



Reinvent Government in an Data Intelligence Era

Molly Just-Behr, Director of Industry Solutions

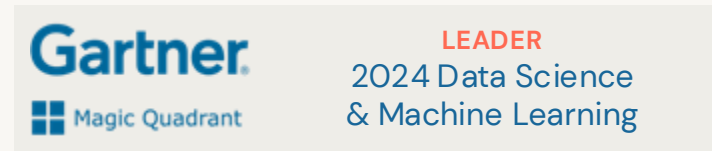
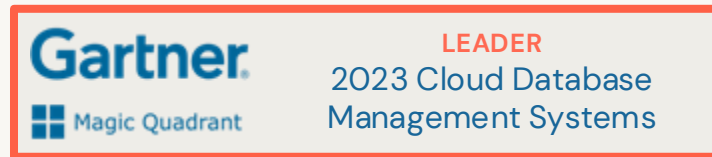
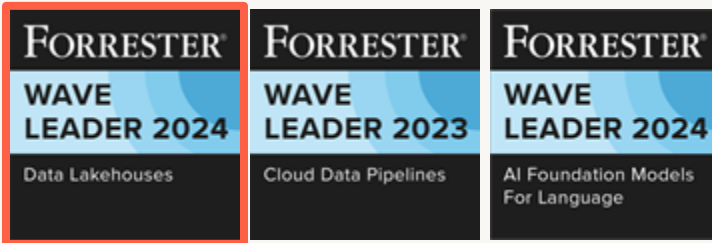
September 18, 2025





databricks

The Data + AI company



Analytic
Stream
Processing



12,000+
global customers



Inventor of the
lakehouse
and pioneer of
generative AI

Creator of:



DELTA LAKE



Databricks Federal

- **Deep compliance experience:** available at FedRAMP Moderate and High, DOD IL5, and higher classification levels
- **History in U.S. Intel:** Databricks is an In-Q-Tel company, first delivered software to the Sponsor in 2017
- **Well-Adopted in PubSec:** Databricks has 100s of customers in U.S. PubSec, including all DoD branches, FBI, DHS, and many of the intelligence agencies

