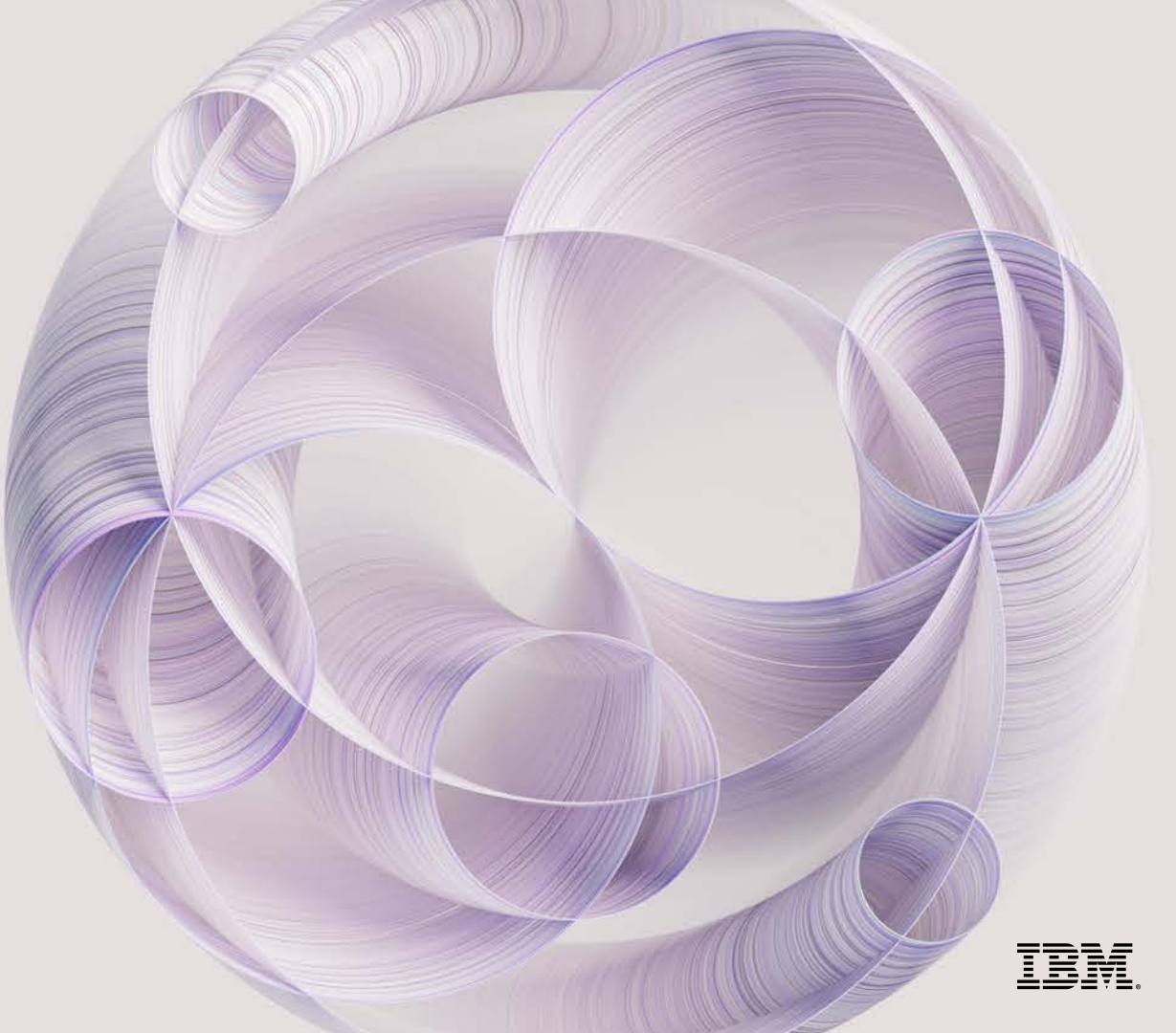
Watsonx.data



DACH - Partner Roadshow

14./15.09.2023

The platform for AI and data

watsonx

Scale and accelerate the impact of AI with trusted data.

watsonx.ai

Train, validate, tune and deploy AI models

A next generation enterprise studio for AI builders to train, validate, tune, and deploy both traditional machine learning and new generative AI capabilities powered by foundation models. It enables you to build AI applications in a fraction of the time with a fraction of the data.

watsonx.data

Scale AI workloads, for all your data, anywhere

Fit-for-purpose data store optimized for governed data and AI workloads, supported by querying, governance and open data formats to access and share data.

watsonx.governance

Enable responsible, transparent and explainable data and AI workflows

End-to-end toolkit encompassing both data and AI governance to enable responsible, transparent, and explainable AI workflows.

