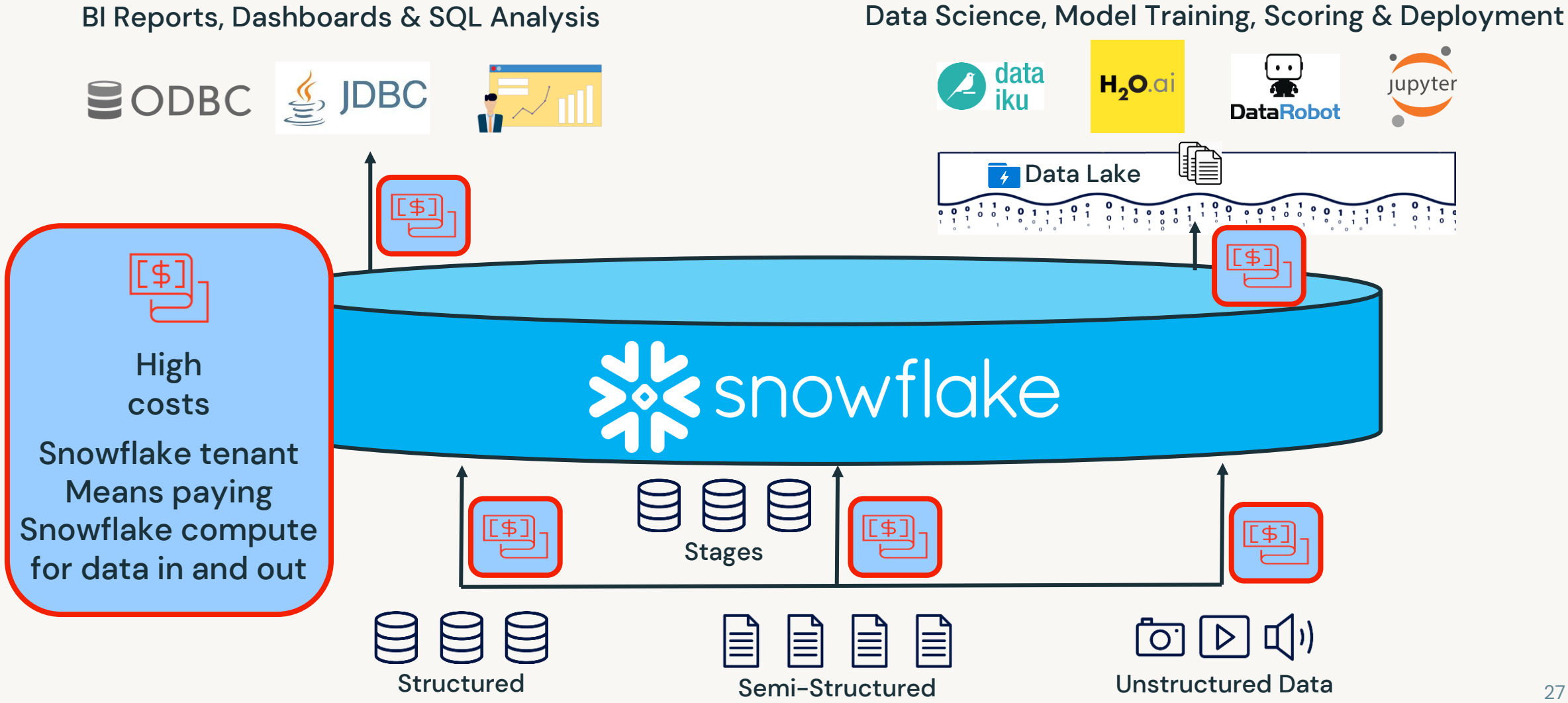
Fivetran · BLITZZ · CONFLUENT · Rivery · Airbyte · Qlik



Data Pipelines

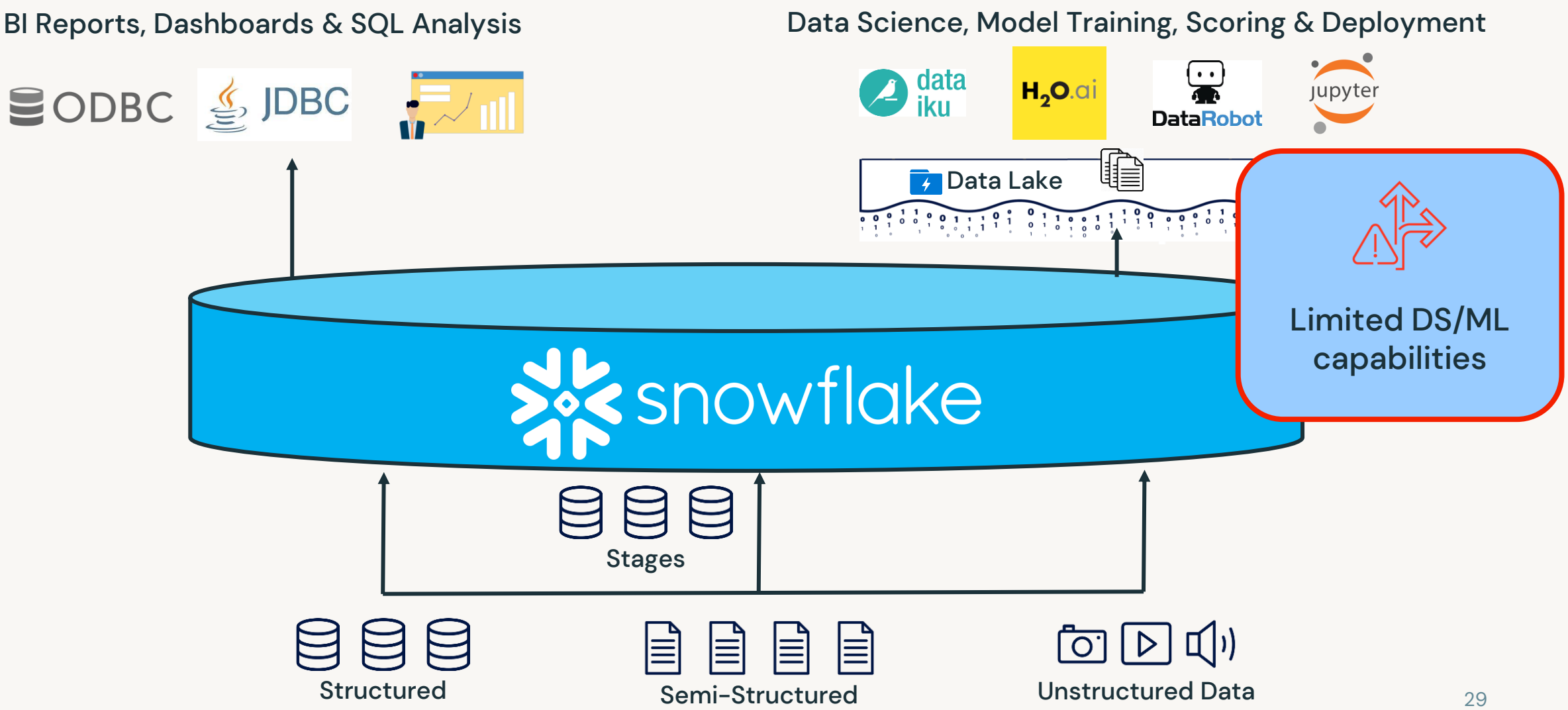
dbt Labs · MATILLION · Informatica · Prophecy · Azure Data Factory