**Data and AI is a top Government
Priority**



Data + AI is a top White House priority



IT Modernization
Act



Move from legacy, on-prem to the **cloud** for performance and scale



Federal Data
Strategy



Federate and catalogue all data and provide easy and reliable access

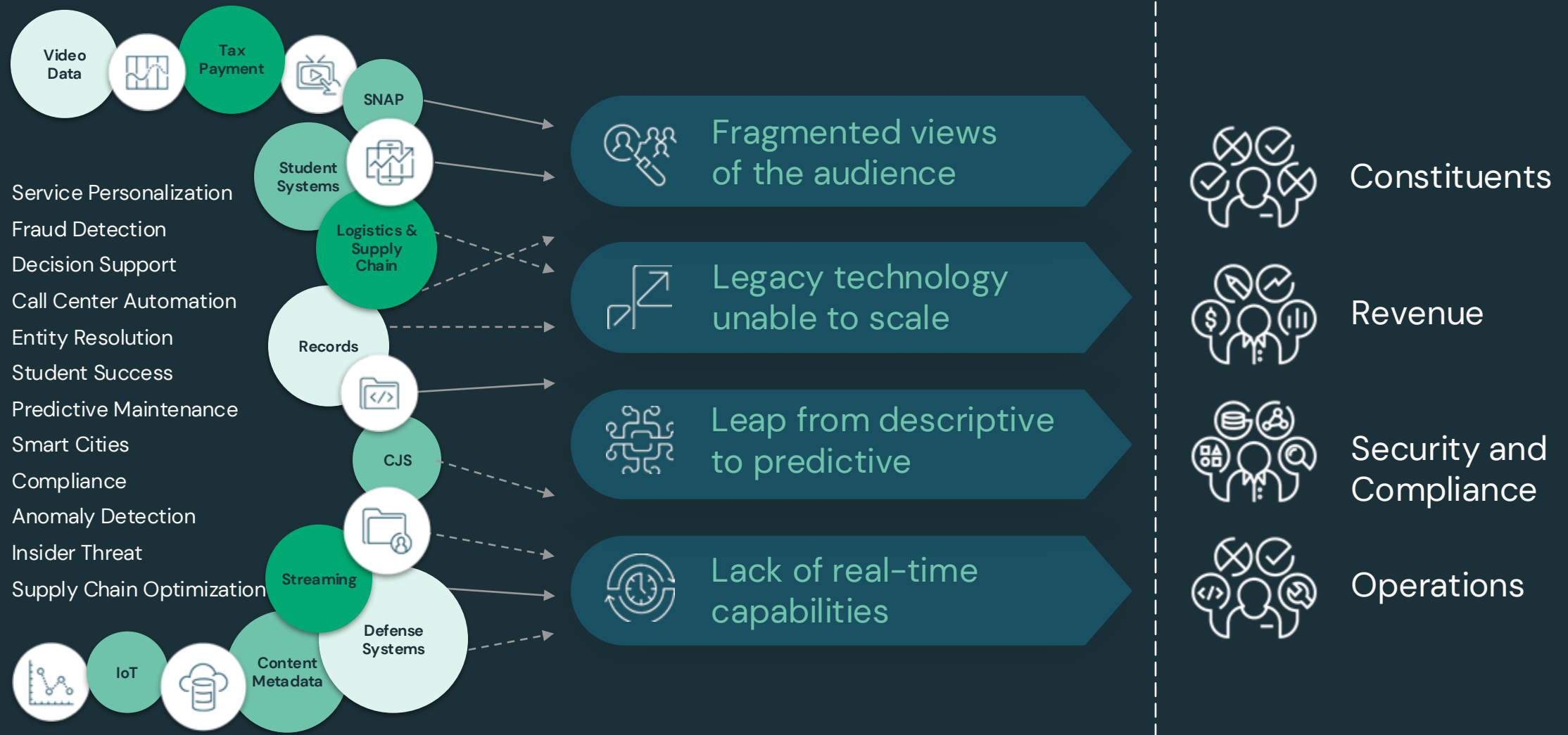


AI Executive Order



Accelerate adoption of **data and AI** to maintain global leadership position

But Government struggles with Data and AI



The data estate is fragmented

Governance



Orchestration
and ETL



Data Lake



Data
Science



Streaming



Data
Warehouse



Generative
