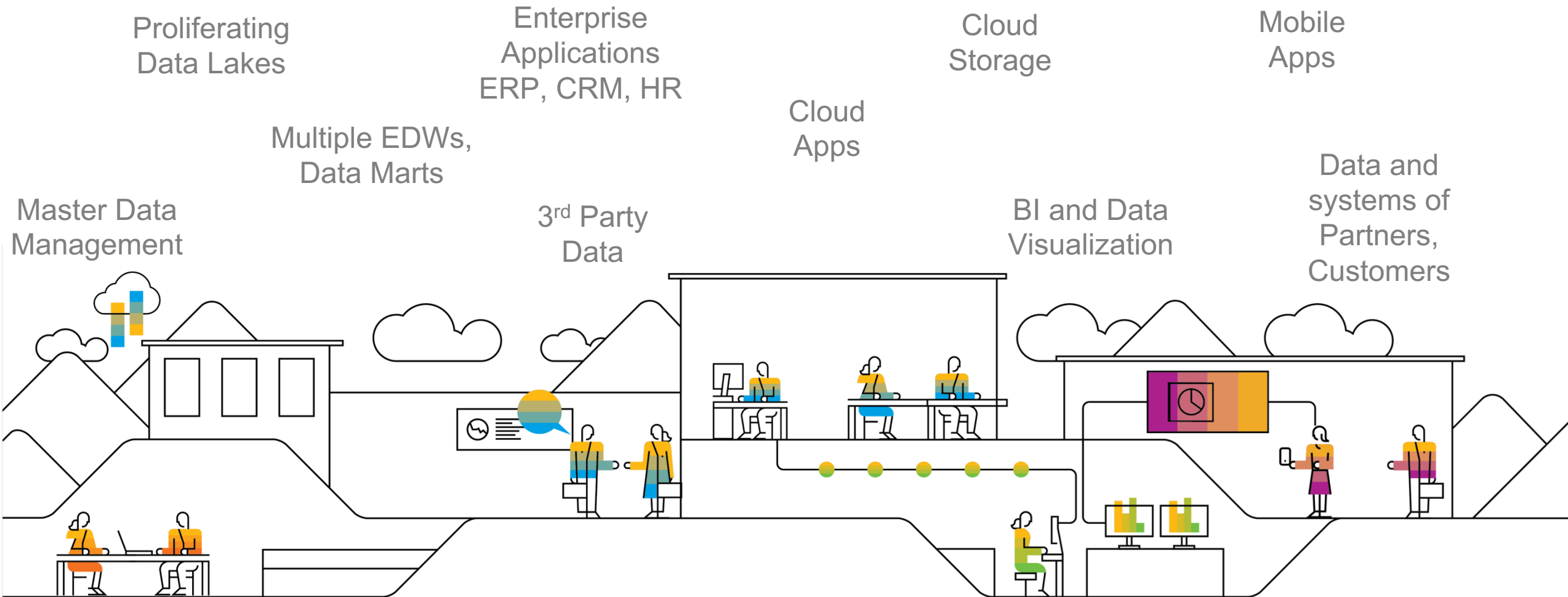


# SAP Data Hub

**Freedom of Data in a Diverse Landscape**



# Enterprise data landscapes are growing increasingly complex



# Modern Landscapes

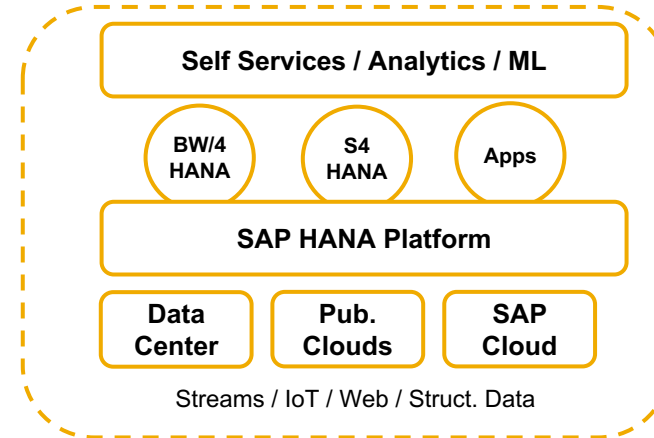
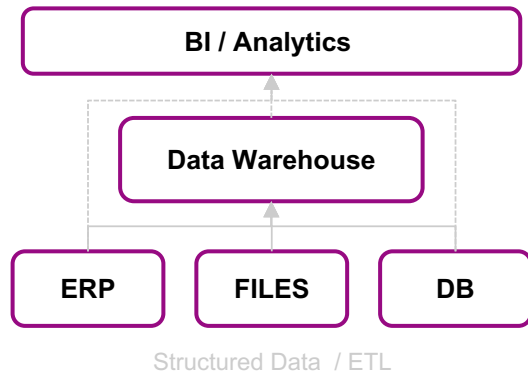
„Big“ Data is transforming customer landscapes

From centralized, relational, on premise DWH approaches...