Issues with the Snowflake Data Cloud

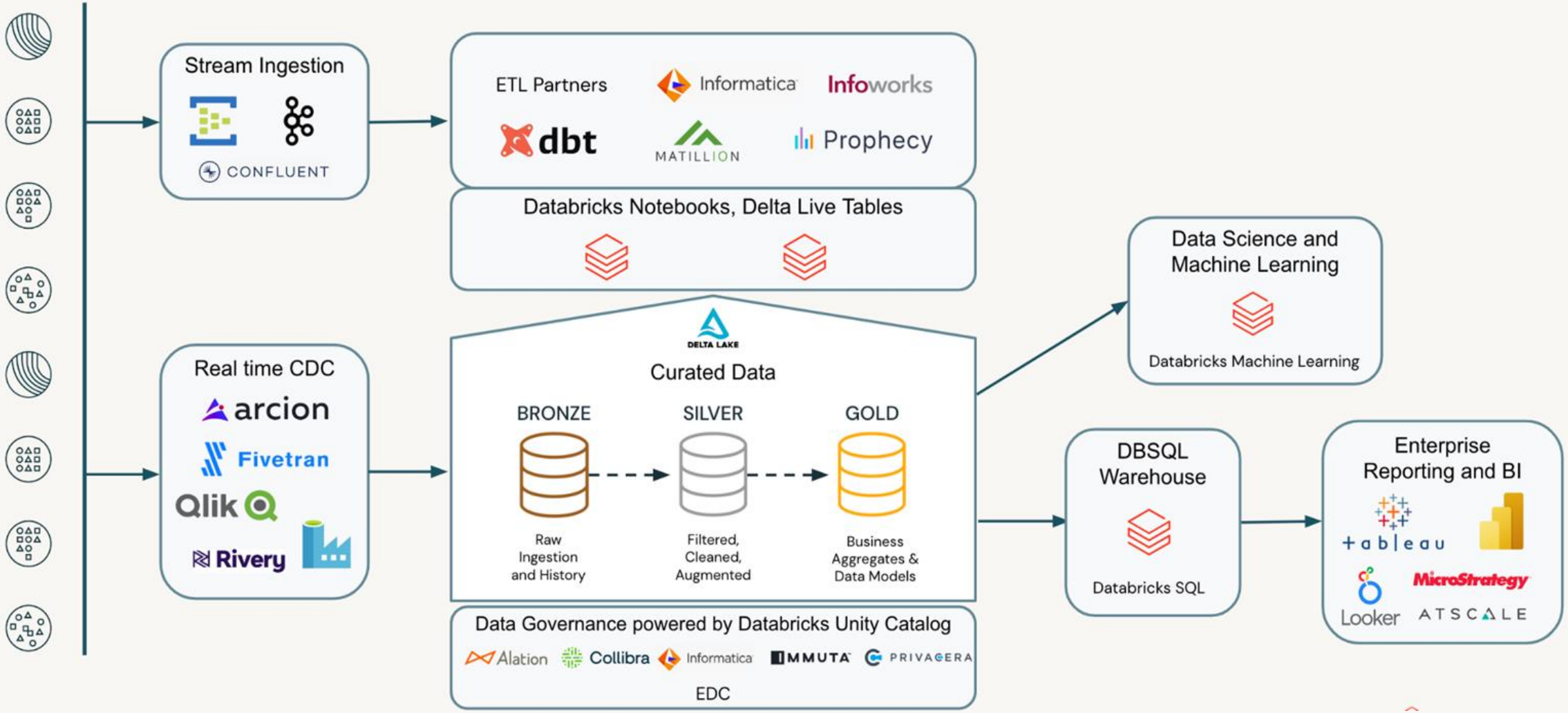


What pain will the customer feel from copying data out of Snowflake for non-SQL workloads?

