



# SAP's Strategy for Big Data and Enterprise Information Management

Andreas Wesselmann, SAP SE  
Session ID # 83238

# About the Speaker

## **Andreas Wesselmann**

- SVP T&I Big Data, SAP SE
- Globally responsible for R&D of Big Data and Data Management solutions

# Key Outcomes/Objectives

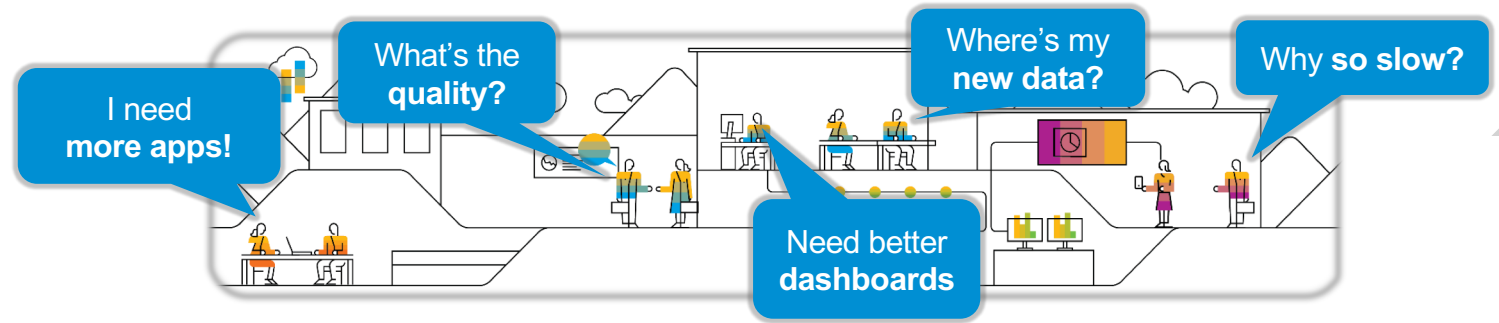
1. Understand SAP's Data Management Strategy and its role in the Intelligent Enterprise
2. Know how the existing EIM products fit into this strategy
3. Learn how SAP Data Intelligence can help you to make the Intelligent Enterprise intelligent

# Agenda

- The customer situation today
- SAP Data Management: The Big Picture
- The Approach for EIM: Integrate and Innovate
- Summary and Outlook



# Enterprise data landscapes are growing increasingly complex



## LANDSCAPE CHALLENGES

### GOVERNANCE

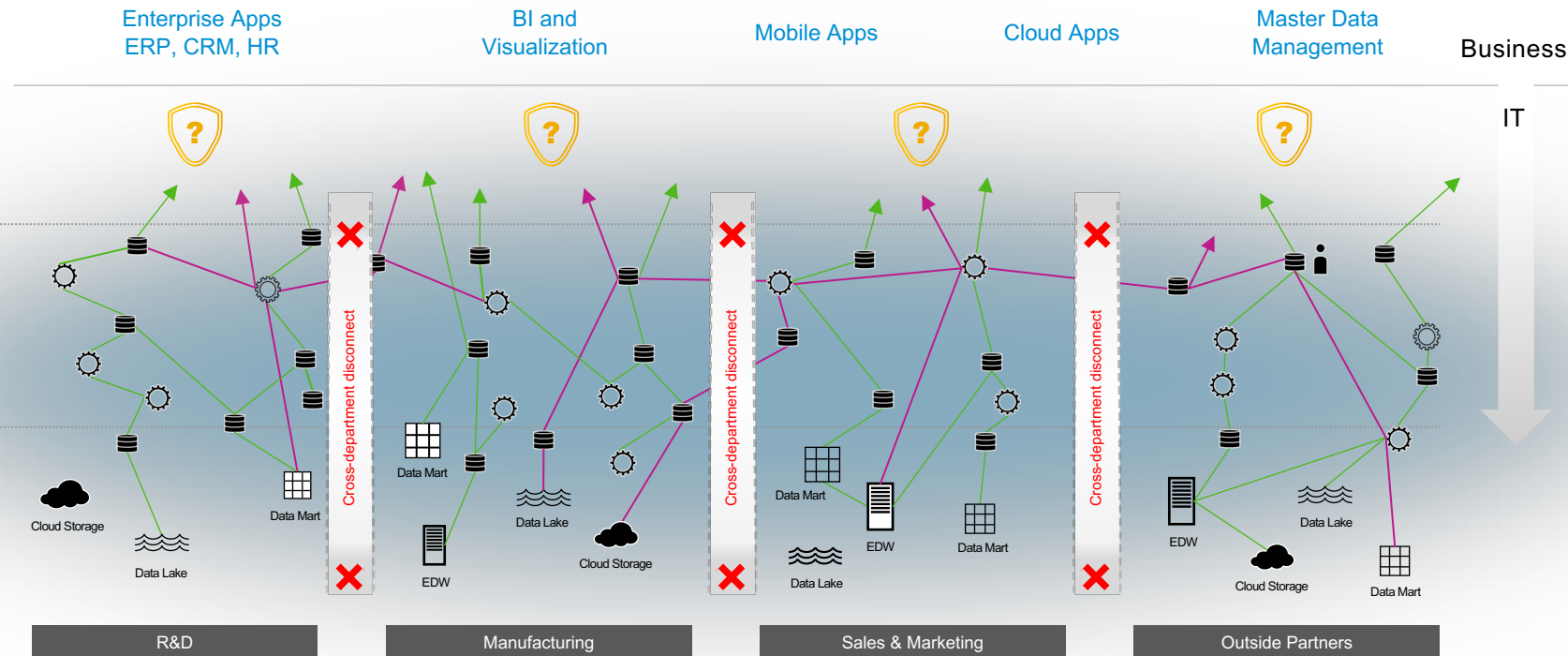
Lack of security and visibility. Who changed the data? What was changed? Who is accessing it?