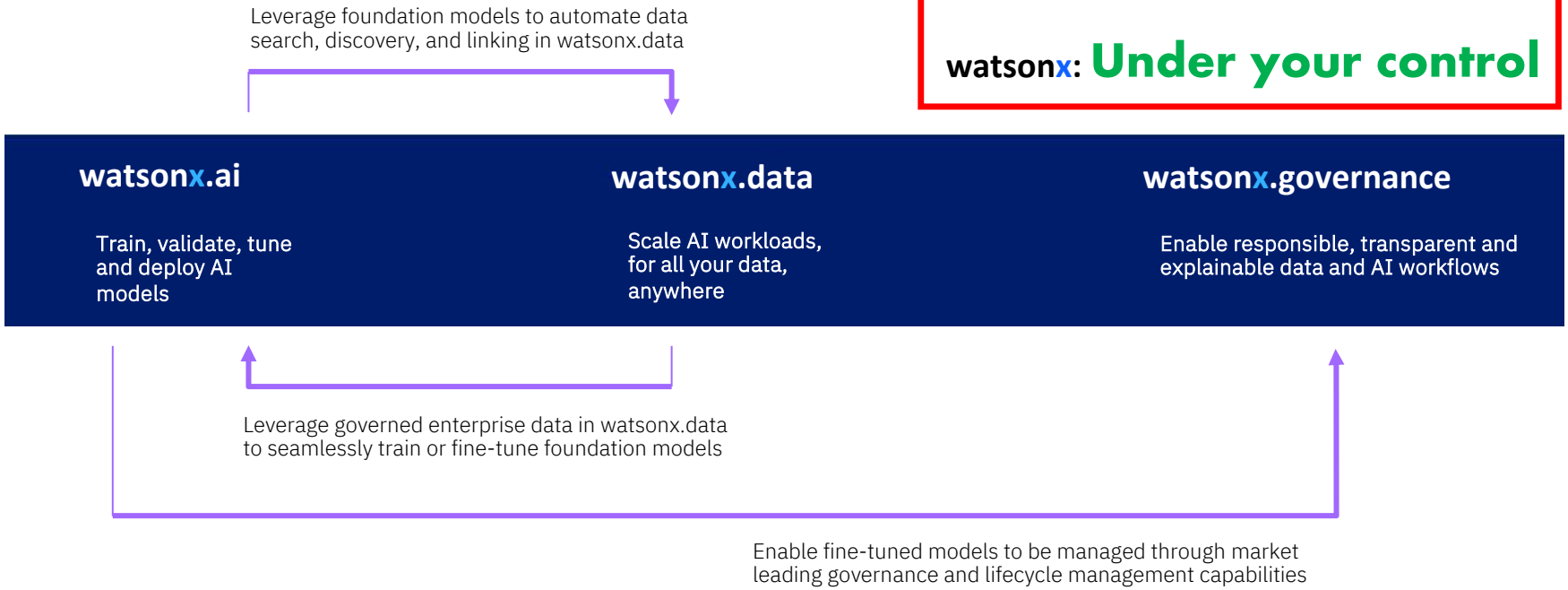
Put AI to work with **watsonx**

Scale and accelerate the impact of AI with trusted data.

