

# **Azure Modern Data Architecture**

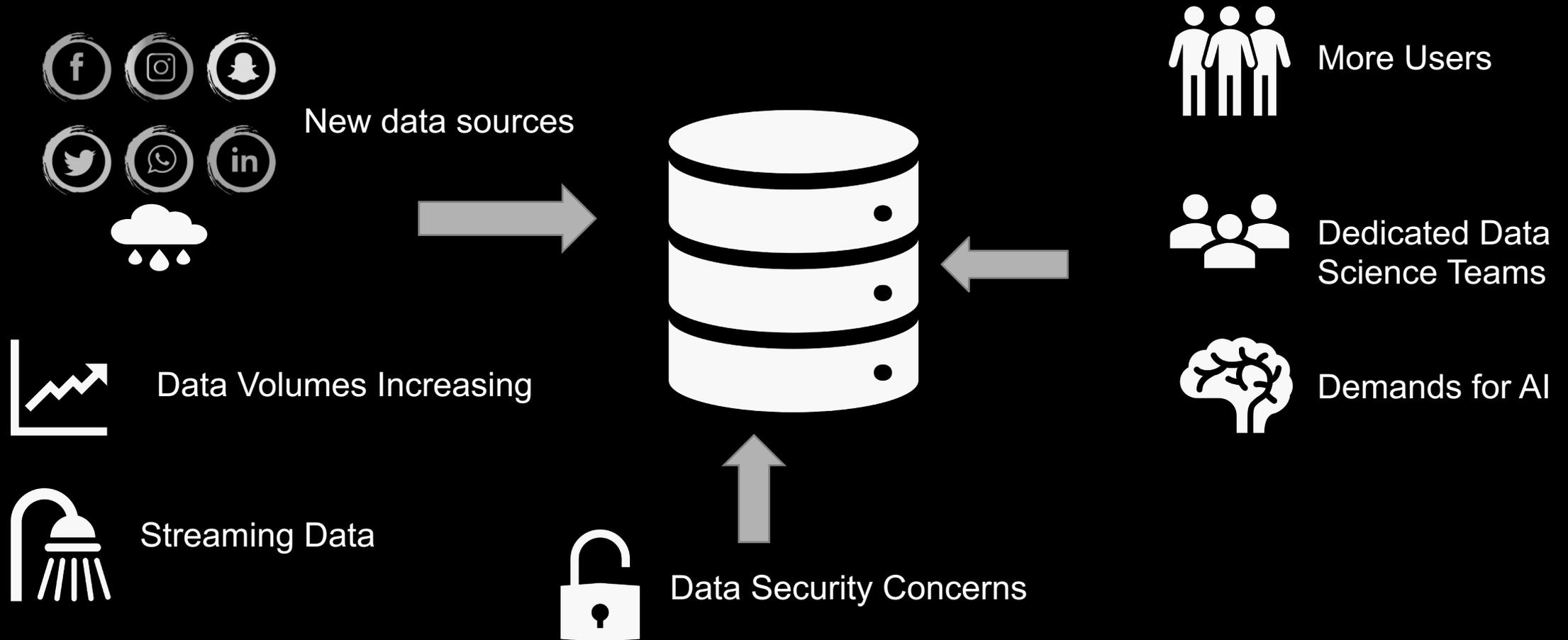
# Introduction to Data Thirst

- We believe that data-driven intelligence has the power to accelerate performance.
- Data Thirst helps companies and institutions to set up their data for effective use, to unearth hidden information, and to embed insight into better, faster and simpler decision workflows.

# Current Warehousing State of the World

- Relational Data Warehouse
  - Processes once a day
  - Frequently overruns
- ETL with using GUI tools
  - Testing is difficult
  - Not reusable
- Scale up only
  - No scale out options
- Change is complex

# Data Warehouse Pressure



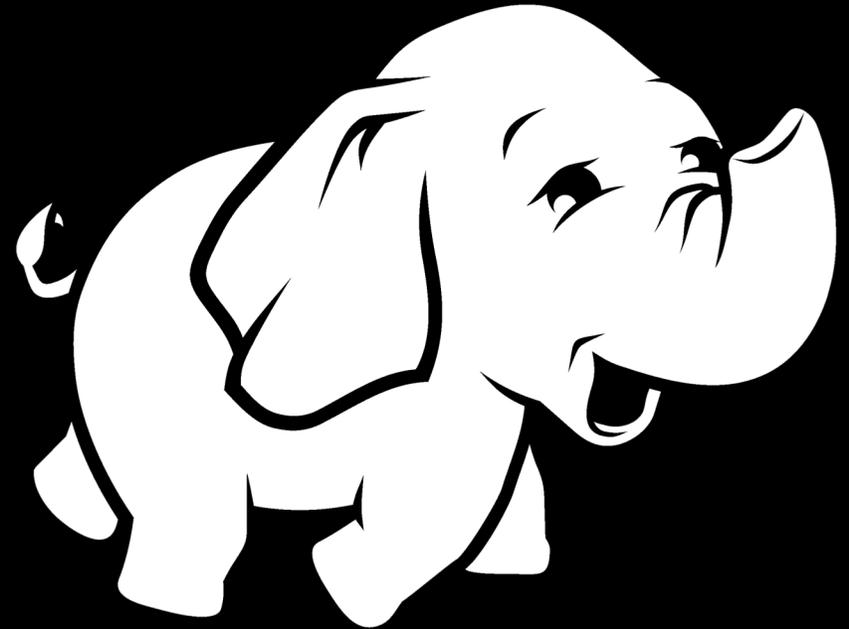
**🚰 Data Thirst**

# Net Effect

- Decisions are not data driven
- Unhappy end users
- Lack of insights
- Shadow IT
- Data Leaks