AI



BI

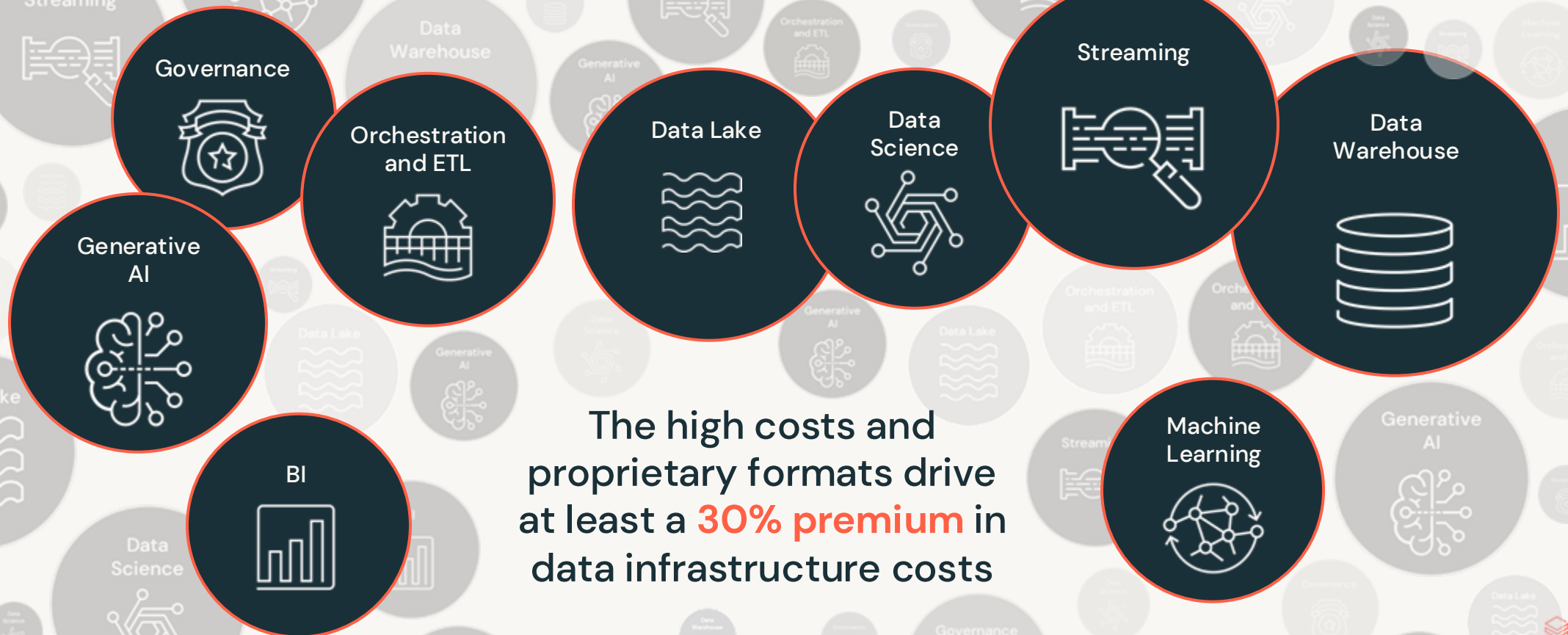


It's a complexity nightmare
of **high costs** and
proprietary formats

Machine
Learning



It's usually worse than it seems



The high costs and proprietary formats drive at least a **30% premium** in data infrastructure costs

Reinvent Data and AI with the Data Intelligence Platform



We pioneered the lakehouse

Combines the best of data warehousing and data lakes

What Is a Lakehouse?



by Ben Lorica, Michael Armbrust, Reynold Xin, Matei Zaharia and Ali Ghodsi

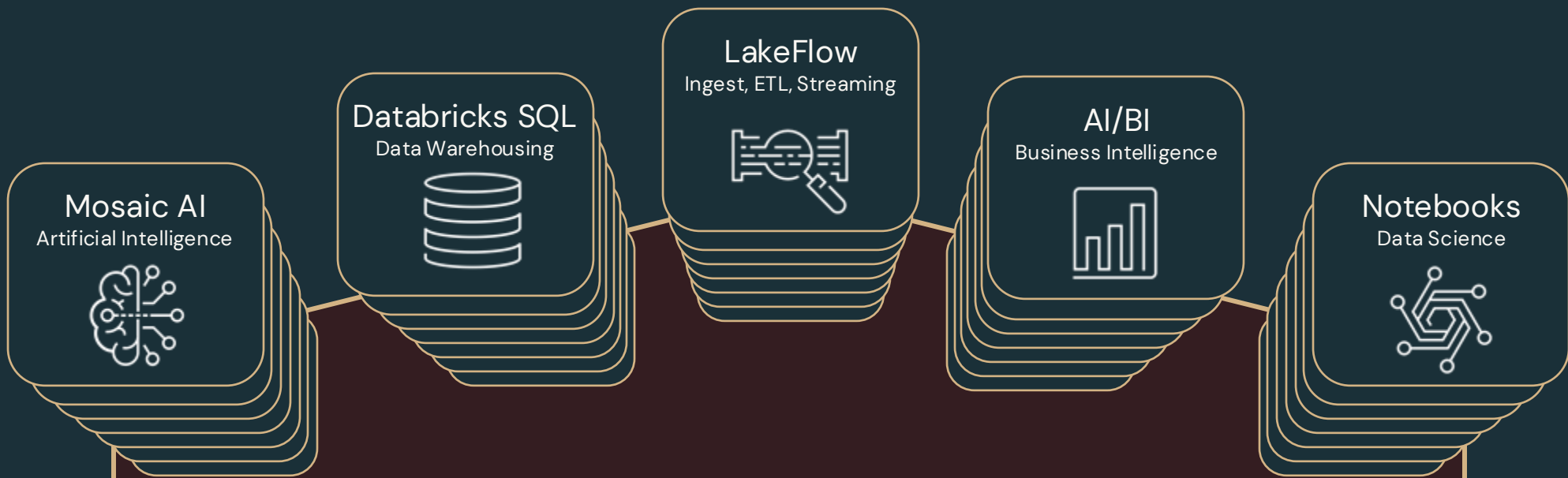
January 30, 2020 in [Engineering Blog](#)