watsonx: Your data

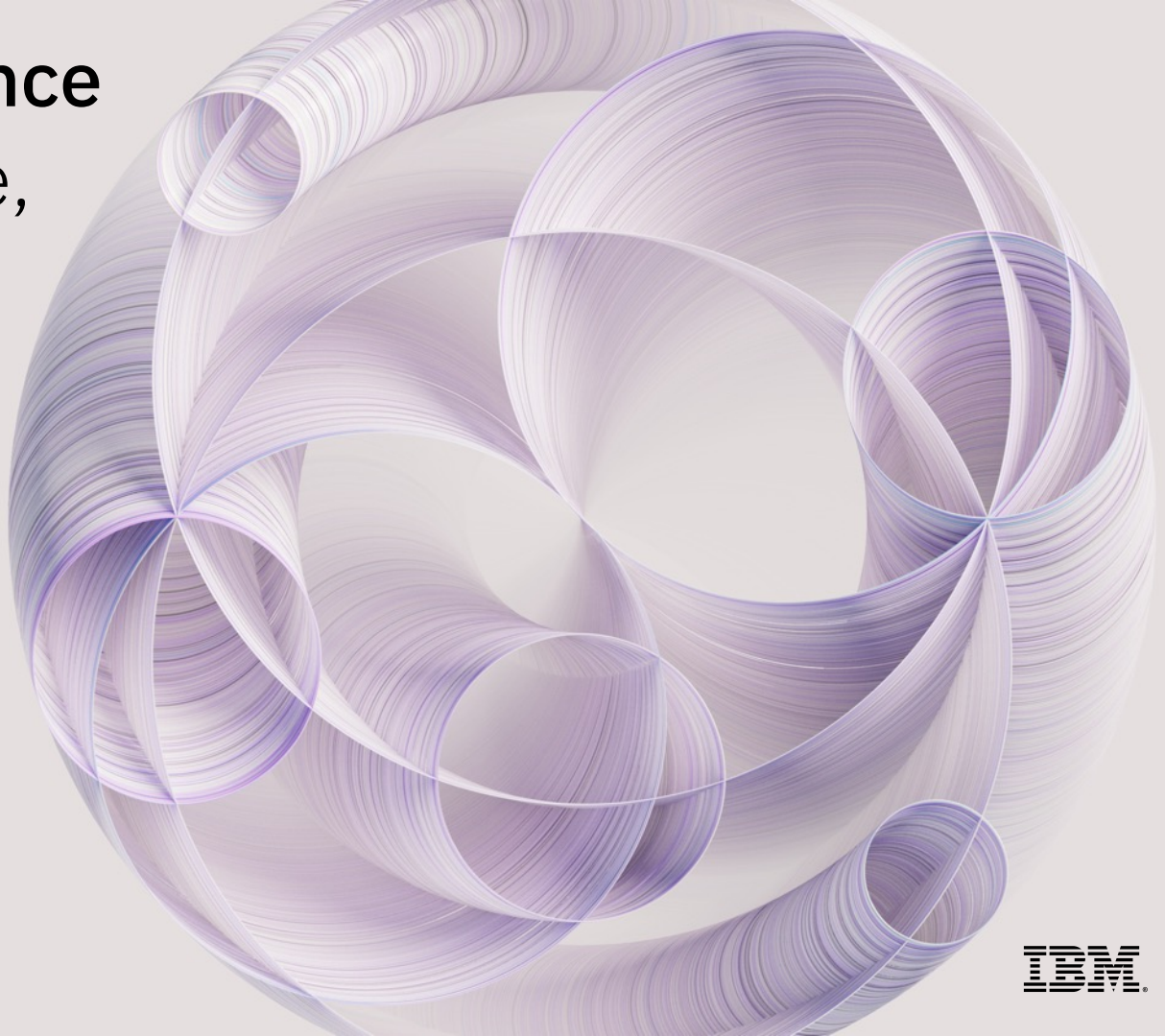
watsonx: Your infrastructure

watsonx: Under your control



Watsonx.governance

Enable responsible, transparent and explainable AI workflows

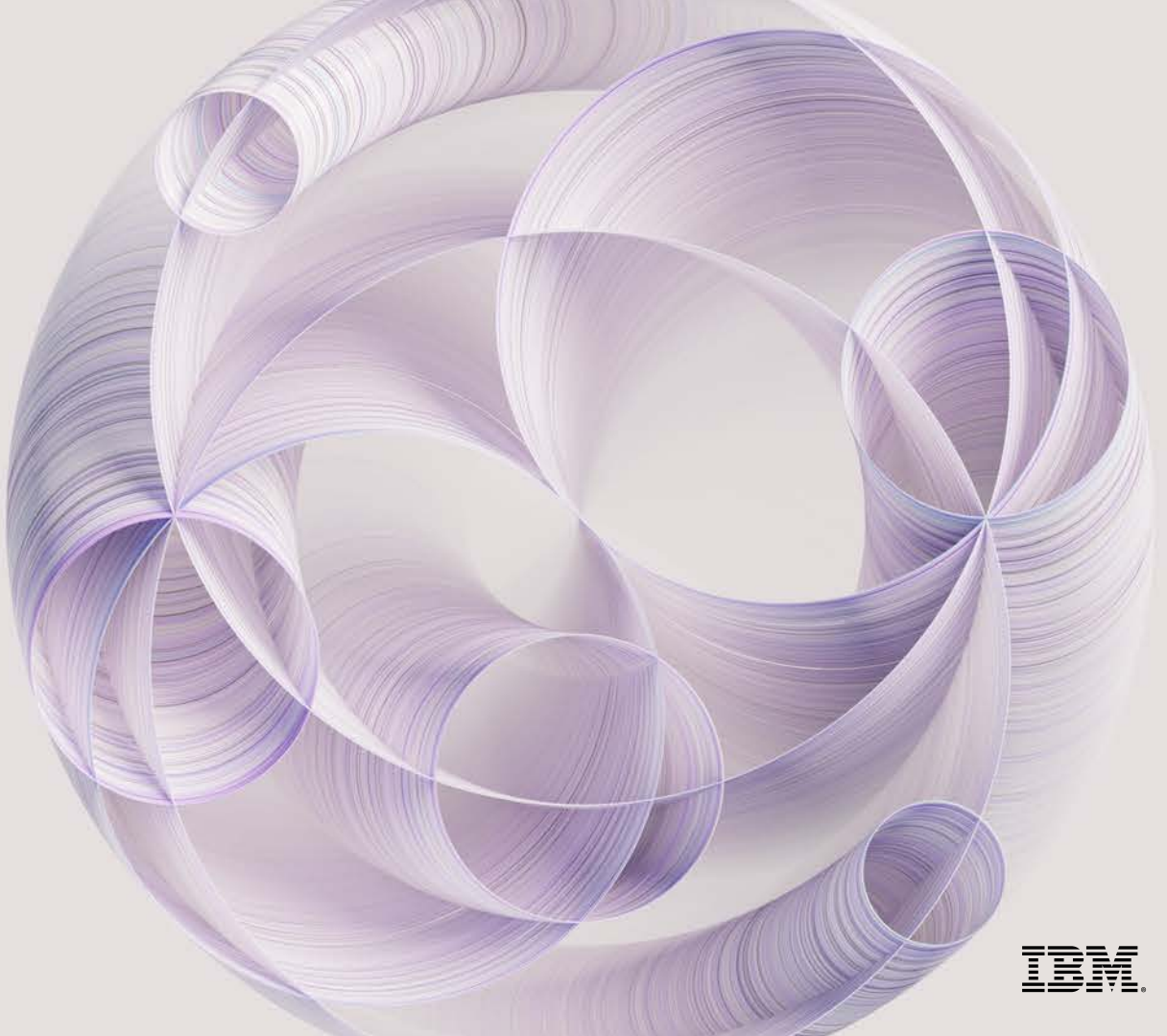


IBM watsonx.governance

end-to-end automated AI lifecycle governance toolkit built to mitigate risk and improve compliance

