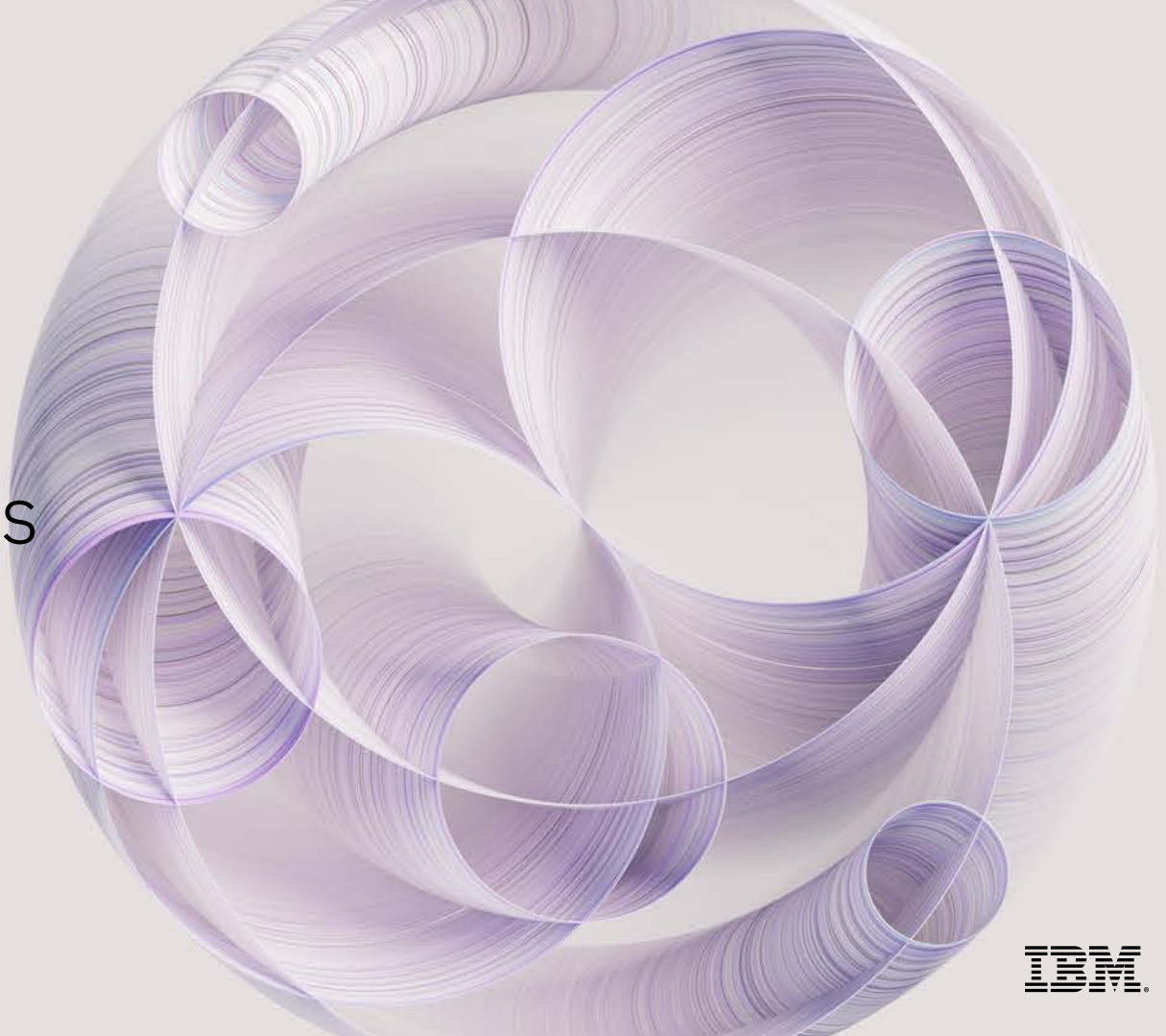
# Watsonx.data

## Day 2
## Adoption Patterns



# Warehouse Offloading for Cost Optimization

## Warehouse offload narrative

- Competitive and IBM (Netezza + Db2)
- Talk track – Client is concerned with the spend on traditional warehouse today – looking to optimize for both performance and cost
- **Value prop:** Cost optimization and openness through the shared meta layer and fit-for-purpose engines
- Example: Snowflake write-intensive workloads moving to Spark and/or Presto, thus educing cost of Snowflake virtual data warehouses

## Key Information to Gather

- Understanding current customer warehouse workload
  - A workload analyzer is also being developed to help understand the customer's current workload
  - What workload can be offloaded to Presto or Spark
  - What associated data need to be offloaded as well?

# Data Store for AI + BI Workloads

## Data Store for AI narrative

- Generative AI and BI have distinctly different data store requirements
- Talk track – Leveraging watsonx.data as the data store for AI and BI
  - Ability to persist files of various type in object store
  - Ability to process ultra wide tables, distributed queries
  - Unlimited snapshotable storage  with Iceberg table format
  - Spark as an ai training engine

- **Value prop:**
  - **Consolidation of data stores. One copy of data for multiple uses**
  - **Ability to handle the different modal of data**

- Focusing on improving this in the future with items such as vector databases

## Key Information to Gather

- Where are they preparing data for AI?
- Where are they training and deploying their models?
- Overlap in the data used for their AI training vs BI needs?

# Governed Data Access with Data Fabric

## Governed Data Access and Sharing with WKC

- Governing disparate data sources is difficult and providing self serviced access while being secure is challenging
- Talk track – Enable your organization to share data freely without concerns over access or governance
- Value prop:
  - Global governance policy that are enforced locally to reduce time and effort sent managing governance policies

## Key Information to Gather

- Are they a current WKC customer or looking to adopt?
- What are their current governance solution today

# Replacing Warehouse or Hadoop

## Why are we not replacing?

- Watsonx.data offers engines that are great at supporting large queries over modern data formats and it will not offer the same SLA as a traditional warehouse
- Customers may have petabytes of data in their data lake and will unlikely move all of it in a short time frame
- Workload migration may be more challenging and difficult than the data migration

## Why may it come up as a watsonx.data use case?

- Other data lakehouse vendors such as Dremio and Databricks will message how they can replace a data lake and a warehouse with their solution
- Customer have a desire to consolidate and reduce cost

## Future

- The team is actively working on the Velox engine, which will bring warehouse “like” performance to the presto engine
- Customers may slowly shift their data from their existing data lake to watsonx.data once we prove out data access to their lake and the value our integrations brings