### LIMITED TOOLS

Lack of enterprise readiness. High effort to productize complex data scenarios across data landscape

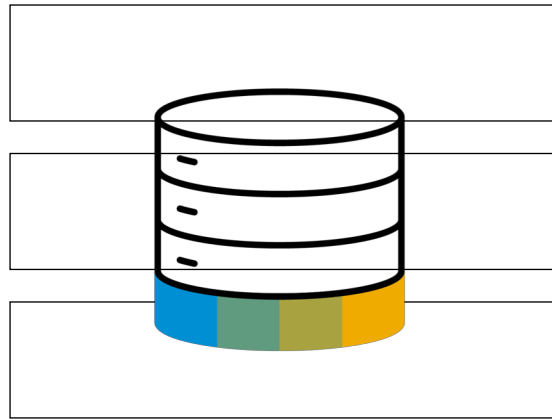
### MISSING LINK

Between Big Data and Enterprise Data. Data is kept in silos across the enterprise.

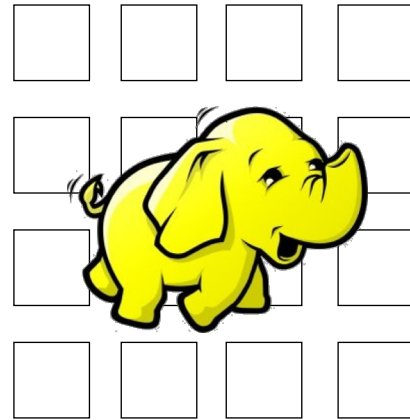


# New challenges require new technologies

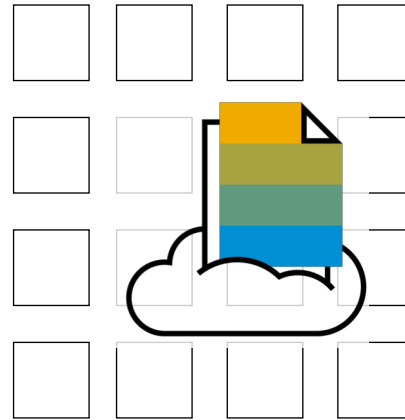
Distributed systems in a distributed landscape



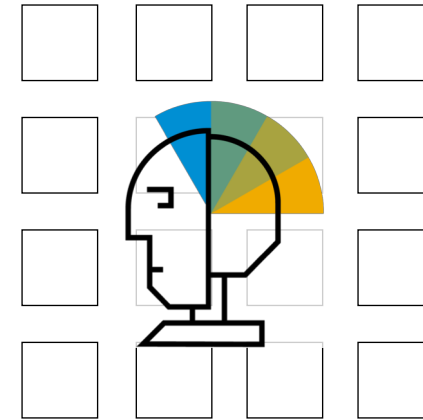
Existing Systems



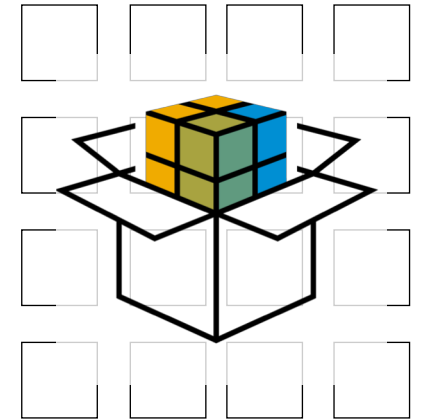
Hadoop/Spark



Cloud Storage  
(i.e. AWS S3)