Watsonx.ai

Train, validate,
tune
and deploy AI
models
with confidence

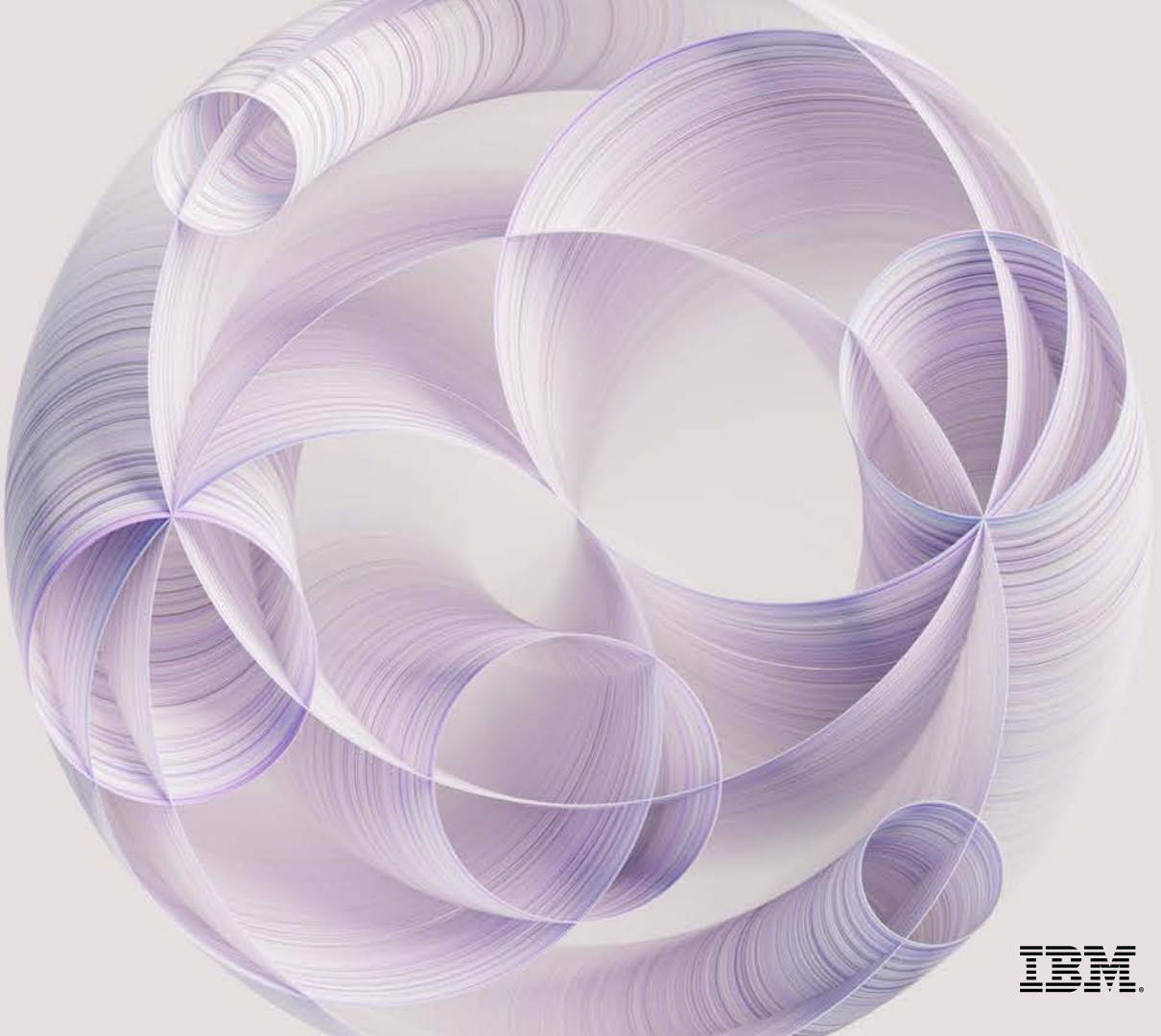


IBM watsonx.ai

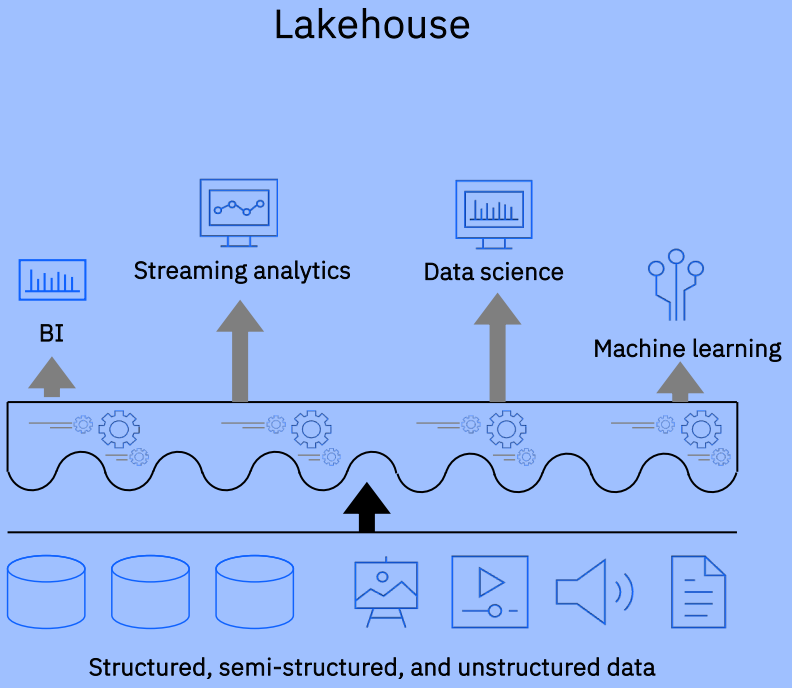
A next generation enterprise studio for AI builders to train, validate, tune and deploy both traditional ML and new generative AI capabilities, powered by foundation models