Share this post



Over the past few years at Databricks, we've seen a new data management architecture that emerged independently across many customers and use cases: [the lakehouse](#). In this post we describe this new architecture and its advantages over previous approaches.

Data warehouses have a [long history](#) in decision support and business intelligence applications. Since its inception in the late 1980s, data warehouse technology continued to evolve and MPP architectures led to systems that were able to handle larger data sizes. But while warehouses were great for structured data, a lot of modern enterprises have to deal with unstructured data, semi-structured data, and data with high variety, velocity, and volume. Data warehouses are not suited for many of these use cases, and they are certainly not the most cost efficient.



Lakehouse



ICEBERG



Data Lakehouse

An open, unified foundation
for all your data



Generative AI

Easily scale and use data and AI



Data Intelligence Platform

Democratize data + AI across
your entire organization



Unify and Own the data



Unify and own your data

**Adopt the lakehouse
data formats that are
right for your
organization**

Eliminate unnecessary costs
from multiple copies of data

Break free from lock-in
with open formats you control



Iceberg



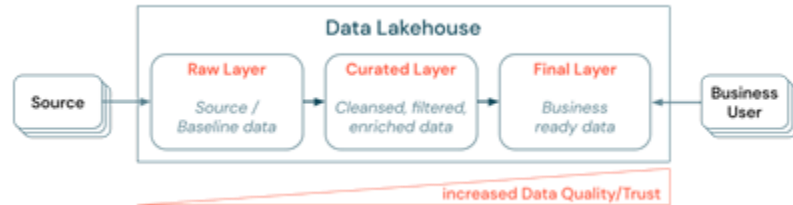
Delta Lake



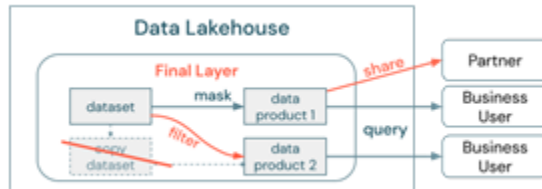
Parquet

Lakehouse Principles for Modernization

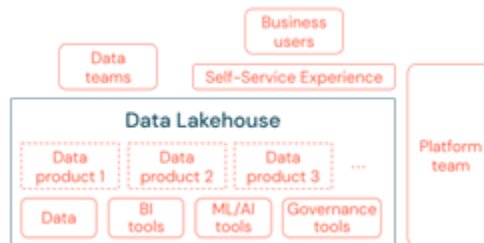
1. Curate Data and Offer Trusted Data-as-Products



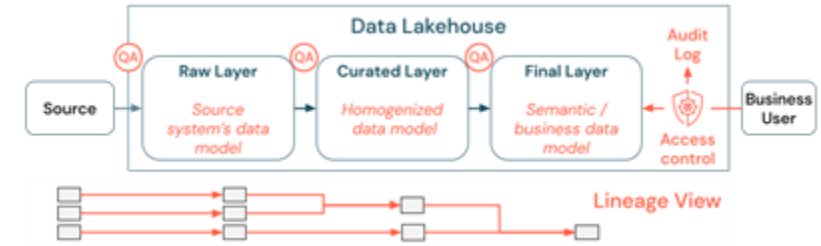
2. Remove Data Silos and Minimize Data Movement