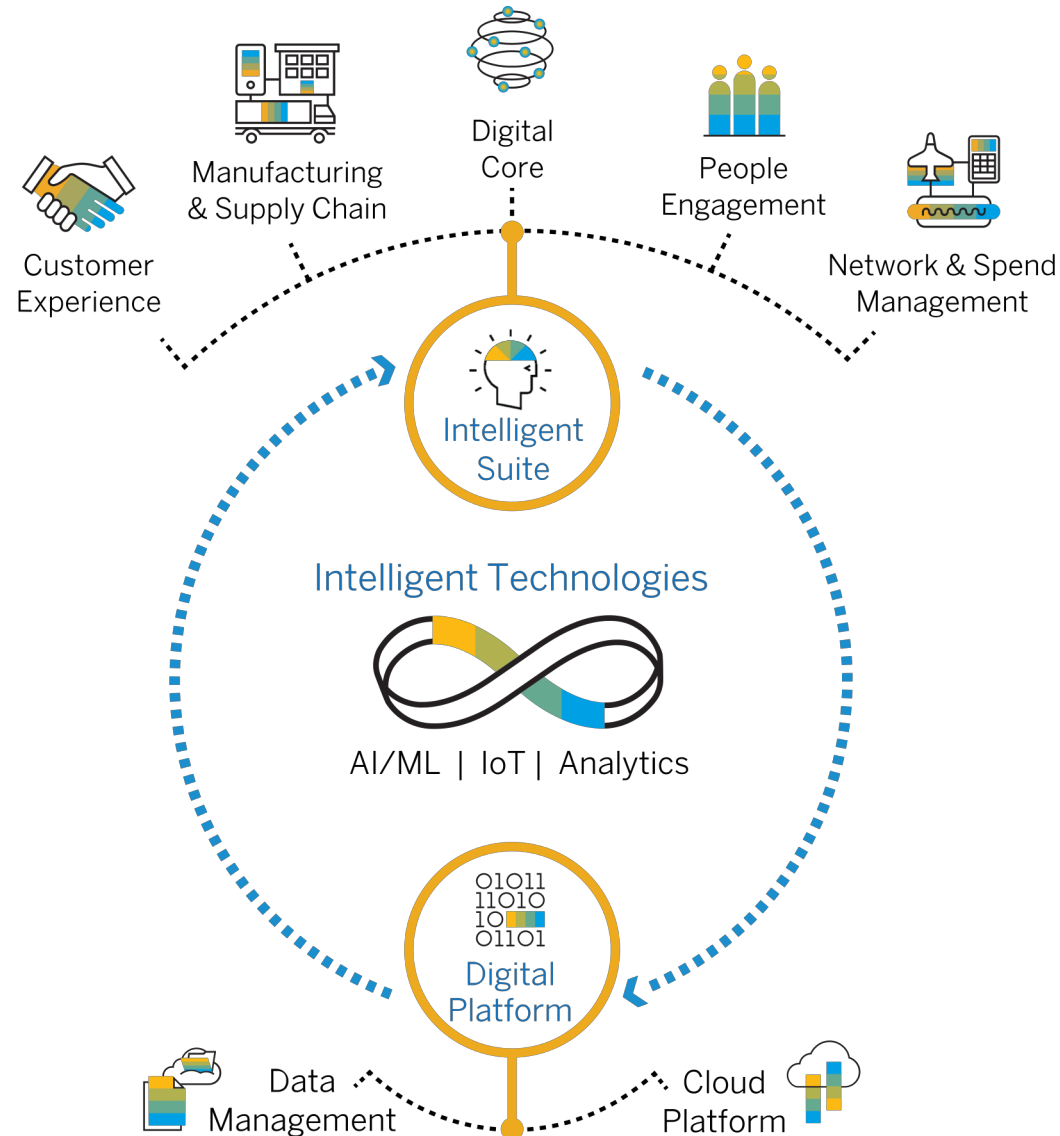


Machine Learning  
(Python, Spark, Tensorflow)



Containers  
(Kubernetes, Docker)

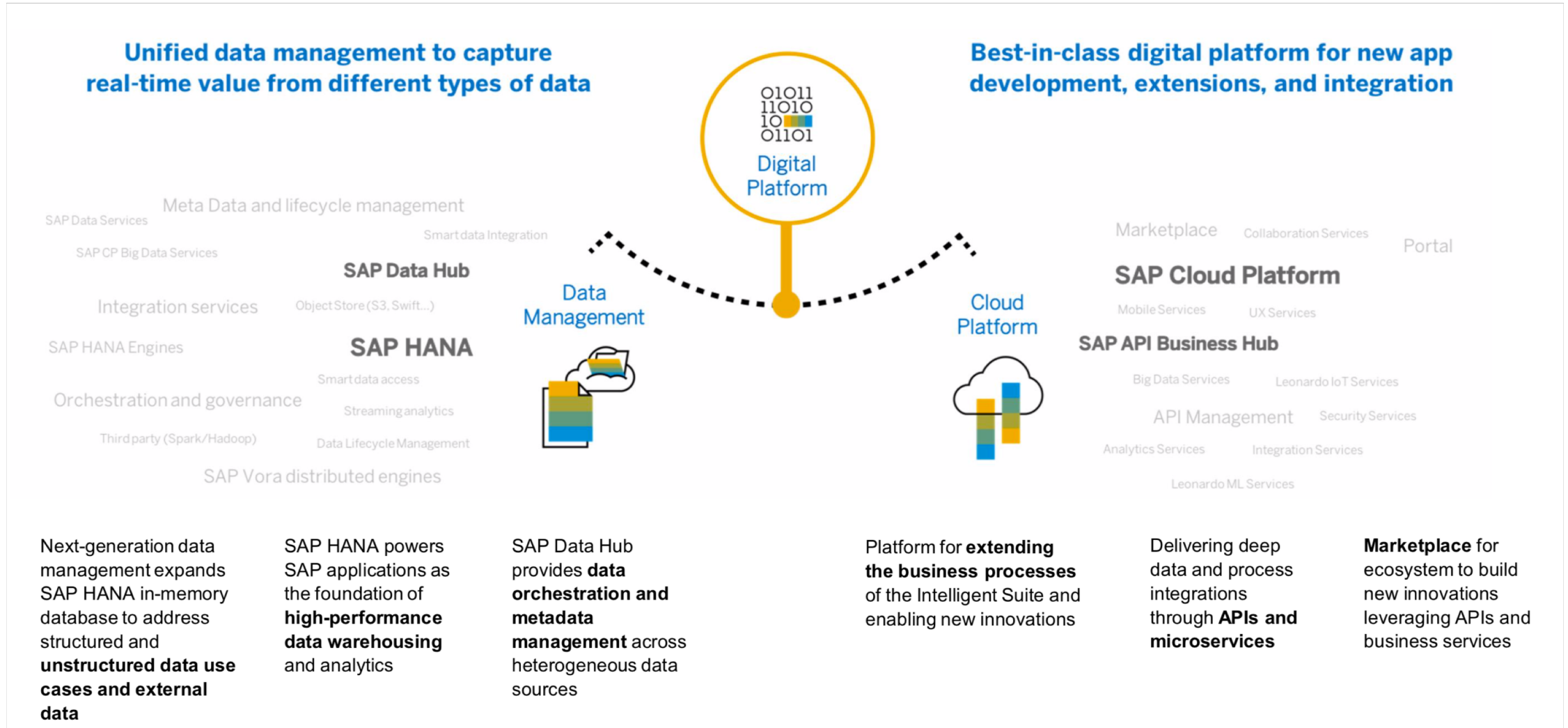
# SAP Strategy – Deliver the Intelligent Enterprise