Watsonx.data

Scale AI workloads, for all your data, anywhere



A lakehouse combines the best qualities of a data lake and a warehouse into a **single integrated platform** to support highly complex data transformations and a **wide variety of use cases**



Why Lakehouse

Current Challenges and Opportunity

Data Warehouse Challenges

- Proprietary data formats
- Vendor lock in
- SQL-centric
- Less flexible
- Elasticity scale limitations
- Expensive
- Loosing next gen applications

Hadoop* Data Lake Challenges

- Poor in place Performance
- ACID/DML lacking
- Failure to address real time requirements
- Rigid pipelines – unable to handle evolution
- Narrow user focus mainly Data Science & ML
- Expensive to expand to generic BI and Analytical use cases
- High Skill to maintain and operate
- Storage and compute scaled together.

Data *Lake*
+
Ware*house*

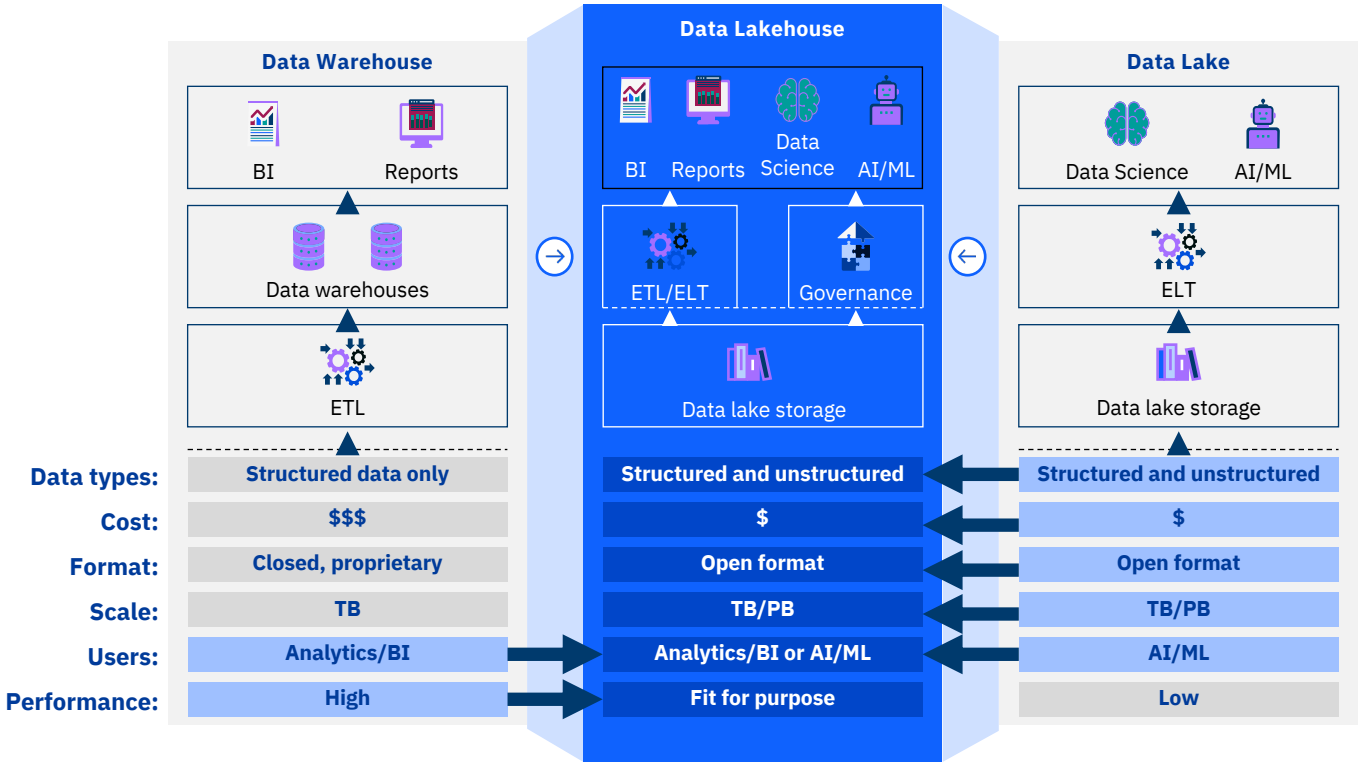
Lakehouse *(Tablestakes)*

- ACID-compliancy - ensure consistency for multiple engines to concurrently read/write data
- Compute Storage separation (and low cost, highly elastic storage anywhere)
- Data Versioning
- Regulatory compliance
- In-built Governance with policy enforcement
- Open Data for applications – shared across vendors and technology