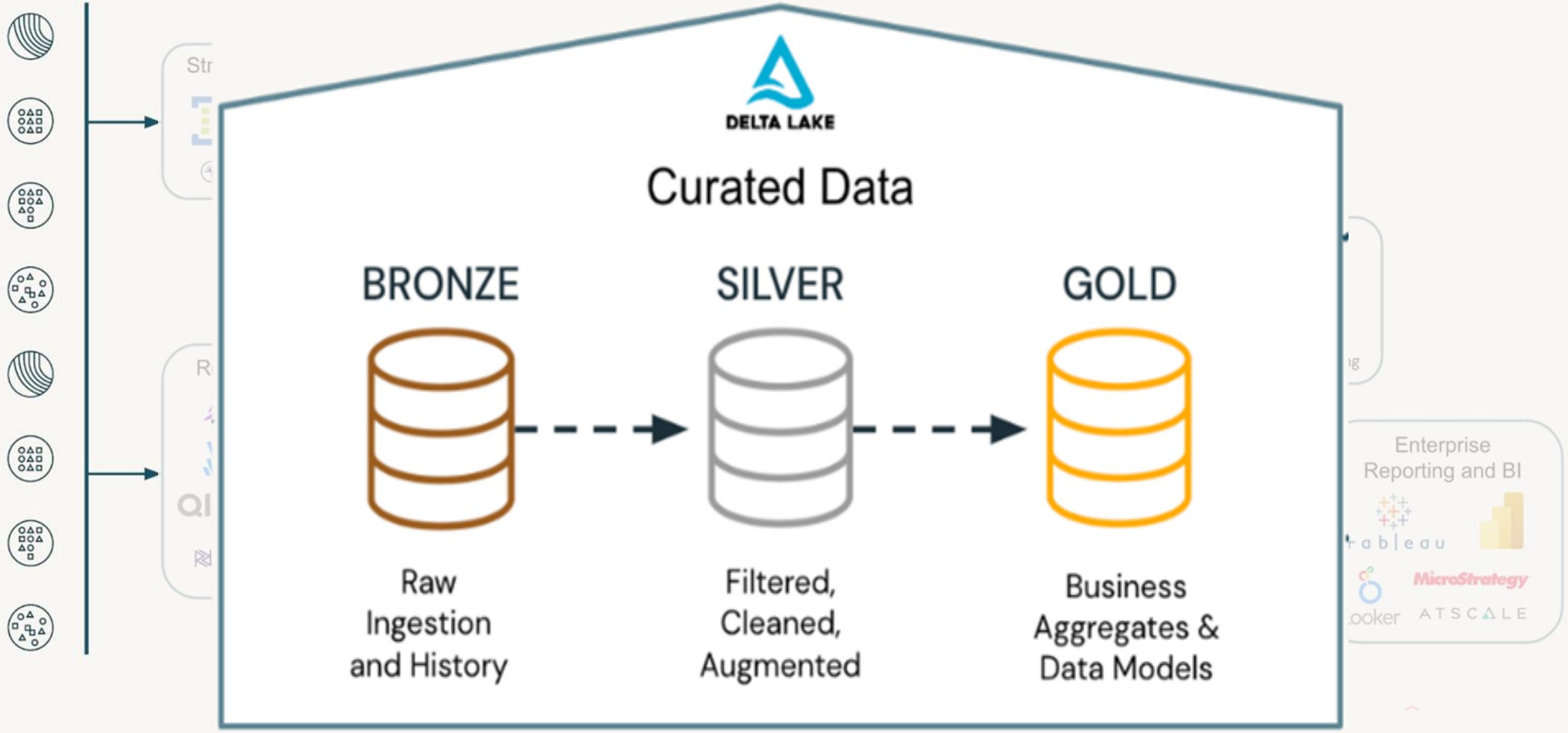
Issues with the Snowflake Data Cloud



Modern Data Warehousing on Databricks

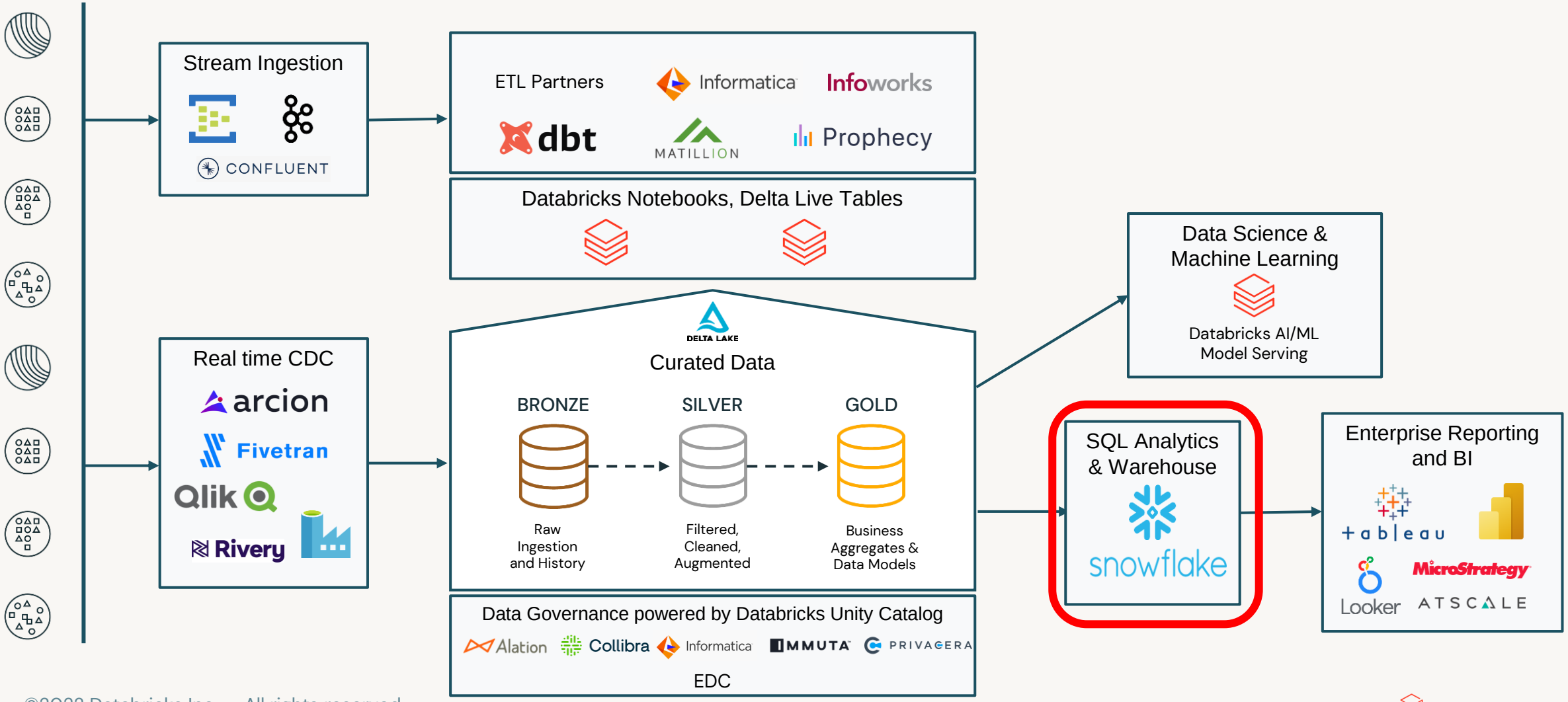


Modern Data Warehousing on Databricks



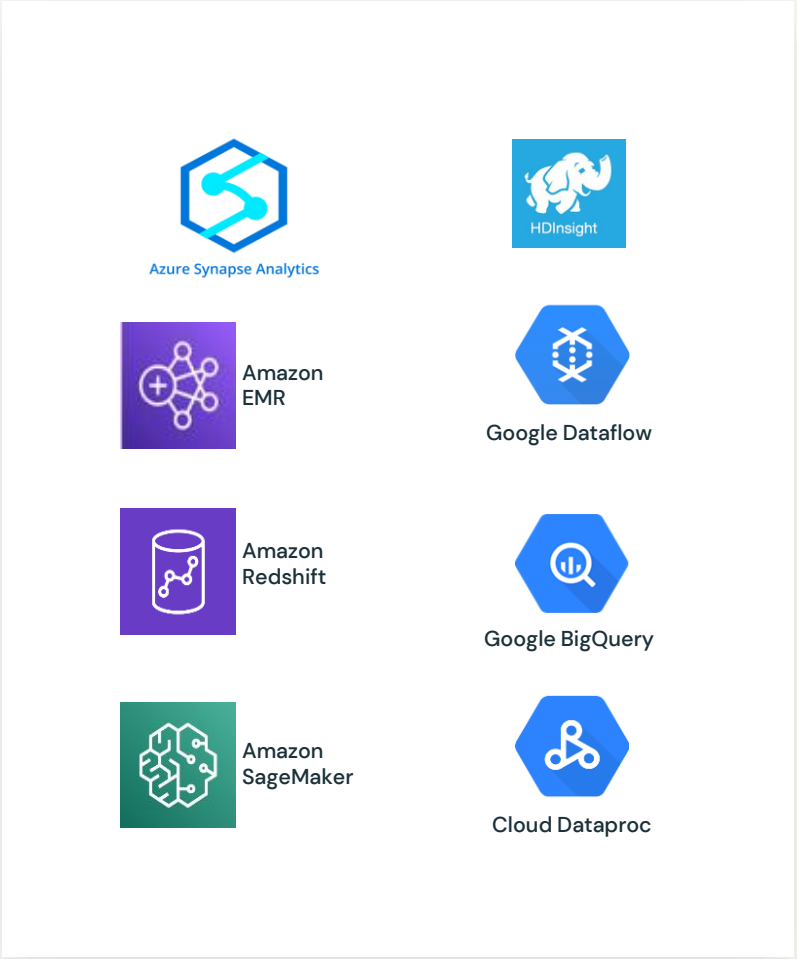
How can Databricks co-exist with Snowflake?

Modern Data Stack on Databricks



What about using cloud native tools?