...to modern distributed (cloud) Data Platforms

## Characteristics

App. Server & DB  
OS & Hardware  
ETL Driven  
Structured Data



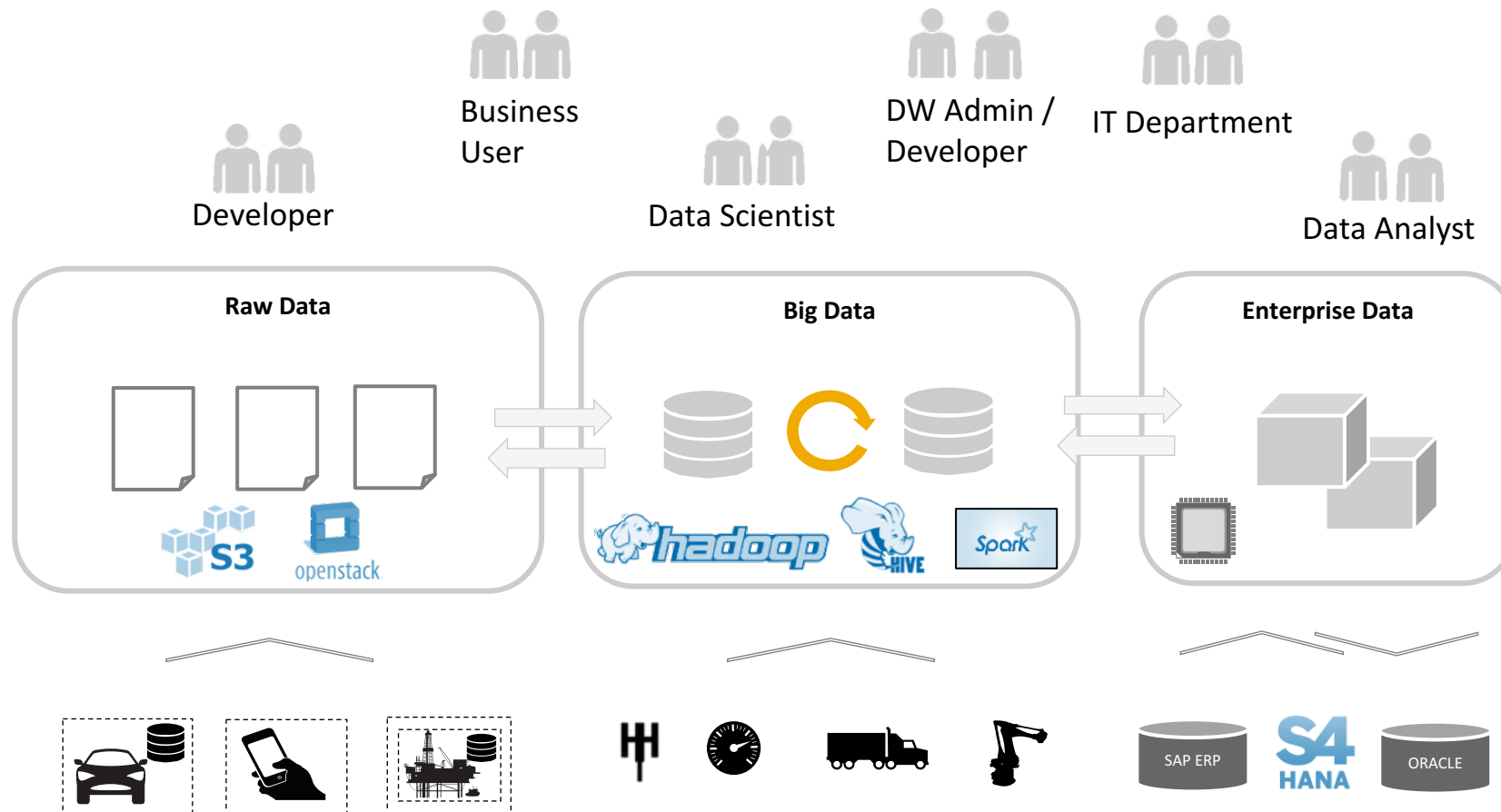
## Characteristics

Serverless Computing  
Containerized Software  
Distributed Data  
Data Driven  
Any Format

## Key Drivers:

- Challenges for traditional architectures due to multi-structures, large data volumes, landscape scale outs
- Growing Cloud / Data Lake / IoT Adoption

# Modern Landscapes –Example



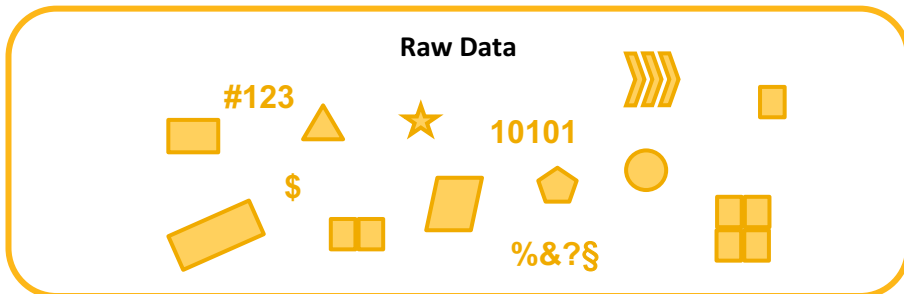
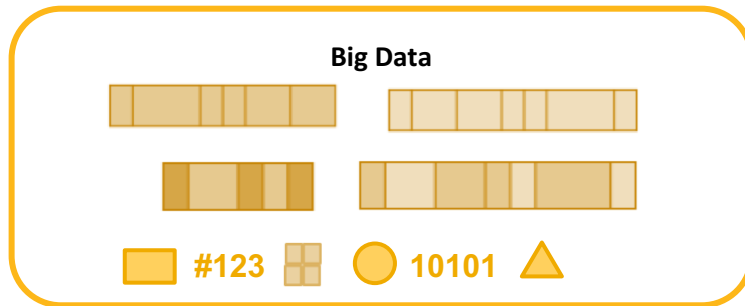
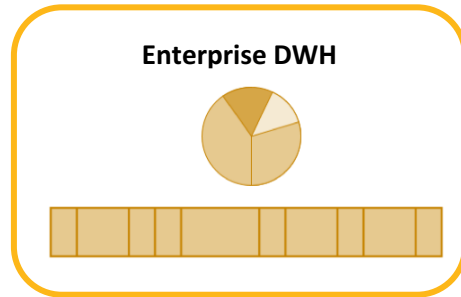
## Key Challenges & Questions

- Which technology for which data?
- How can collaboration between different groups with individual requirements be facilitated?
- How can huge and ever-growing data volumes be handled?
- Rapidly changing landscape – where is the next opportunity?

Another set of silos  
are not the answer!!!

# Modern Landscapes

## Data Perspective as biggest challenge



Analytical  
Modeling

Joins / Unions  
De-normalization

Matching /  
Duplicate  
Check

Meta Data Extraction  
and Generation

Enrichment

Cleansing

(Re-) Formatting

Anonymization / Masking

Parsing

Filtering

Search

Data  
Validation

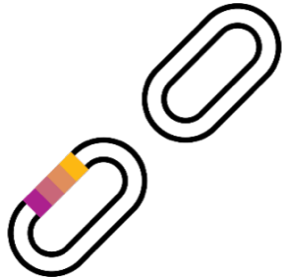
Data Stream

## Refining Insights out of diverse Data

- Structuring the unstructured.
- Handling large volumes of data efficiently
- ETL or DWH are not the answer
  - Data formats & granularity
  - Data Streams and flexible structures
  - Deploy logic at the data, not data at the logic
- Need to ensure integrated analytics across the enterprise.
- Early insights on each level of data needed
- Automation across layers needed