  - Key use cases to support:
    - OLAP
    - ETL
    - ML

* bulk of the market w.r.t Data Gravity

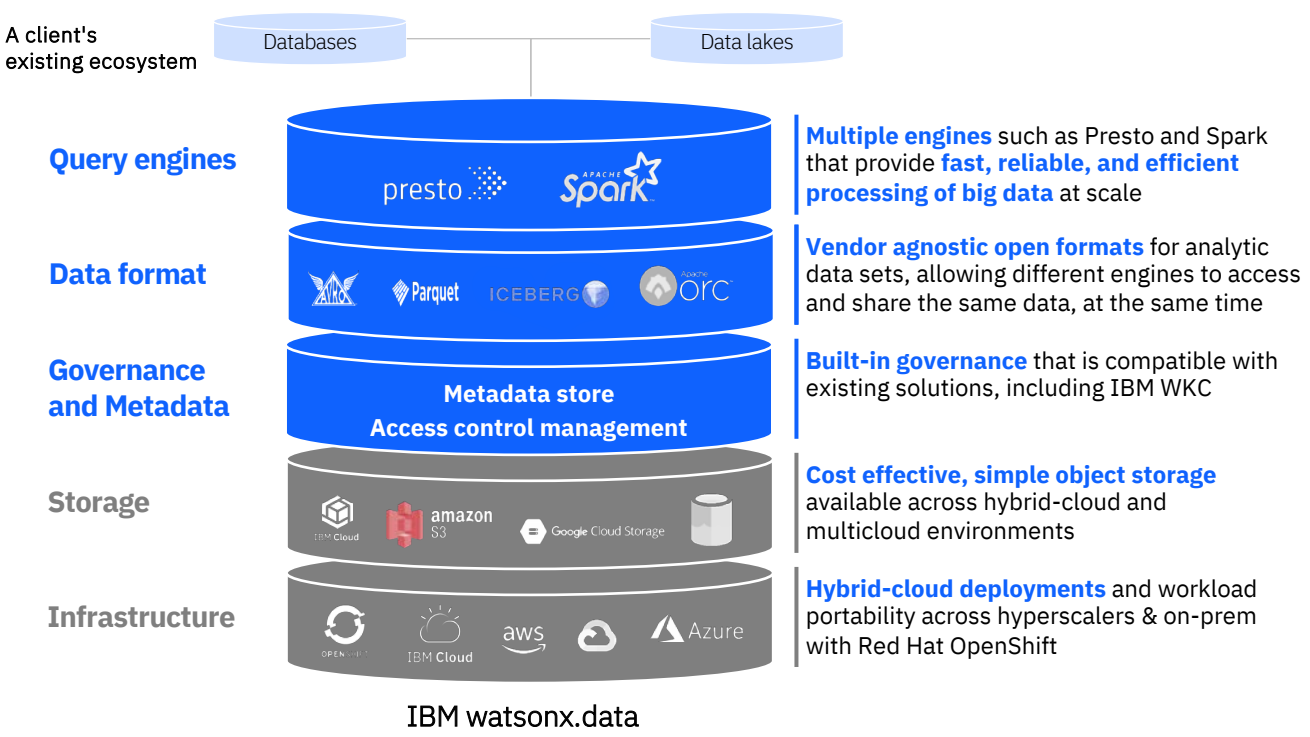
Lakehouses are meant to be a new class of data store that combines the best of data warehouses and data lakes



However, first generation lakehouses still have key constraints that limit their ability to address cost and complexity challenges:

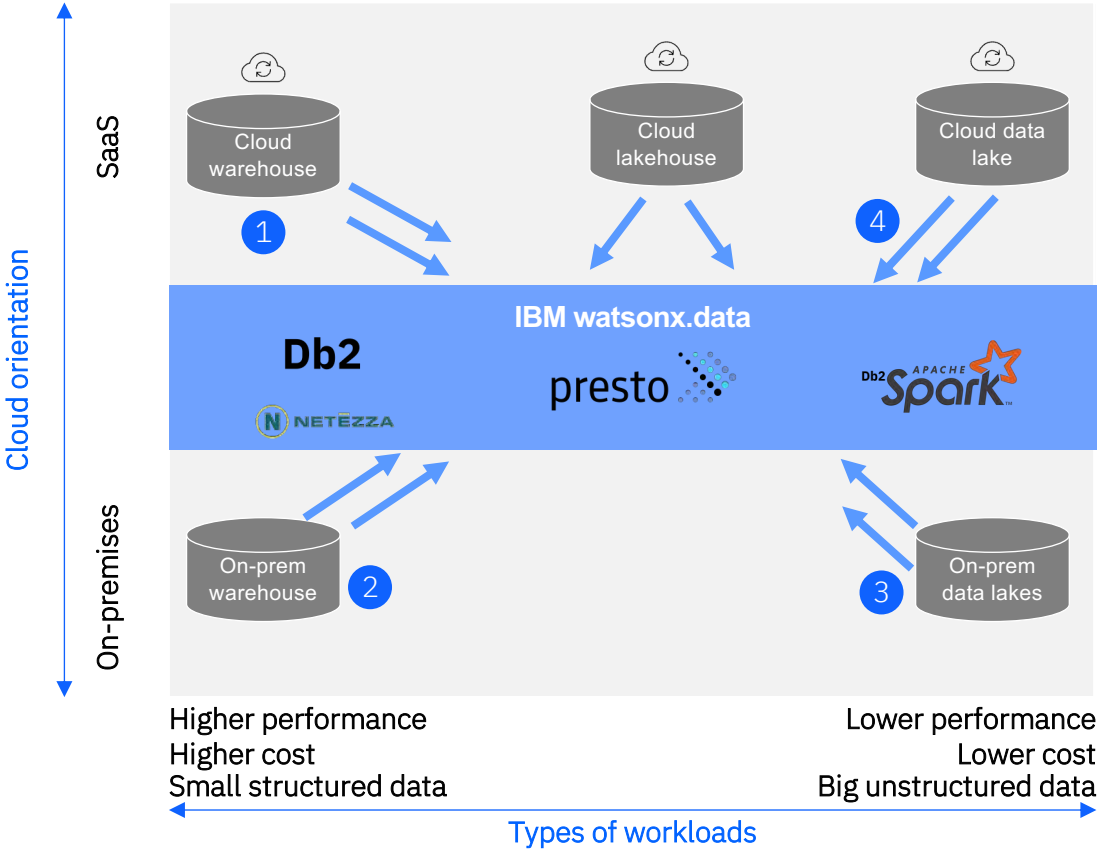
1. Single query engines set up to support limited workloads –typically just BI or ML
2. Typically deployed over cloud only with no support for multi-/hybrid cloud deployments
3. Minimal governance and metadata capabilities to deploy across the entire ecosystem