Cloud Provider Native Services



Advantages

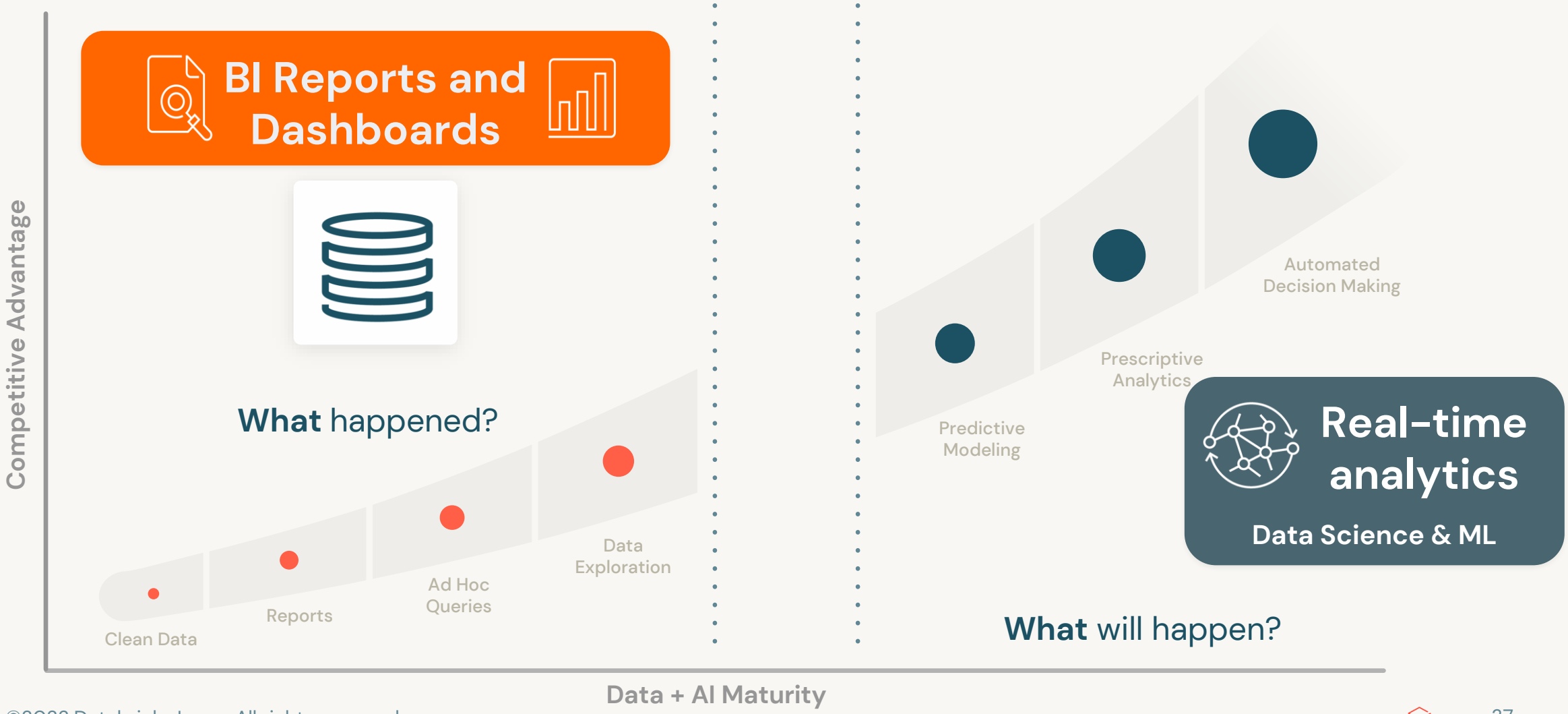
- A large collection of different services to meet your needs
- Easy to deploy
- Single bill makes it easy for administrator

Disadvantages

- Have to maintain many services and connections between them, e.g. ETL Tools with data warehouses, data science tools, etc. > complexity increases
- Cloud services limited to the cloud provider > but multi-cloud is where customers are moving > not cost effective to move data into a single cloud > better to work on it where it lands

What do we mean by DIY?

Different workloads across the Data Maturity Curve



Do it yourself? Stitch them all together!

Data Warehousing	Data Engineering	Streaming	Data Science & Machine Learning
 Data Analysts	 Data Engineers	 Data Engineers	 Data Scientists
"I want best of breed."	"I want best of breed."	"I want best of breed."	"I want best of breed."

Really hard to do!

Large workforce needed

Teams need to talk to each other

One approach: stitch them all together

Data Warehousing

Data Analysts

- Google BigQuery
- snowflake
- Azure Synapse Analytics
- teradata.
- SAS
- Amazon Redshift

Data Engineering

Data Engineers

- Apache Hadoop
- CLOUDERA
- Amazon EMR
- Apache Spark
- Cloud Dataproc
- Apache Airflow

Streaming

Data Engineers

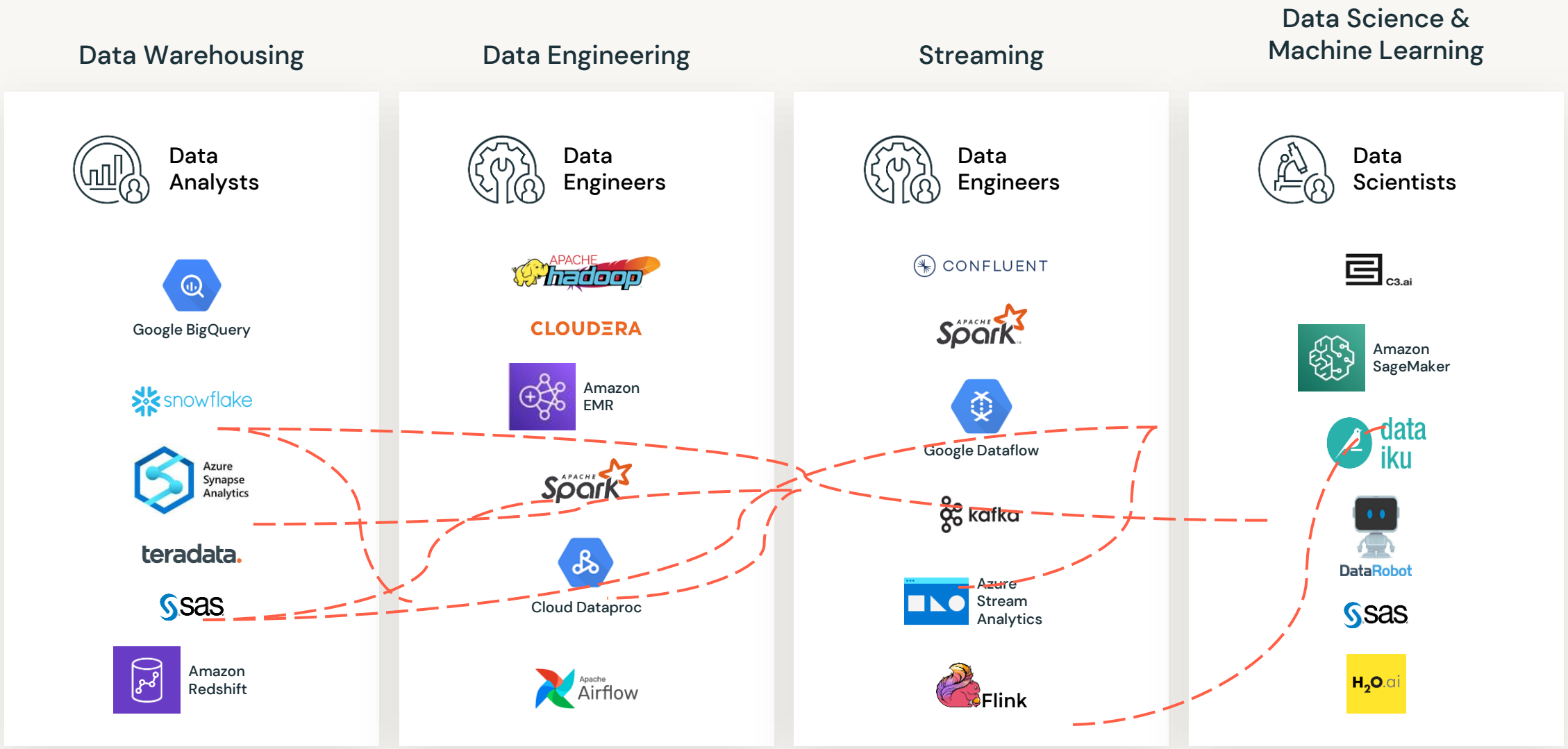
- CONFLUENT
- Apache Spark
- Google Dataflow
- kafka
- Azure Stream Analytics
- Flink