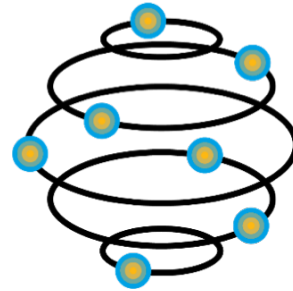
# Challenges

Overcoming silos, complexity to drive better operations and insight



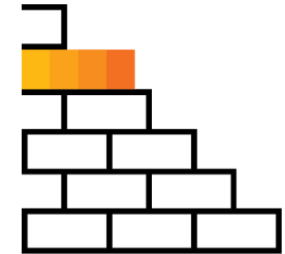
## Missing Link

between Big Data and  
Enterprise Data



## Lack of Enterprise Readiness

of Big Data solutions

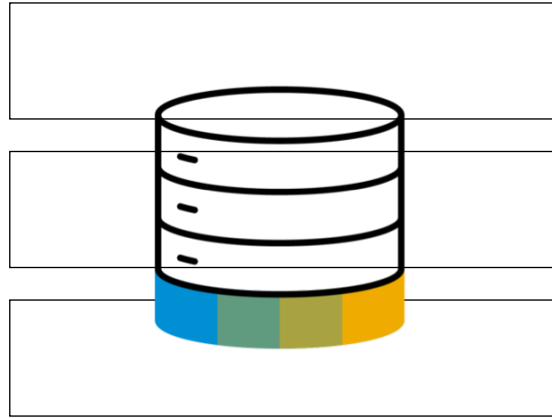


## Limited Tools = High Effort

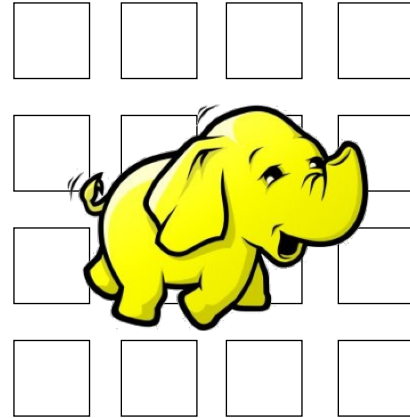
to productize complex data  
scenarios across data landscape

# New challenges require new technologies

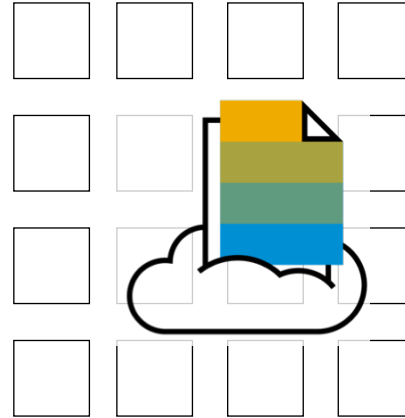
Distributed systems in a distributed landscape



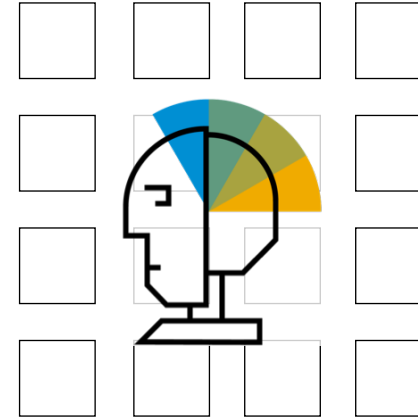
Existing Systems



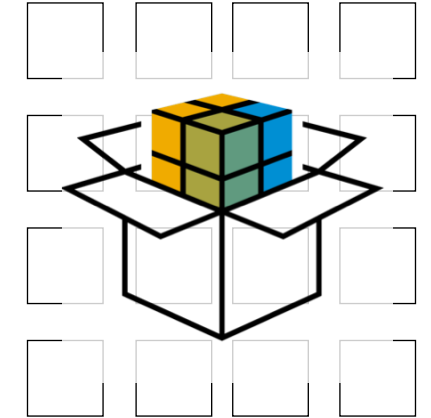
Hadoop/Spark



Cloud Storage  
(i.e. AWS S3)



Machine Learning  
(Python, Spark, Tensorflow)



Containers  
(Kubernetes, Docker)