3. Democratize Value Creation through Self-Service Experience



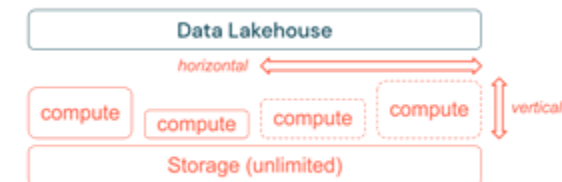
4. Adopt an Organization-wide Data Governance Strategy



5. Encourage the Use of Open Interfaces and Open Formats



6. Build to Scale and Optimize for Performance & Cost

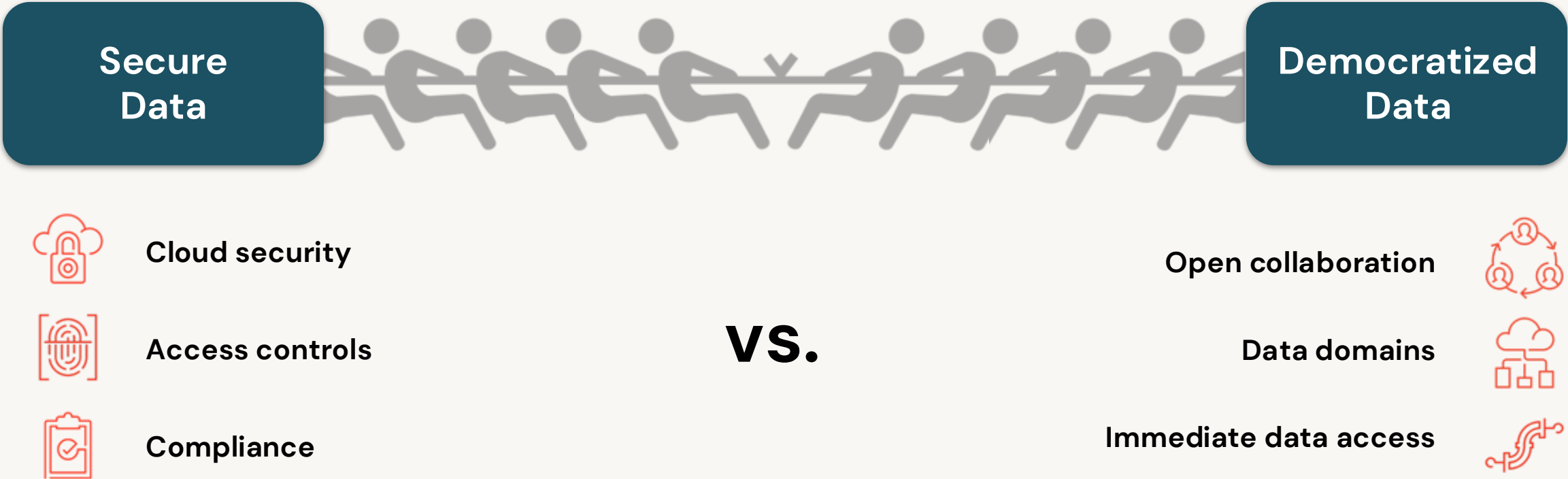


Make Governance a strategic asset



Why Governance? Majority of Public Sector workloads are underpinned by sensitive or highly confidential data

Government Agencies want **two seemingly opposing goals**: 1) secure, well-governed data for highly-sensitive workloads, and 2) improved access to data across Agencies



Data security and data democratization have historically been at odds with each other...

Governance is needed to allow the right access to these sensitive workloads and empower data users across Agencies

Unity Catalog allows Agencies to have **both** data security and democratized data

Secure Data



Cloud security



Access controls



Compliance



Democratized Data

Open collaboration



Data domains



Immediate data access



&

...but Unity Catalog lets you have the best of both worlds!