Data Science & Machine Learning

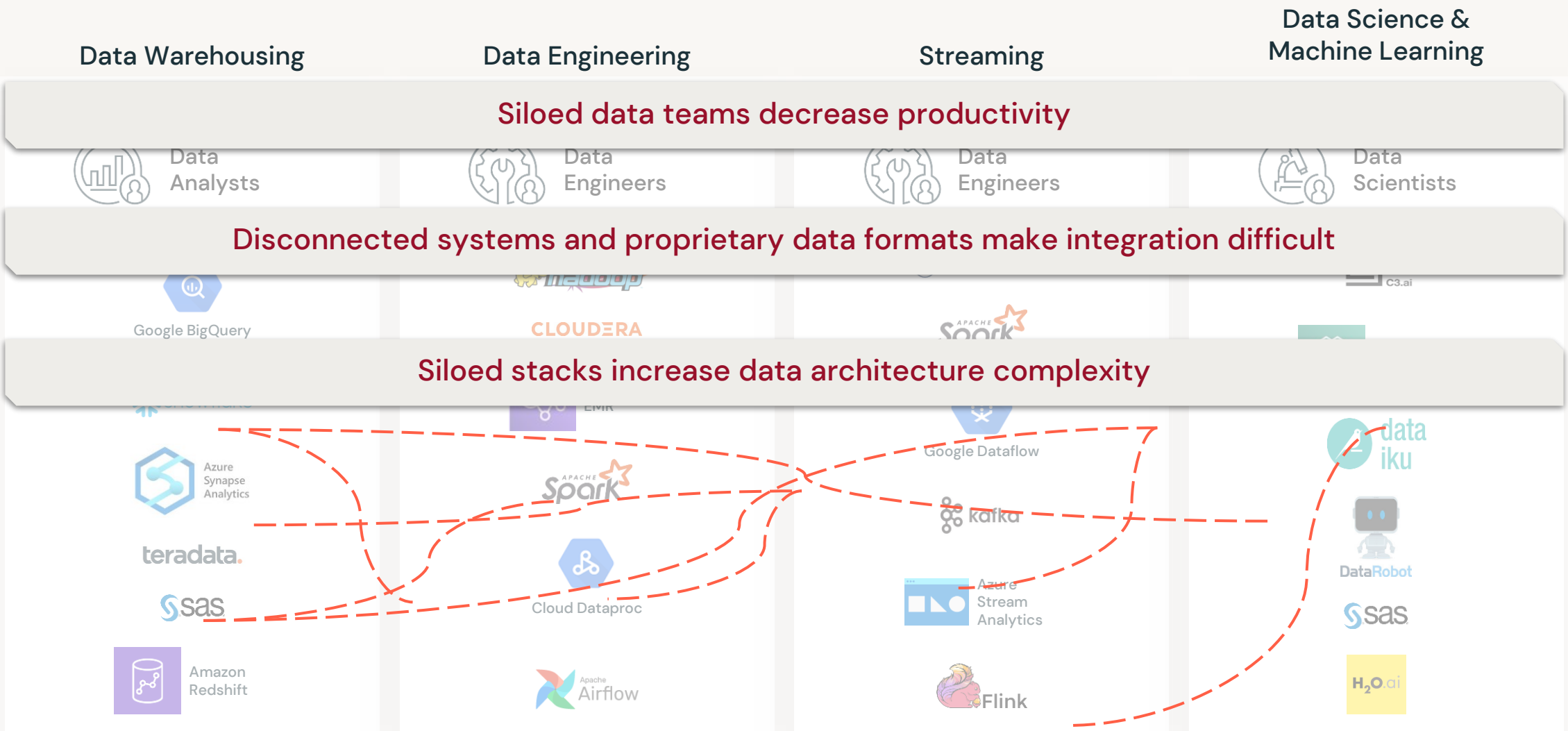
Data Scientists

- C3.ai
- Amazon SageMaker
- dataiku
- DataRobot
- SAS
- H_2O.ai

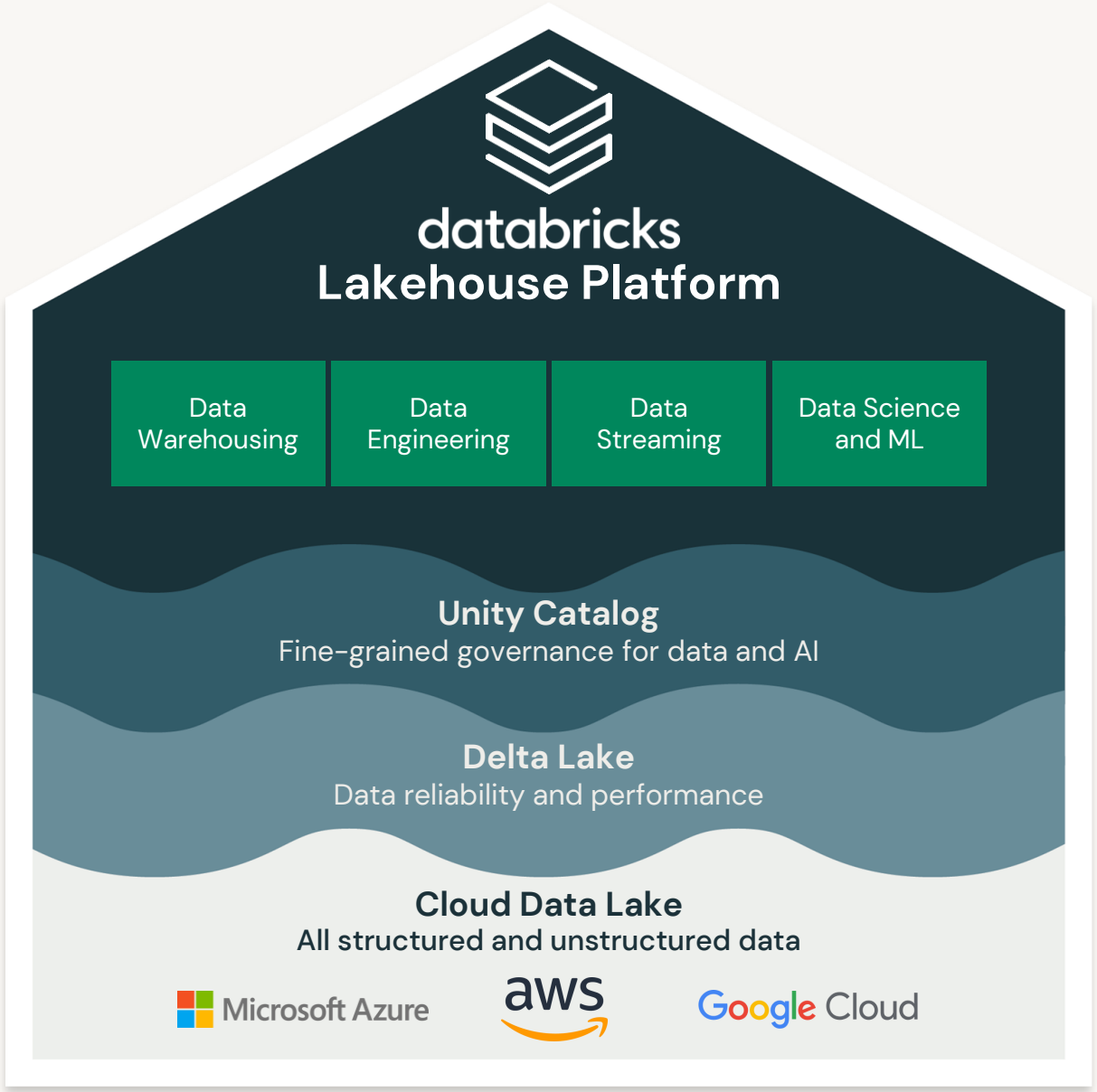
One approach: stitch them all together



One approach: stitch them all together



Why Databricks?