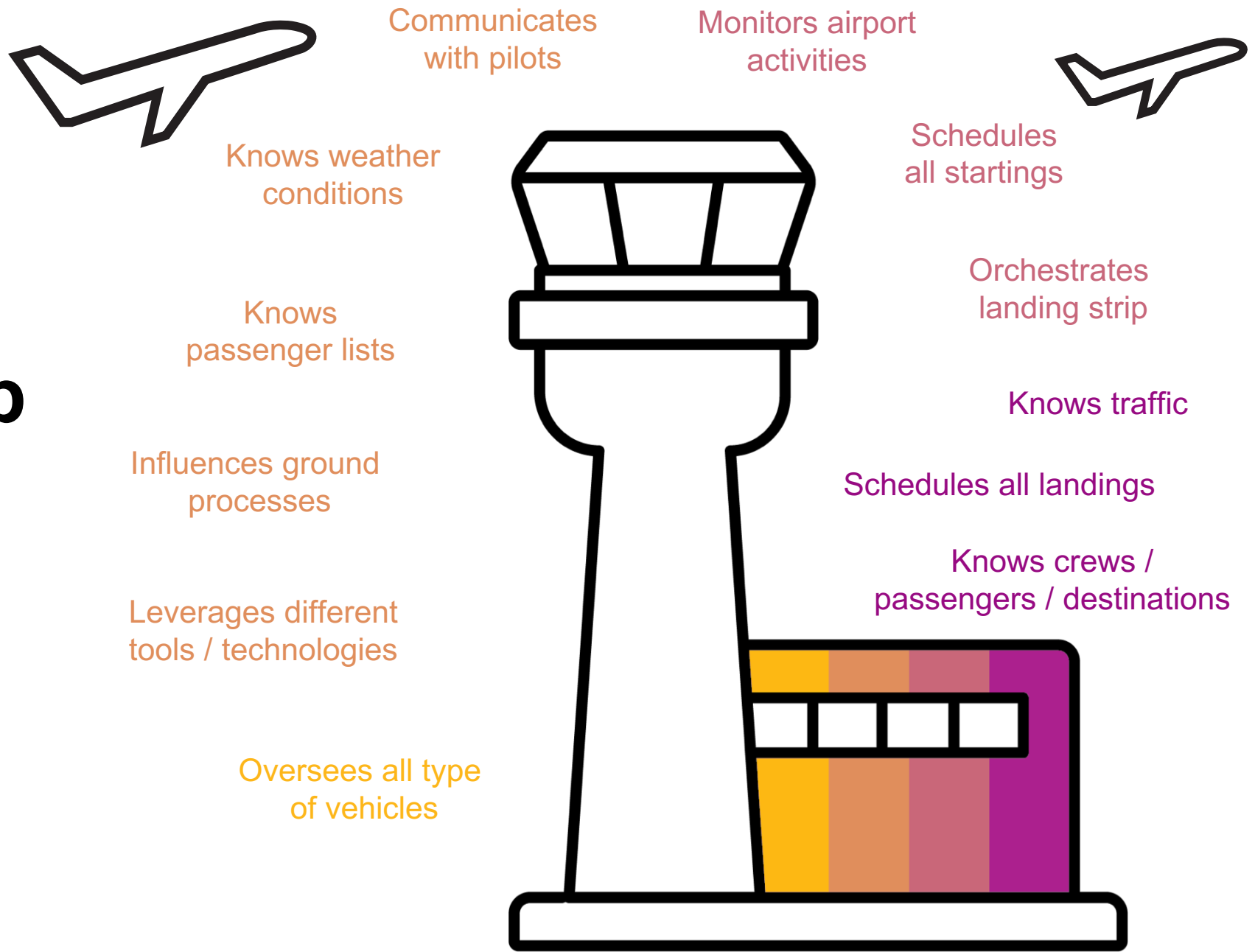
# SAP Data Hub

## Benefits

- 1** Accelerate and scale your data projects
- 2** Gain fast, relevant business insights
- 3** Build agile, data-driven applications
- 4** Enterprise visibility and governance