# Hadoop & Big Data

- Complex
- Over Hyped
  - Edge cases were awesome
  - Not for the masses
- Expensive Clusters
- Drove innovation
- Clear technology winners:
  - Spark
  - Hive
  - HDFS



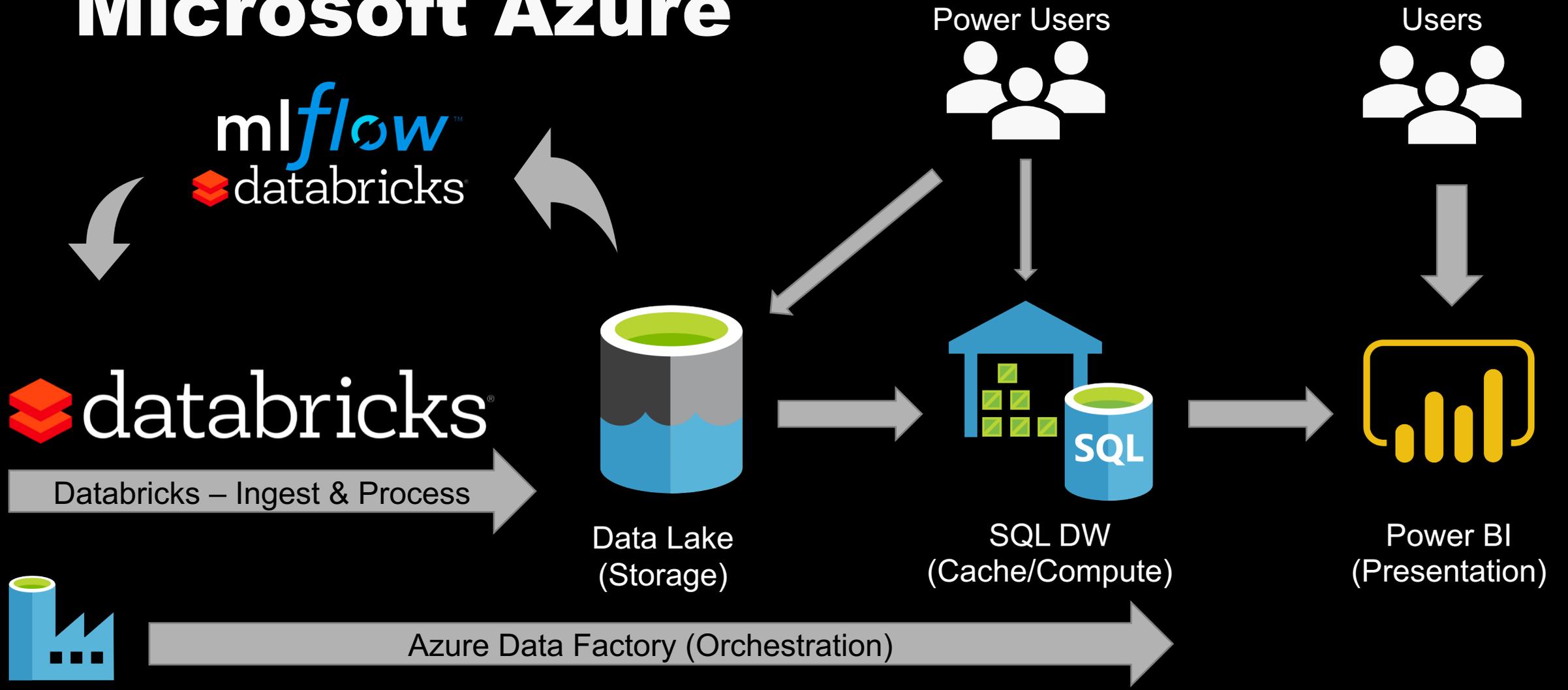
# Goals

- Treat data as an Asset
  - Owned
  - Secured
- Throw away nothing
- Simple Processes
- Modular Systems
- Immutable & Idempotent Processes
- Support for tomorrow

# Modern Data Architecture

- Cloud based Platform
- Modular Components
- Managed Platform as a Service
- Ease of New data sources/flows
- Massive Scale-out Processing

# Microsoft Azure



**Data Thirst**

# Why is it better?

- Separate Storage & Compute
- Scale out
- Modular
- Best of Big Data and Traditional Warehouses
- Support for all types of data
- Develop using in SQL, Scala, Python or Java

# Data Lake Governance

## Sourced

Raw data from source.  
Immutable.  
Inconsistent and unfriendly to use.

## Cleansed

Common format.  
Deduplicated.  
Validated.  
Removed headers/footers

## Integrated

Conform different data sources.  
Handle Slowly Changing Dimensions.  
Generate Keys & Lineage.

## Presented

Curated data.  
Star Schemas.  
User friendly.  
BI Tool friendly.

Source Systems



 Data Thirst

# Why Data Thirst?

- Delivered Modern Data Warehouse's in various industries
- We bring experience
- Developed Patterns and Practices to get you started quickly
- Developers certified by Microsoft & Databricks
- Microsoft Gold Partner

# Modern Data Warehouse by Data Thirst

- Extract - Everyone ingests data the same way
- Transform:
  - Custom to every customer
  - Similar Patterns
  - Business Logic/Rules
- Load - Everyone writes data in the same way
- Data Thirst's Modern Data Warehouse framework provides:
  - Extract & Load Solutions
  - Reusable Patterns & Templates for Transformation

# Questions

- Contact: [hello@datathirst.net](mailto:hello@datathirst.net)