Databricks Lakehouse Platform

Simple

Unify your data warehousing and AI use cases on a single platform

Multicloud

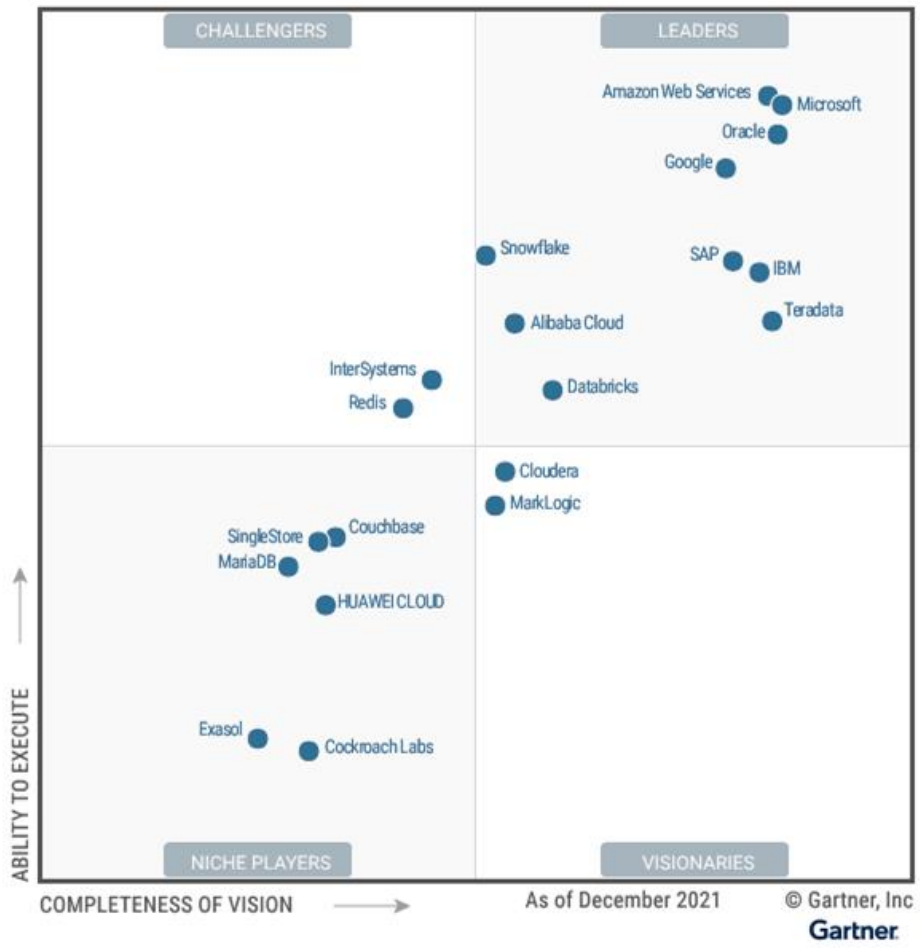
One consistent data platform across clouds

Open

Built on open source and open standards

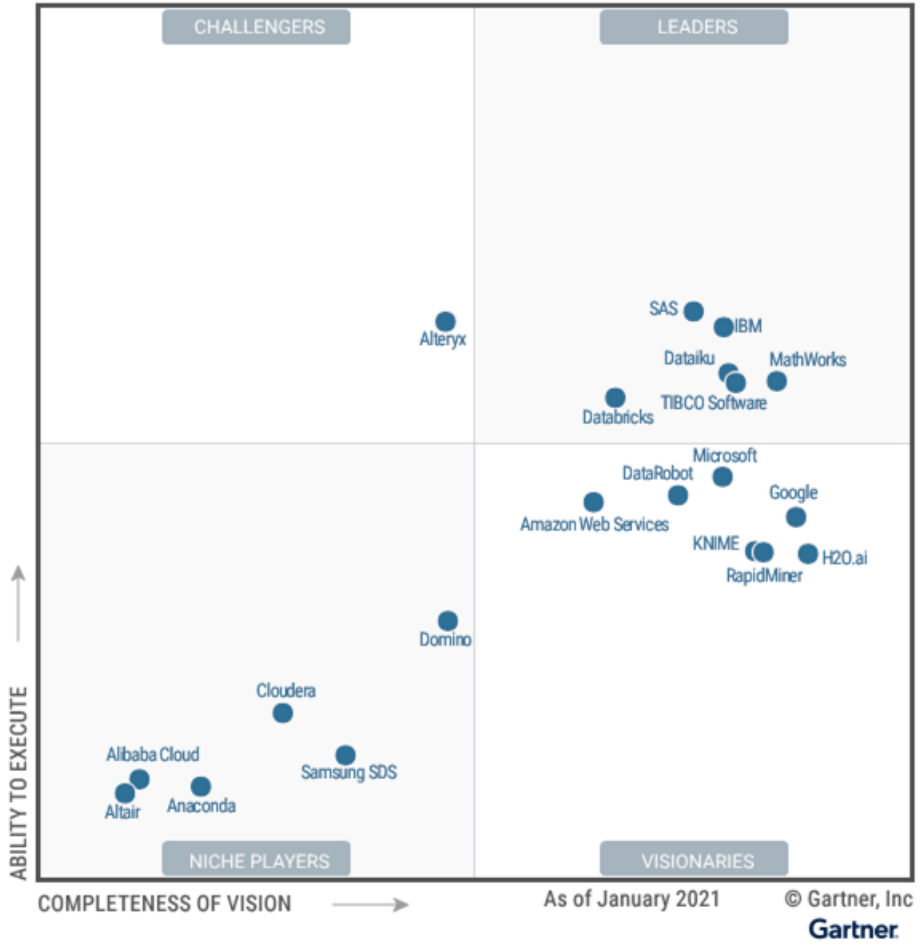
Only platform recognized by Gartner as Leader in two MQs