# SAP Data Hub

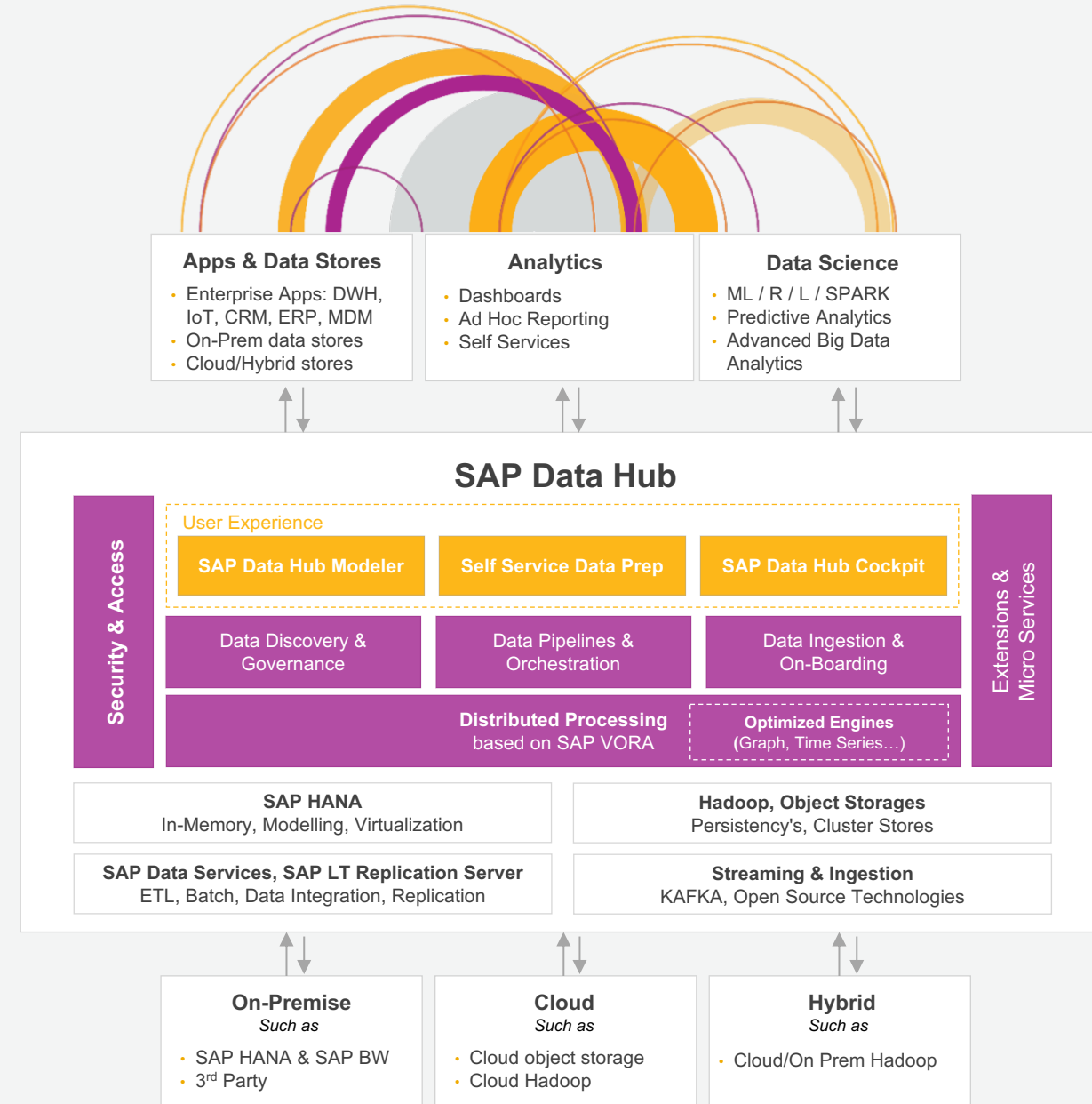


# SAP Data Hub

## Overview

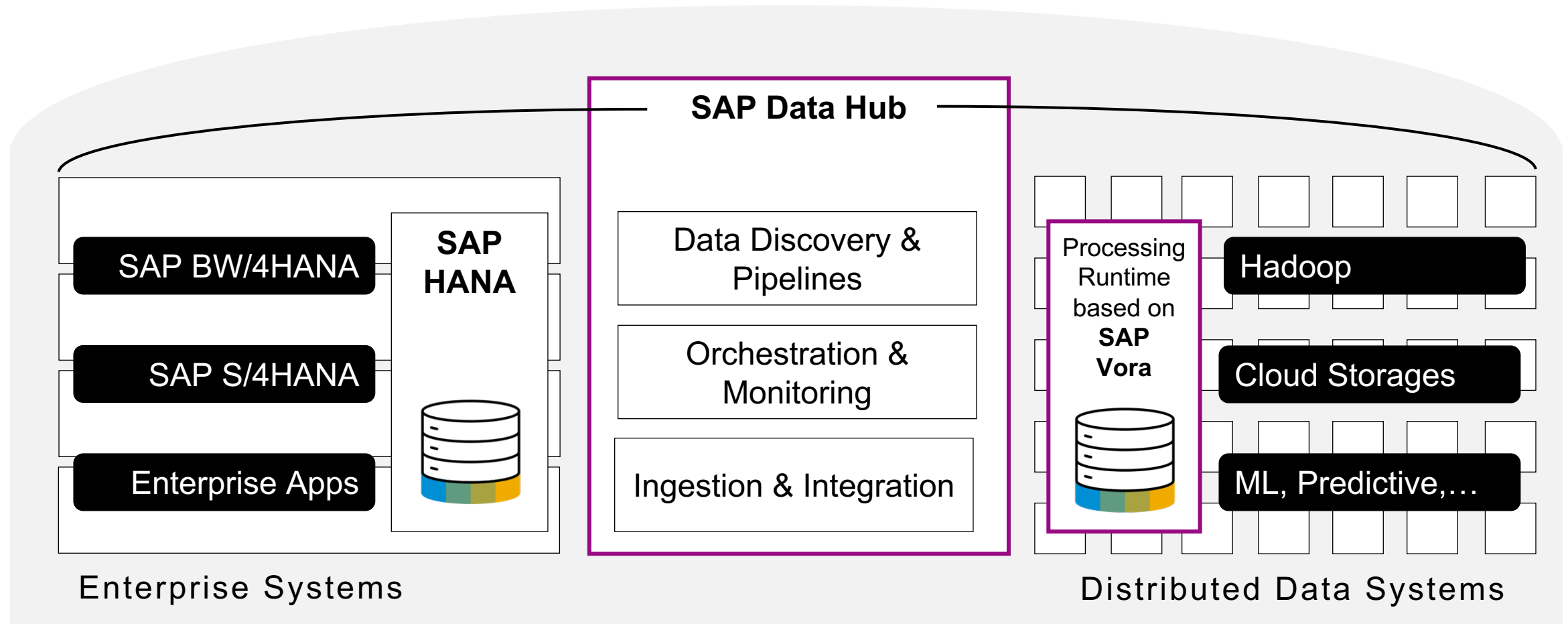
### Define data driven processes across complex enterprise landscapes

- Access on-premises, cloud, or hybrid data sources – SAP or non-SAP (Amazon, Hadoop)
- Leverage robust enterprise integration capabilities
- Connect easily to SAP data management and application solutions as data sources
- Connect to SAP and non-SAP applications and analytic solutions as endpoints