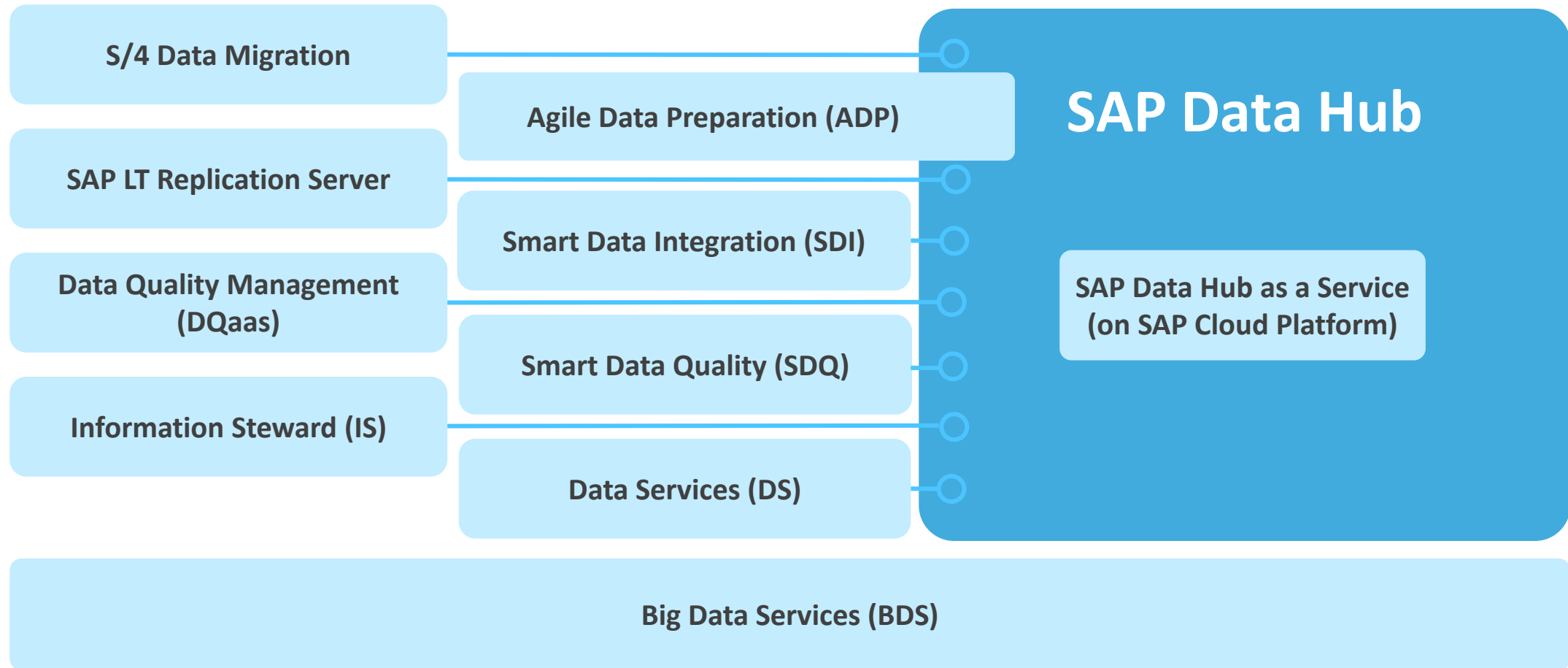
THE INTELLIGENT ENTERPRISE  
features **3 KEY COMPONENTS**