Figure 1: Magic Quadrant for Cloud Database Management Systems



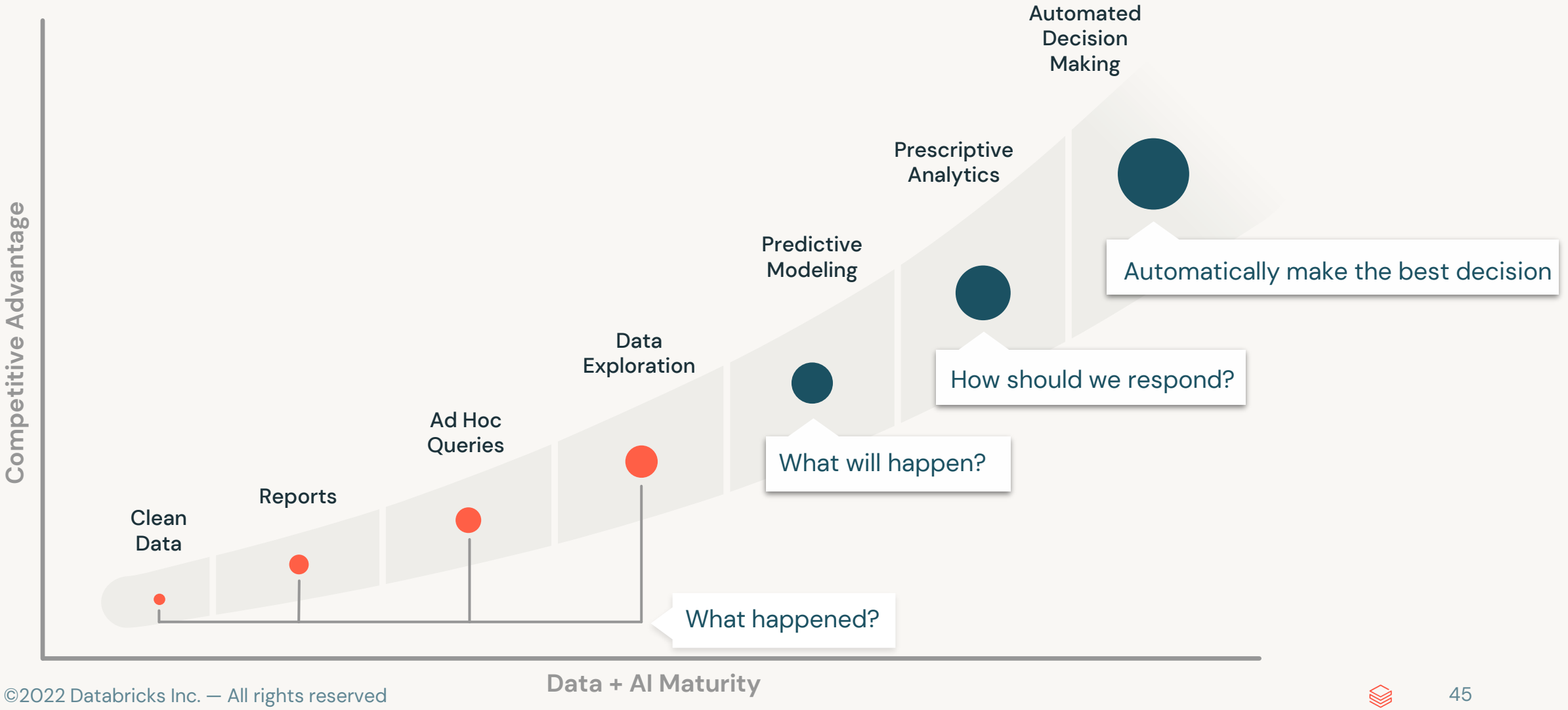
Source: Gartner (December 2021)

Figure 1: Magic Quadrant for Data Science and Machine Learning Platforms



Source: Gartner (March 2021)

Data Maturity Curve: From Hindsight to Foresight

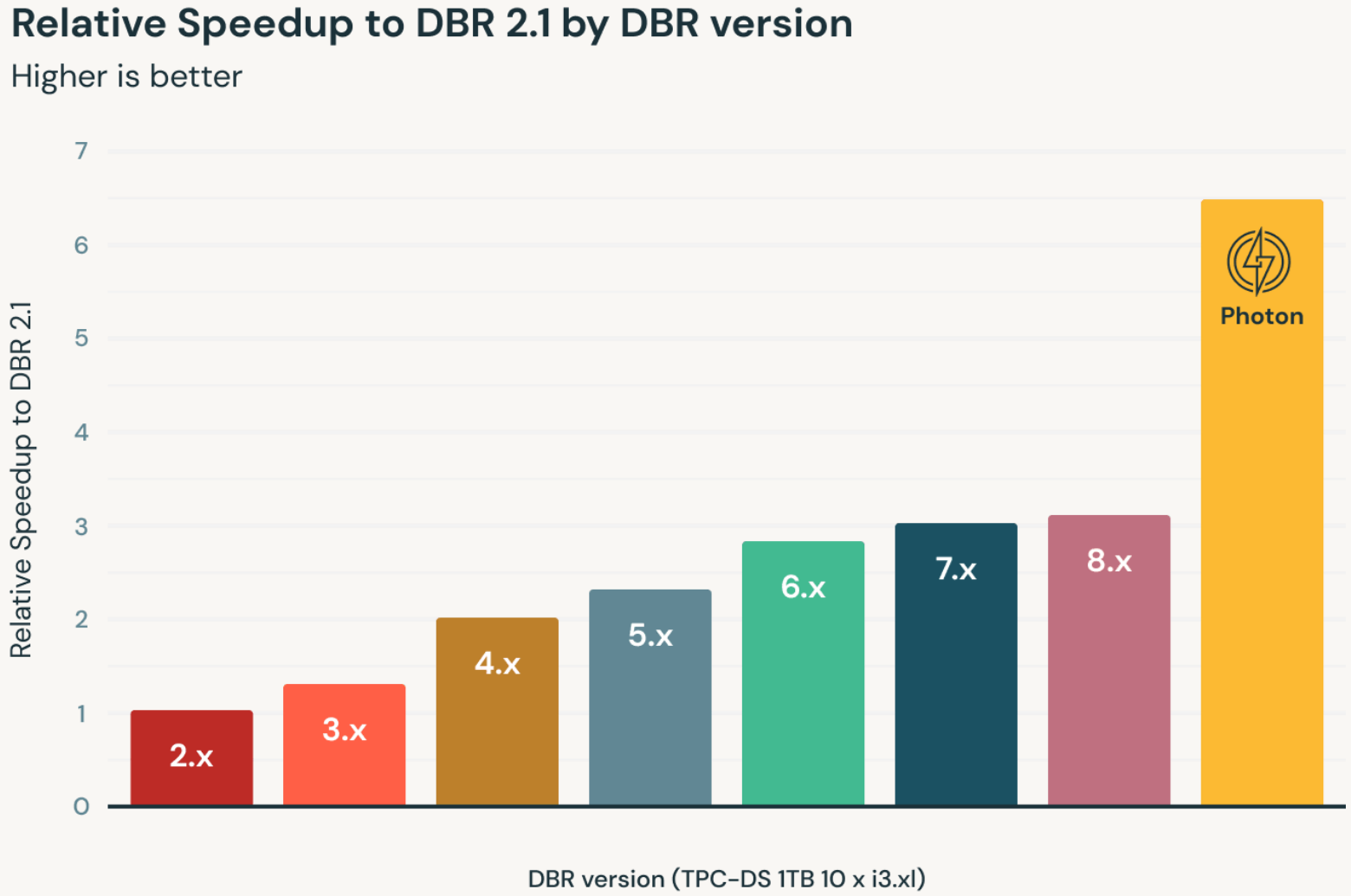


Why build Photon?