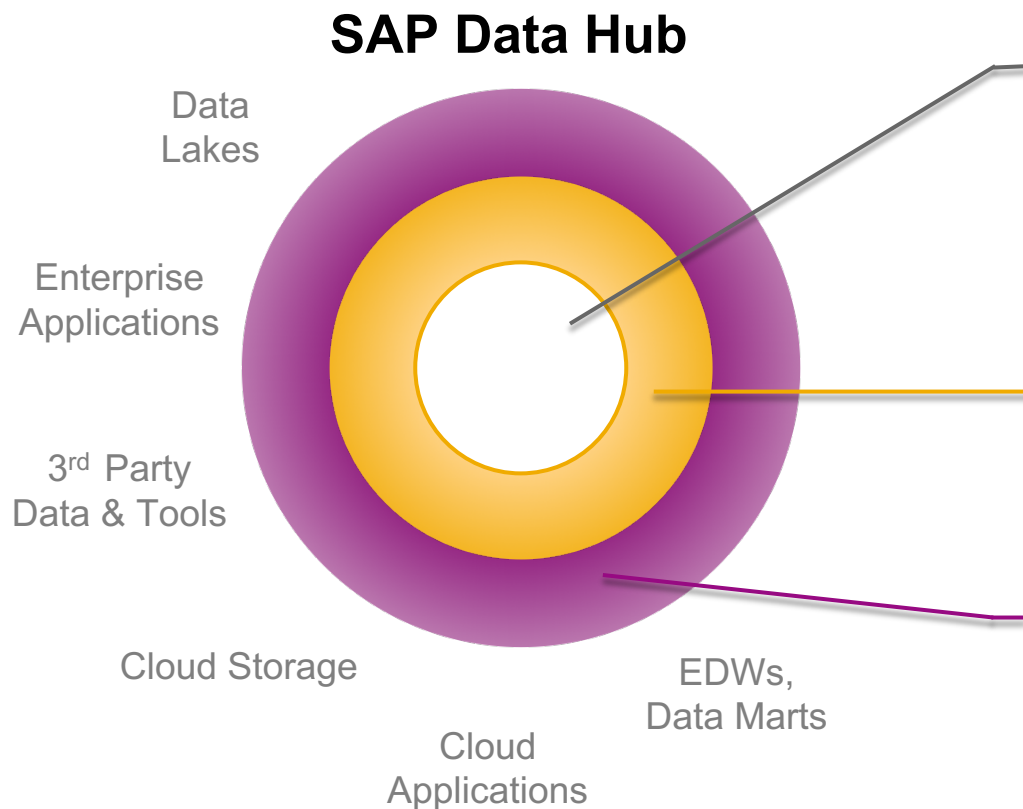
# SAP Data Hub

## Unifying Data Silos



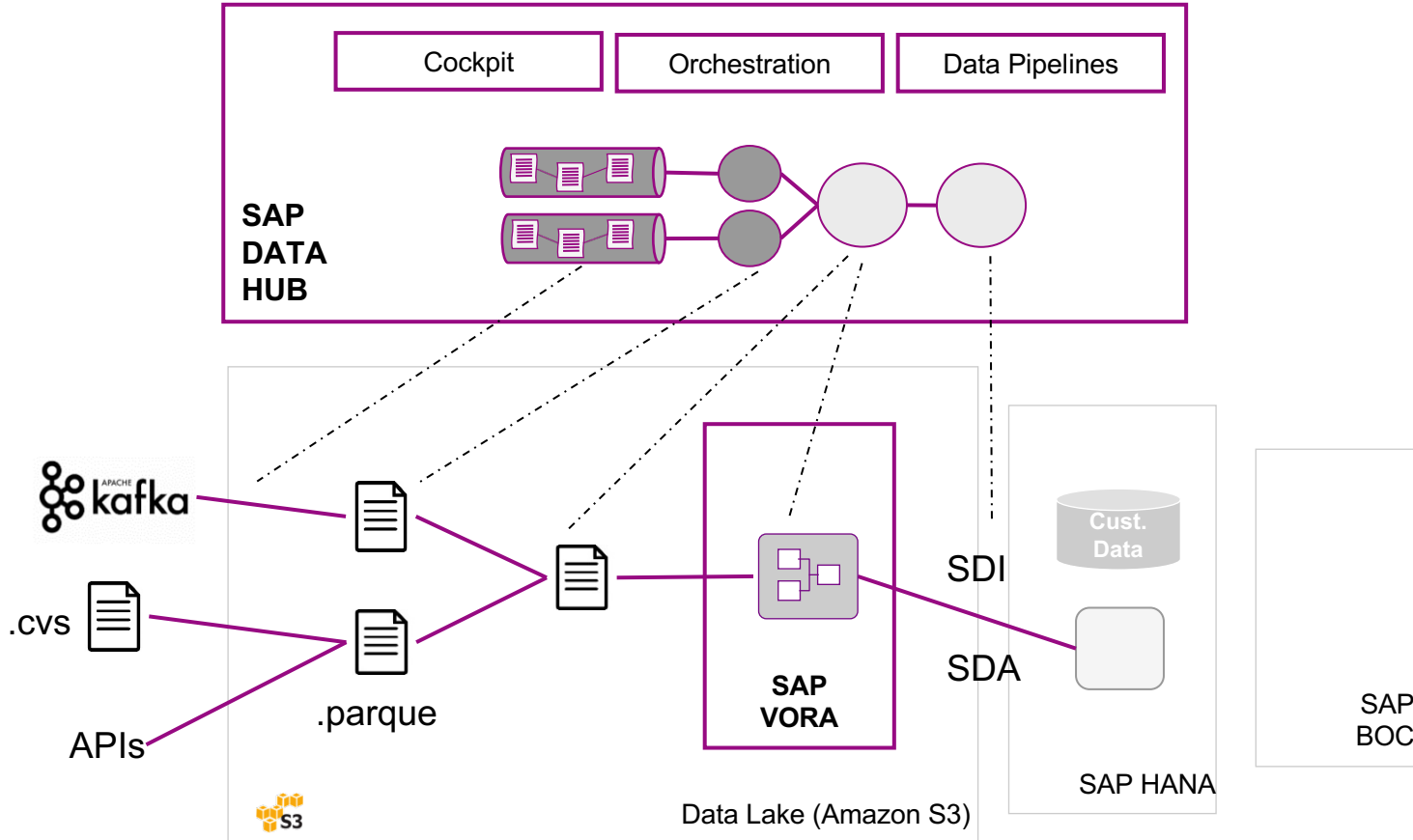
# SAP Data Hub

## Key Functions & Capabilities



- **Data Pipelines** – flow based applications consisting of reusable and configurable operations, e.g. ETL, Preparation, Code Execution, Connectors
- **Workflows** – orchestrate processes across the data landscape, e.g. executing data pipelines, triggering SAP BW Process Chains, SAP Data Services Jobs and many more.
- **Governance** – metadata repository of information stored in the connected landscape. Offering discovery, profiling and search capabilities.