Overview of the key components of IBM watsonx.data: multiple query engines, open table formats and built-in enterprise governance



IBM watsonx.data

IBM watsonx.data enables access to 100% of a client's data and optimizes workloads across their entire ecosystem



1 **Optimize costly cloud warehouses**

Optimize Snowflake (and similar) workloads to reduce costs by leveraging fit-for-purpose query engines and compute resources (cache vs. compute optimized)

2 **Optimize and access on-prem warehouses**

Optimize on-prem workloads using low-cost object storage and fit-for-purpose engines

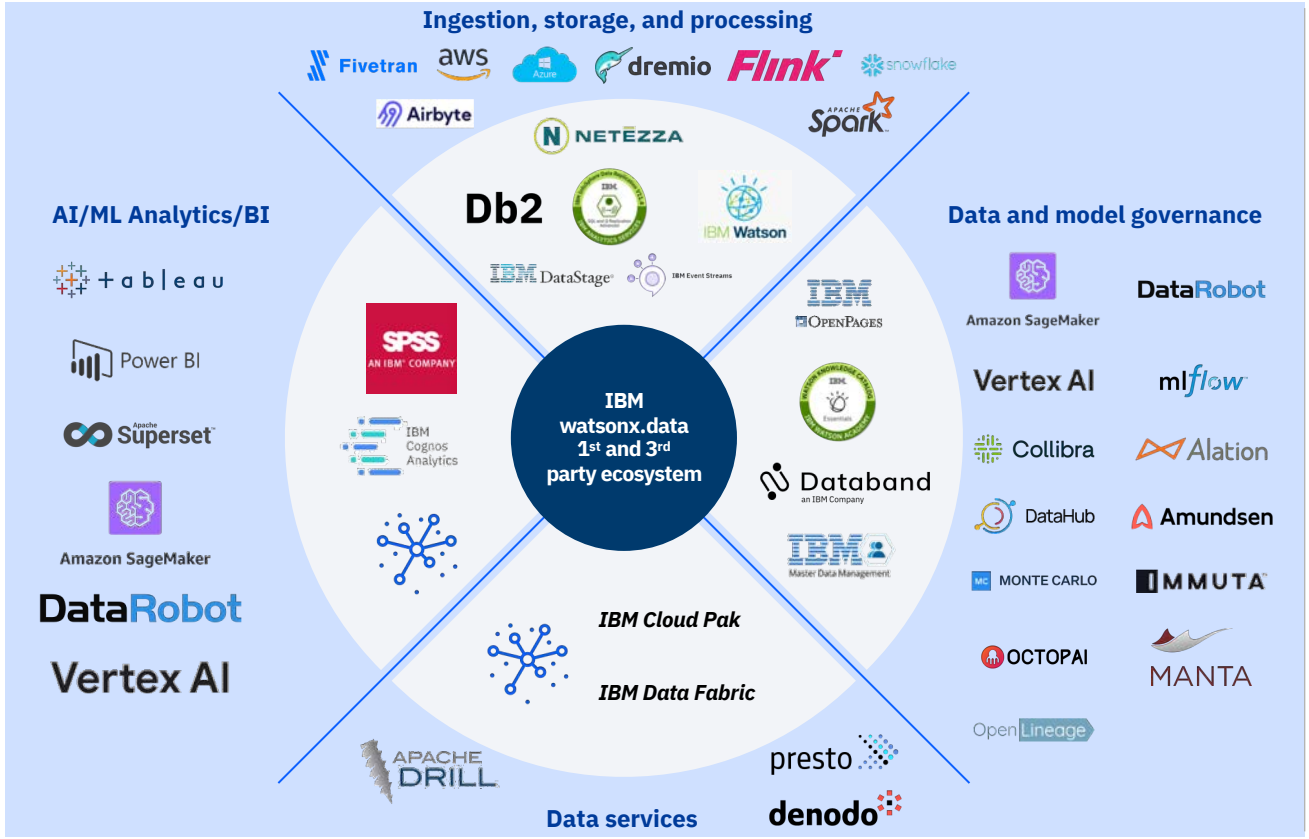
3 4 **Modernize data lakes**

Run existing historical reporting and enable new AI workloads without the cost and complexity of Hadoop

1 2 3 4 **Deploy across hybrid cloud and multicloud**

Seamlessly deploy to both the public cloud and a client's existing on-prem investment, in under 10 minutes

IBM watsonx.data integrates with a robust ecosystem of IBM and 3rd party to enable easy development and deployment of a client's use cases



1st party 3rd party

IBM watsonx.data integrates with both 1st party and 3rd party technologies to support multiple use cases and business needs

Robust 1st party integrations allow users to seamlessly realize the benefits of existing IBM technologies and streamline the flow of data and information between products

Open source and open format provide flexibility and freedom to integrate with popular 3rd party solutions

Artificial Intelligence (AI)

Human intelligence exhibited by machines

AI can be defined as a technique that enables machines to mimic cognitive functions associated with human minds – cognitive functions include all aspects of learning, reasoning, perceiving, and problem solving.