- 1 Intelligent Suite
- 2 Digital Platform
- 3 Intelligent Technologies

# Digital Platform: Unlock data-driven intelligence and innovation

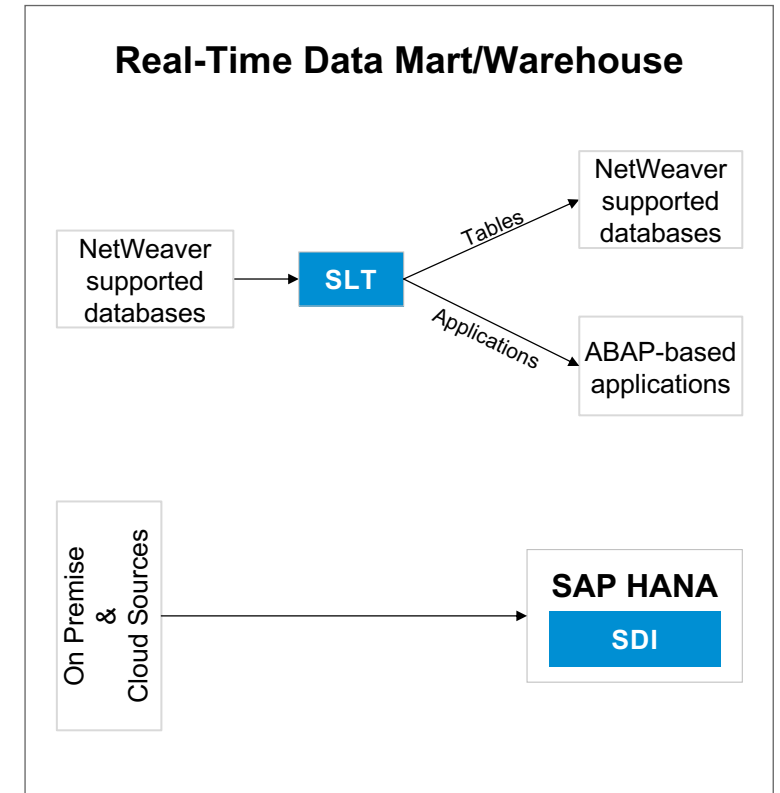
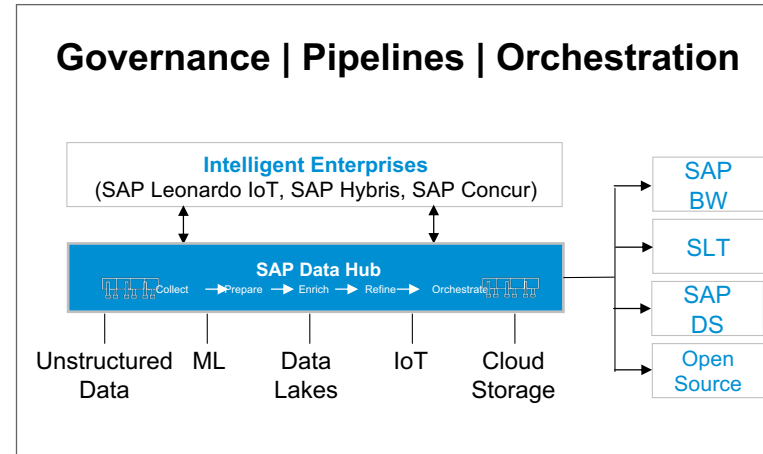
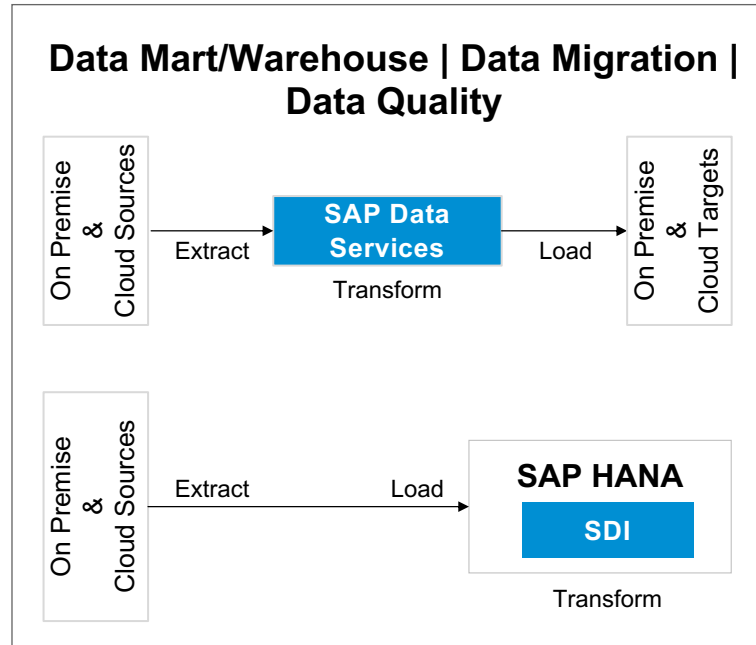