# SAP Data Hub – Streaming and Processing IoT Data



**How to understand customer behavior and to drive insightful decisions?**

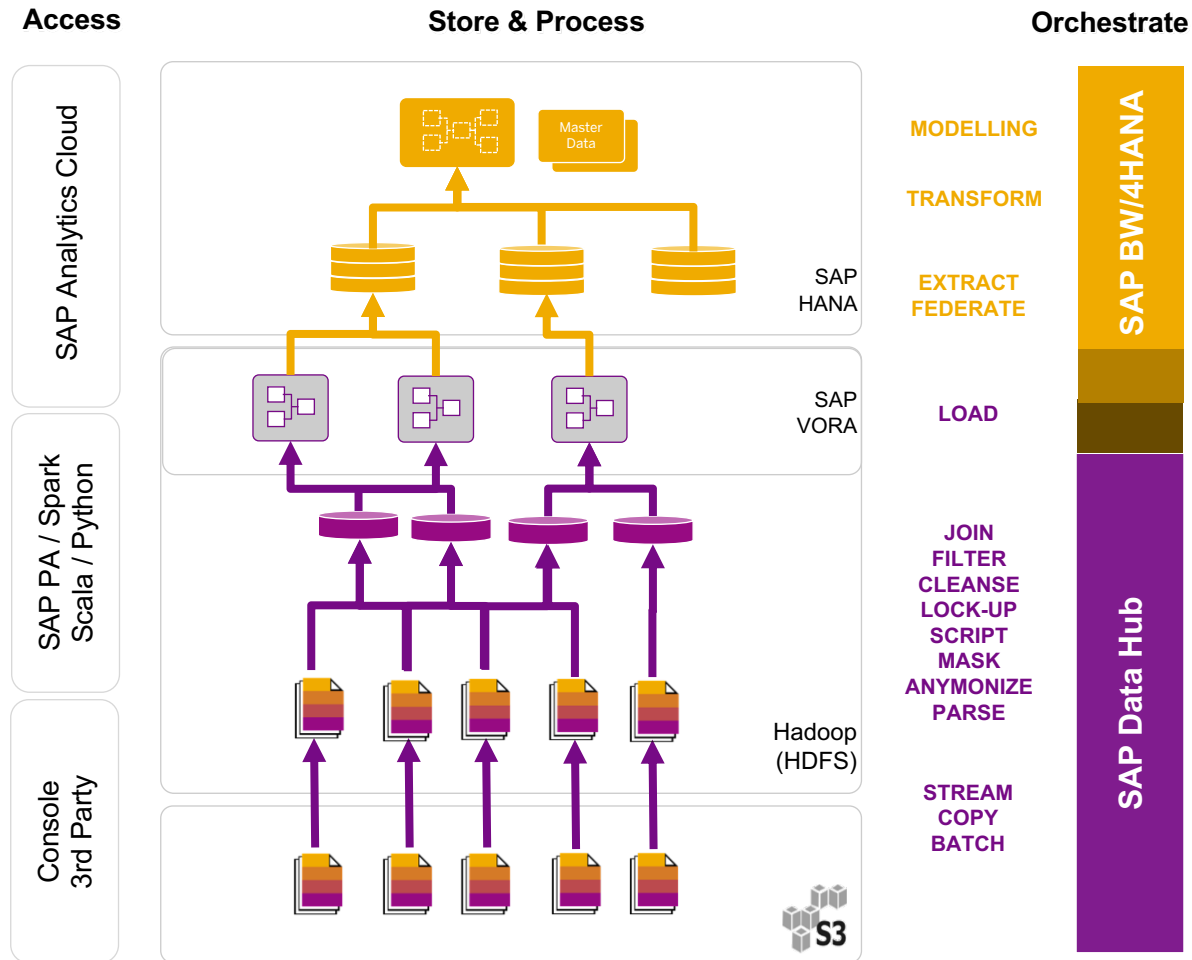
## - Customer Example

- Smart appliances sending sensor data which need to be processed
- ~6 million devices, 16 TB data volume per day

## **Solution with SAP Data Hub**

- Refine business value from data ingestion to enterprise applications
- Visual modeling environment
- Governance and Data Management
- Orchestration and scheduling to define automated data driven processes

# SAP Data Hub – Analyzing Web Log Data with Data Warehouse

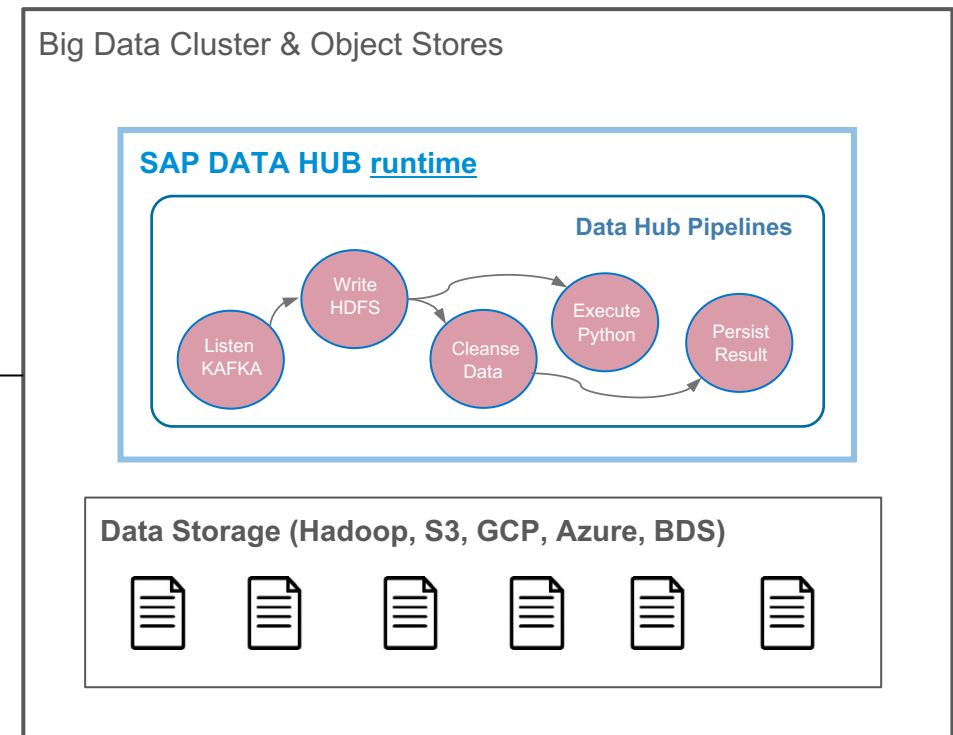
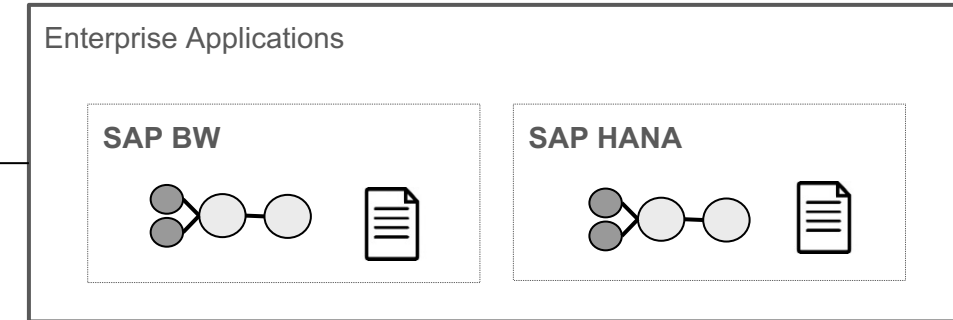
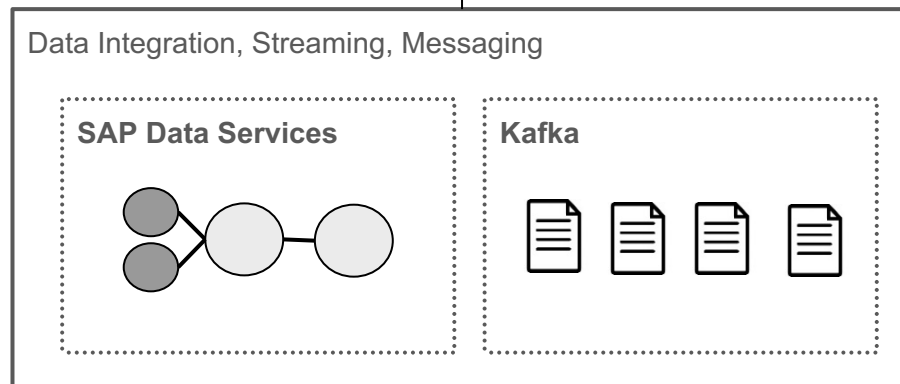
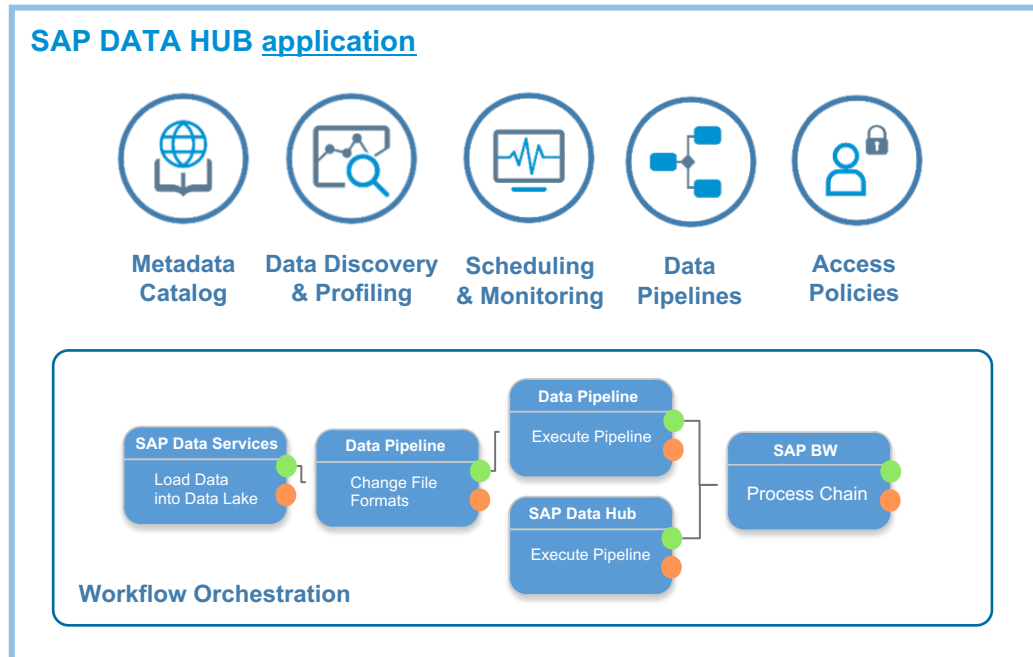