Machine Learning (ML)

Systems that learn from historical data

ML-based systems are trained on historical data to uncover patterns. Users provide inputs to the ML system, which then applies these inputs to the discovered patterns and generates corresponding outputs.

Deep Learning (DL)

ML technique that mimics human brain function

DL is a subset of ML, using multiple layers of neural networks, which are interconnected nodes, which work together to process information. DL is well suited to complex applications, like image and speech recognition.

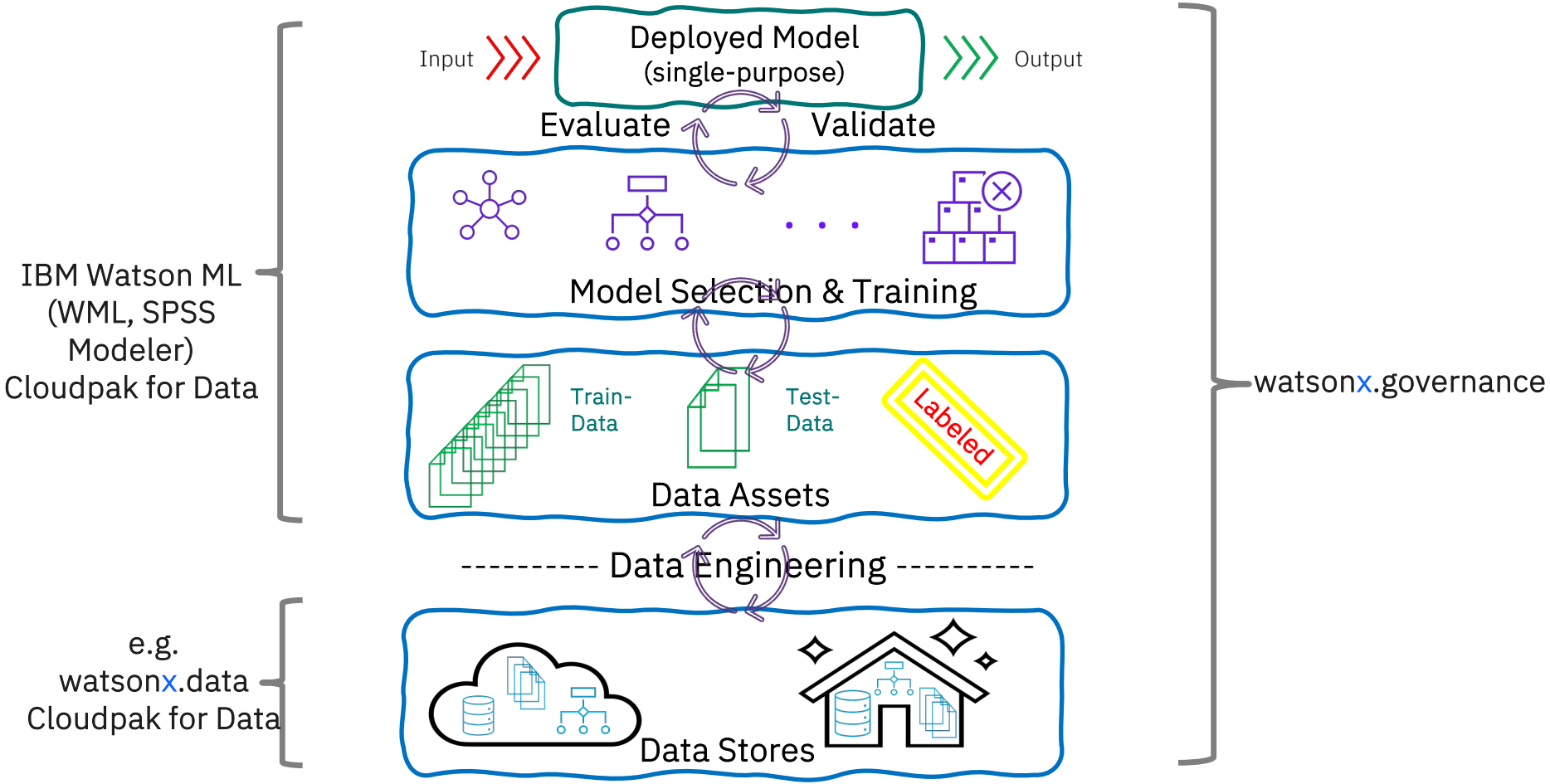
Foundation Model

Generative AI systems

AI model built using a specific kind of neural network architecture, called a transformer, which is designed to generate sequences of related data elements (for example, like a sentence).

1950's 1980's 2010's 2020's

What is watsonx ?



Market dynamics

Major disruptions are driving the growth in the analytics repositories market from on-prem to SaaS and from proprietary to open technologies