# SAP Data Management Portfolio



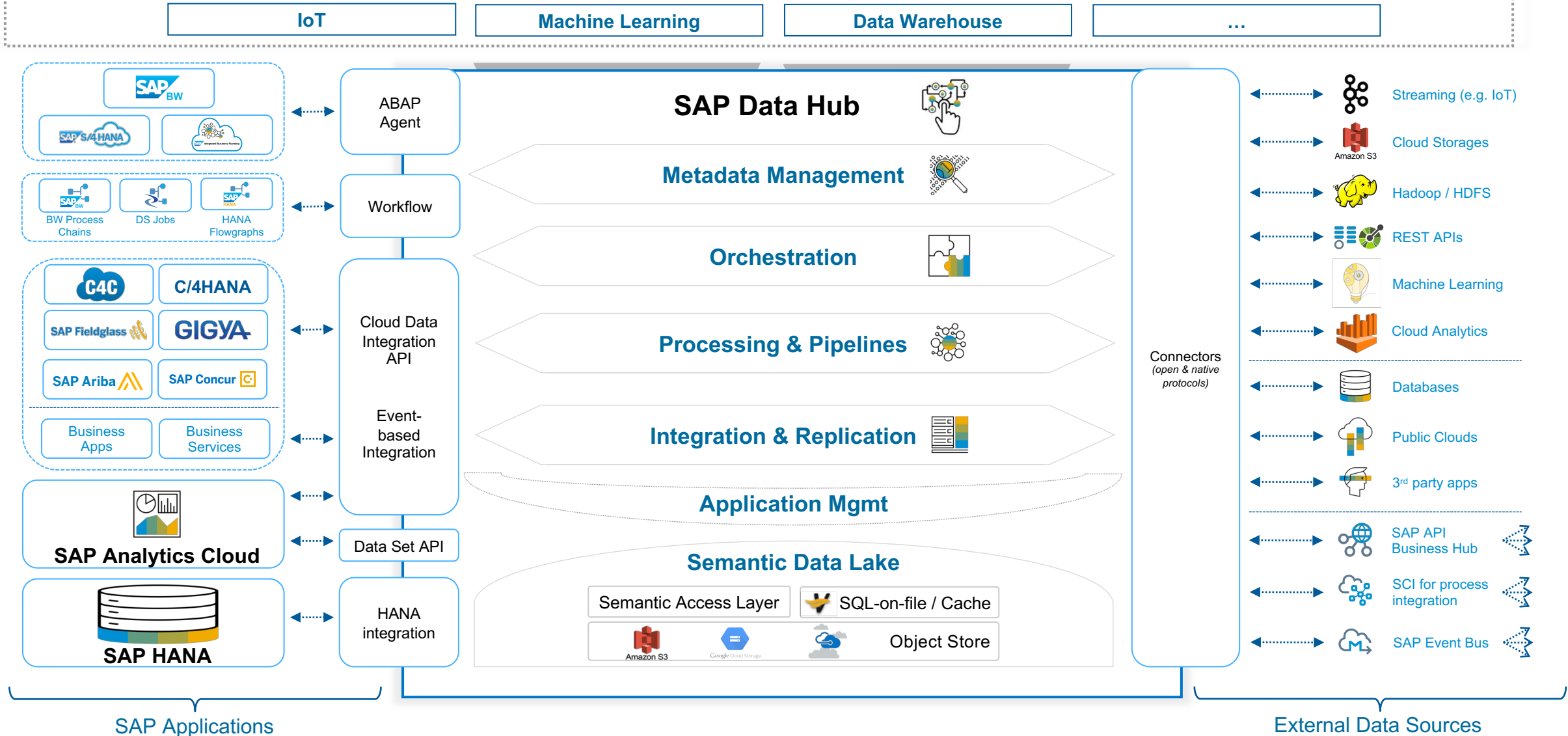
# Data Integration: When to use what?

The main use case scenarios for SAP technologies

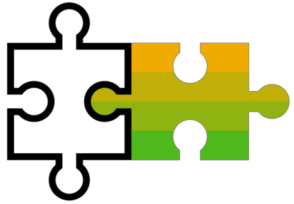


# SAP Data Hub – Unified Data Integration for the Intelligent Enterprise

Data-driven applications



# Key enhancements in SAP Data Hub 2.5



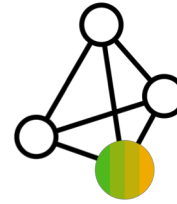
## Enterprise Application Integration

- Standardized interface to integrate SAP Cloud solutions
- Integration model to consume and interact with SAP S/4HANA & SAP Business Suite systems
- Orchestration of SAP Cloud Platform integration to interact with processes



## Meta Data & Data Excellence

- Meta data extraction for SAP S/4HANA & SAP Business Suite systems
- Embedded self-service for data preparation & quality assurance
- Extend anonymization capabilities included in pipeline development model



## Connectivity & Processing

- SQL processing for files
- Node.js as executable environment embedded
- Visualization concept to support pipeline and application development
- Leveraging SAP Cloud Platform Open Connectors to consume external sources



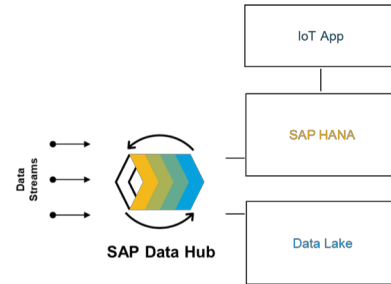
## Deployment & Operations

- Simplified installation process
- Connectivity for High Availability (HA) setups
- Kerberos support for Hadoop services

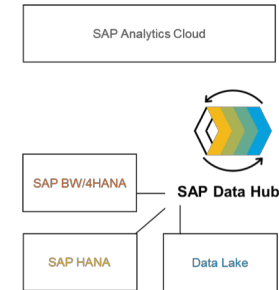


# Patterns and Use Cases

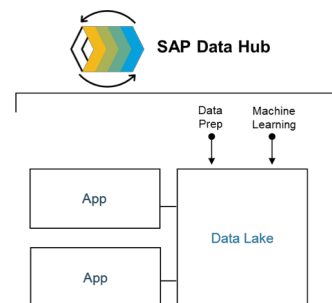
## Overview



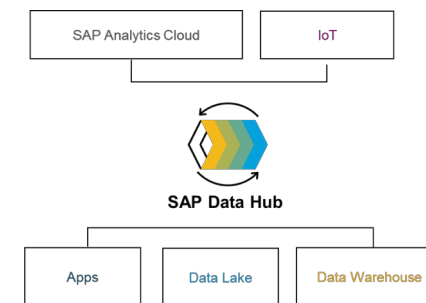
**IoT Ingestion & Orchestration**  
Understand real-world performance