## Example Scenario

- Combine refined big data with enterprise data and corporate master data
- Extract or federate data into SAP BW/4HANA
- Ingest Data into S3 as Landing Zone for data
- Orchestrate and schedule all related processes
- Implement transformations and data pipelines
- Harmonize data structures and look up of reference data
- Execute operations on large data volumes

# SAP Data Hub

## Data Orchestration & Data Pipelines



# External Materials

Main SAP Data Hub Page

<https://www.sap.com/products/data-hub.html>

Franz Faerber (EVP P&I Big Data and SAP Vora) Keynote

<http://events.sap.com/sapandasug/en/session/32274>

Press release and video from Greg McStravick

<https://news.sap.com/new-sap-data-hub-tames-data-landscape/>

TechEd 2017 Demo

<https://www.youtube.com/watch?v=y3XIF-s75Bk>

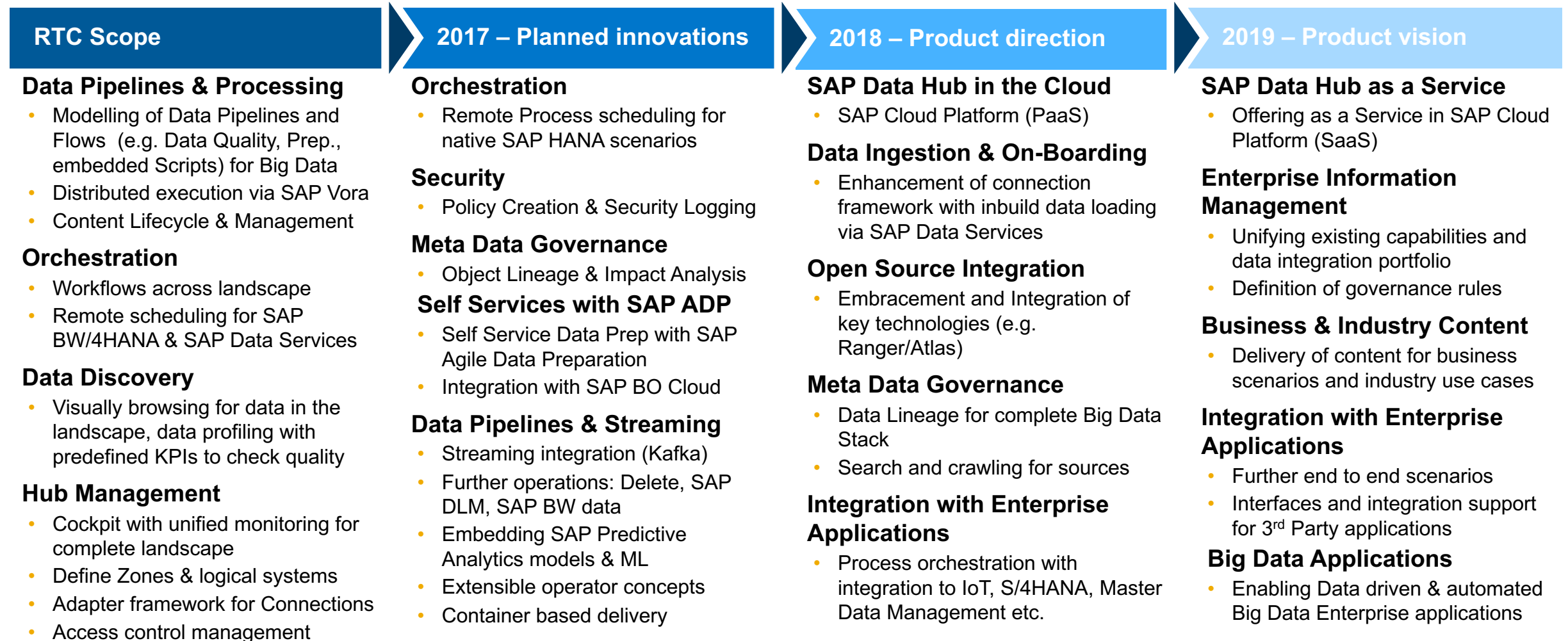
SAP Data Hub Youtube Channel

<https://www.youtube.com/channel/UCLMsNLj0GF0nEw3iixeSEMw>



# SAP Data Hub

## Product road map & vision



This is the current state of planning and may be changed by SAP at any time.

# Thanks!