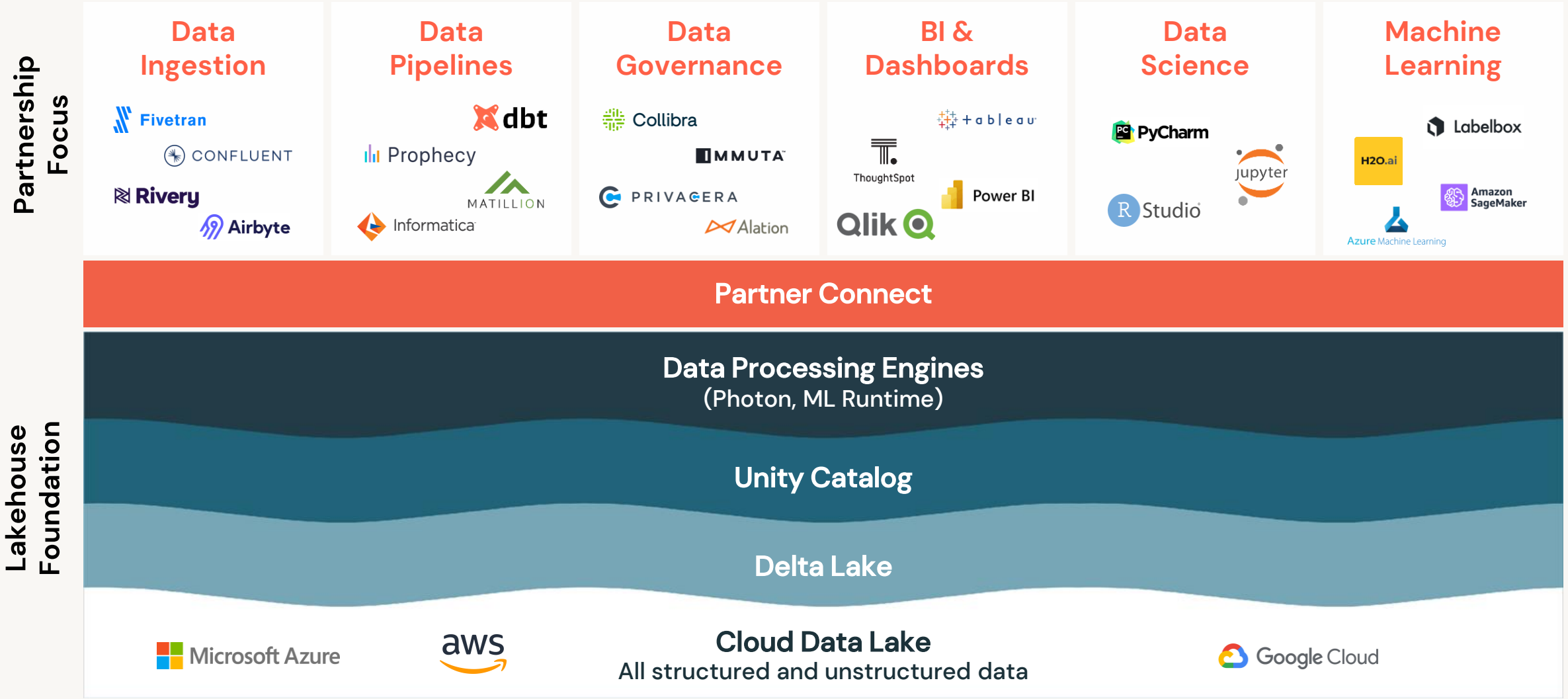
Cheaper and faster

Built for all use cases

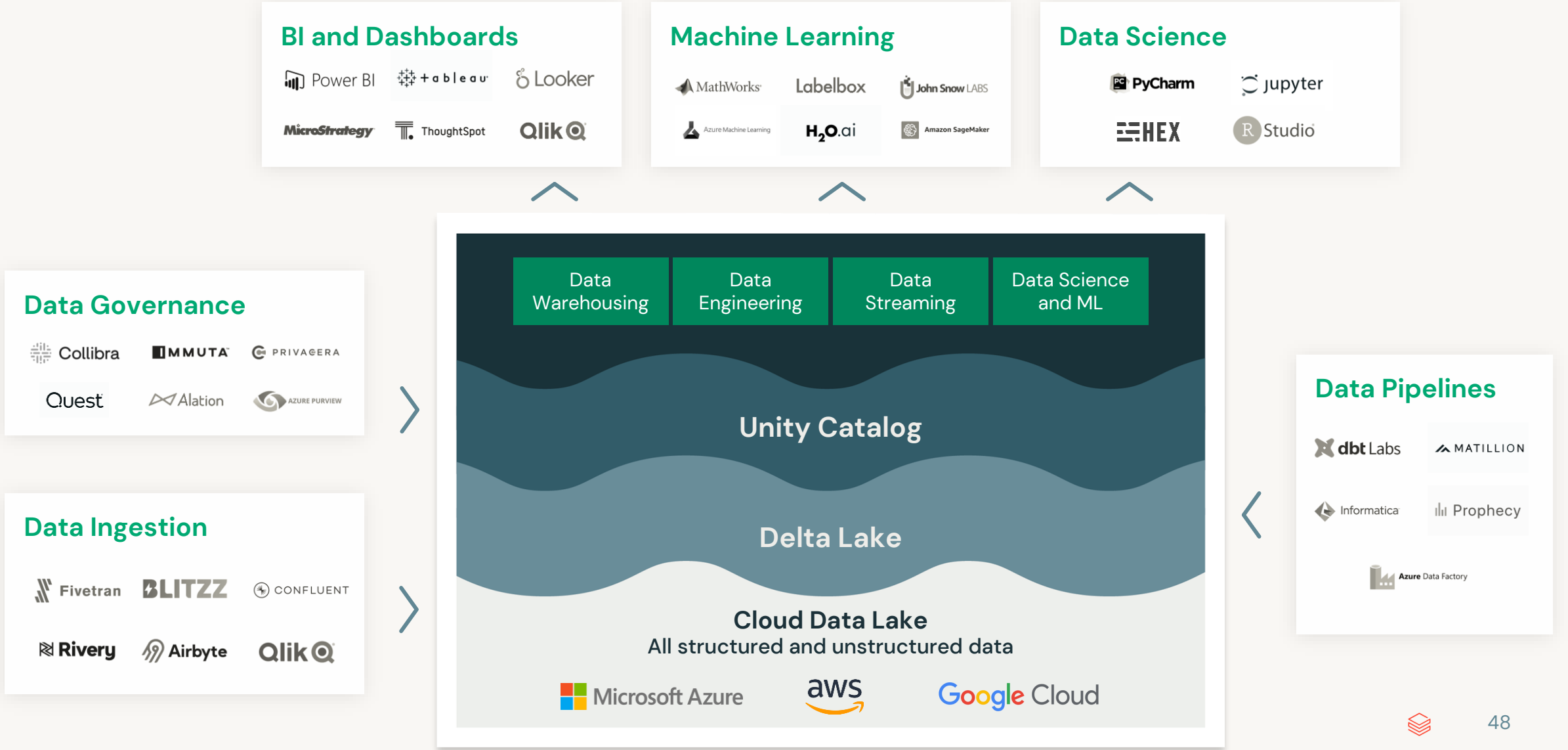
No code changes



Why our ISV Partners are so important



Databricks thrives within your modern data stack



Partner Connect

Gives customers direct access to Partners

Ingest/ETL Partners

Fivetran
dbt
Prophecy
Rivery
arcion
hightouch
MATILLION
HEVO

Data Quality/Data Source Partners

Anomalo
lightup
STARDOG

ML/AI Partners

Labelbox
John Snow LABS
HEX

BI Partners

Power BI
tableau
preset
ThoughtSpot
Qlik

+

25 other partners

Thank you