**Intelligent Data Warehouse**  
Rapidly integrate and leverage new data sources



**Data Science &  
ML Data Management**





**Governance / Data Cataloging**  
Understand and secure your data

# Recent analyst perspectives ...

Leading Integration Offerings  
SAP / NON-SAP DI TOOLS & DI PLATFORMS



DI Tools	NON-SAP	DI Platforms
 SAP Data Services SAP SDI SAP SLT SAP CPI-DS	    	 Azure Data Factory  SAP Data Hub

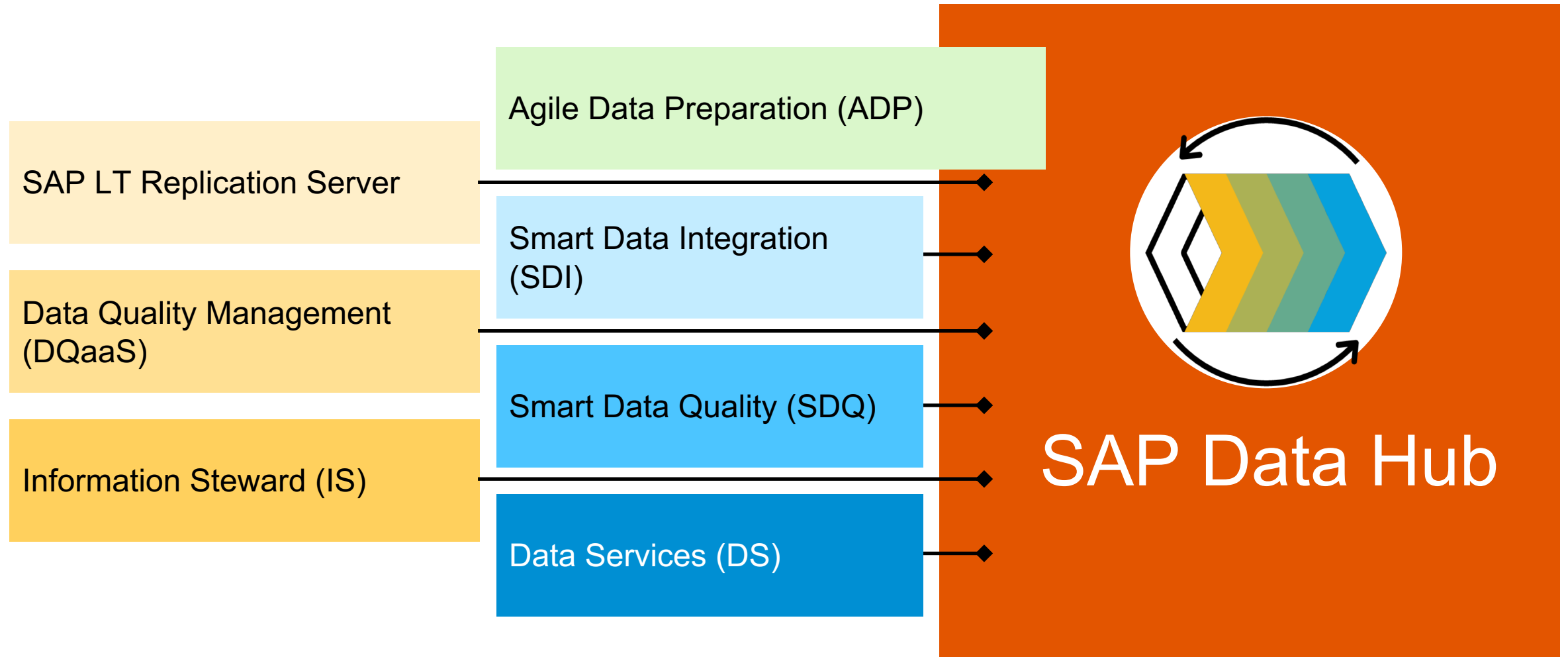
	SAP Tools (SAP DS, SDI, SLT, CPI-DS)	Informatica	DataStage	SSIS	Talend
Monitoring	✓ Web based load monitor	✓ Workflow monitor	✓ Director	✓ SQL Server Management Studio	✓
Error handling	✓	✓	✓	✓	✓
Scheduling	✓ Event based - external Support need	✓ Workflow manager	✓ Administrator	✓ SQL Server Agent Jobs	✓ Scheduling available only in licensed version via Talend Admin Console
Learning Curve	✓ Quick and Easy to learn	✓ Quick and Easy to learn	✗ Moderately Difficult	✓ Easy to learn	✗ JAVA knowledge is a must
Change Control	✓	✓	✓	✗	✓
Auto Documentation	✓	✓	✓	✗	✗
Data Lineage	✓ End-to-End advanced analysis	✓ Detailed forward/backward analysis	✓ Full data lineage reports	✗	✓

✓ All features supported    ✗ Does not support features

<http://visualbi.com/blogs/business-intelligence/battle-eim-etl-tools-sap-bodsslt-sdi-data-hub-vs-informatica-datastage-microsoft-ssis-mulesoft-talend/>

# The Approach for EIM: Integrate and Innovate

Outlook for data integration and processing portfolio



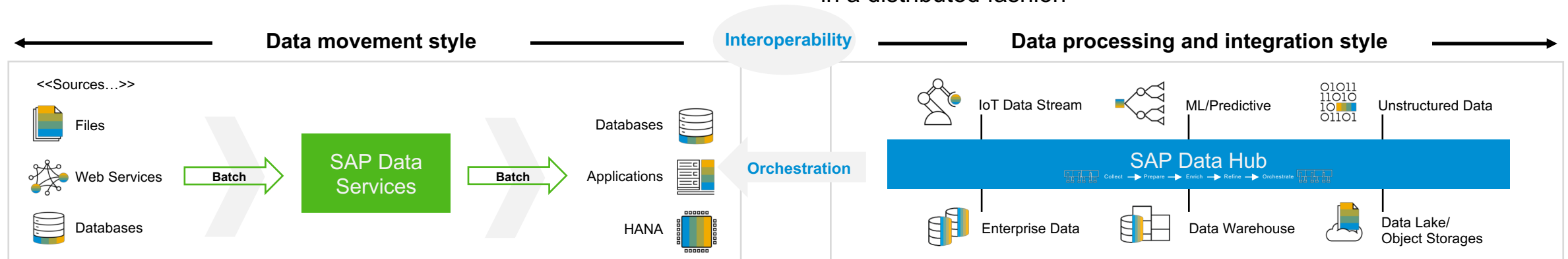
# Complementary solutions: SAP Data Services and SAP Data Hub

## SAP Data Services

Moving application data from transactional sources to data warehouses, including data quality processes and built-in Data Integration transformations

## SAP Data Hub

A pipeline-driven data integration, operations and governance solution for disparate kinds of data (structured, unstructured, streaming, cloud etc.), supporting both integration and processing in a distributed fashion



### Key use cases

- BI / Traditional Data Warehousing, Data Migration, Data Quality

### Characteristics

- ETL in a standalone heterogeneous landscape
- Centralized, on premise & server-based infrastructure
- Relational data focused
- Advanced data transformations & processing (e.g. Join, SQL, DQ...)

### Key use cases

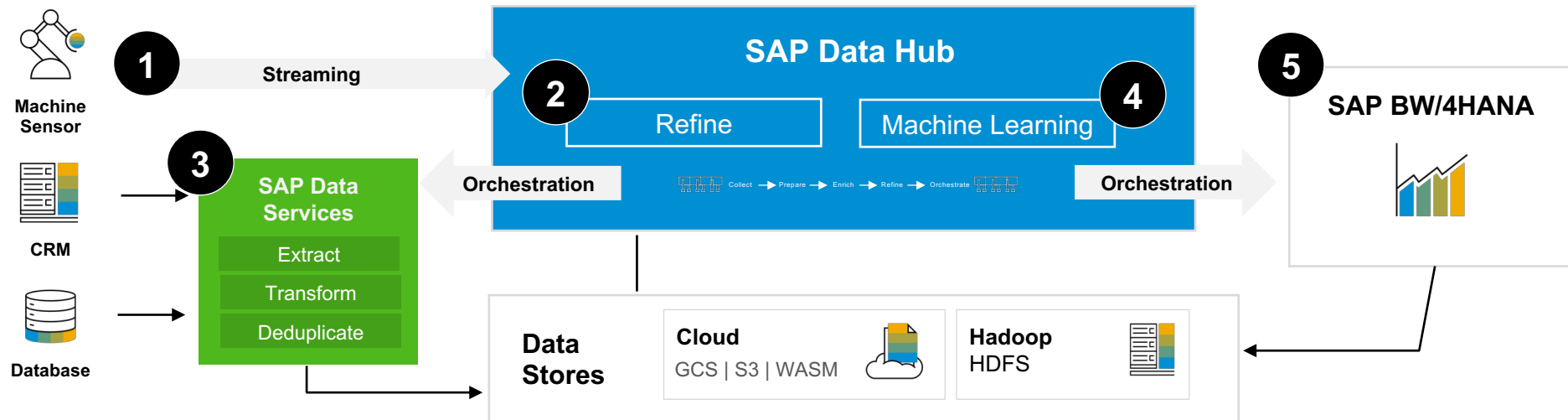
- Data science & Machine Learning, Big Data Warehousing, IoT, Data Tiering, Data Network

### Characteristics

- Support multiple data ingestion methods
- Pipelining and orchestration of data processes in complex landscapes
- Cloud & on-premise deployment, distributed data processing & serverless computing via Kubernetes
- Big data focused (tables/views, object storages, any data formats), for both data at rest and data in motion
- Complex data refinery & processing (e.g. ML, Predictive, Image, custom code)

# SAP Data Services and SAP Data Hub

## Interoperability: Example for combined scenario



1

Ingest large volumes of data (e.g. distance, pace, heartrate, location...) from machine sensors by using MQTT/Kafka operator (SAP Data Hub)

2

Refine data according to purpose and store it in data stores (SAP Data Hub)

3

Acquire additional relevant structured data (e.g. customers, sales, behavioral, demographic) into data stores by remotely orchestrating DS jobs (leveraging existing SAP Data Services investments)

4

Apply machine learning algorithms (e.g. classification, clustering, identifying outliers, etc.) on the data to discover new insights about user characteristics (SAP Data Hub)

5

Invoke process chain to ingest the results into SAP BW/4HANA for further data analysis and reporting (SAP Data Hub)

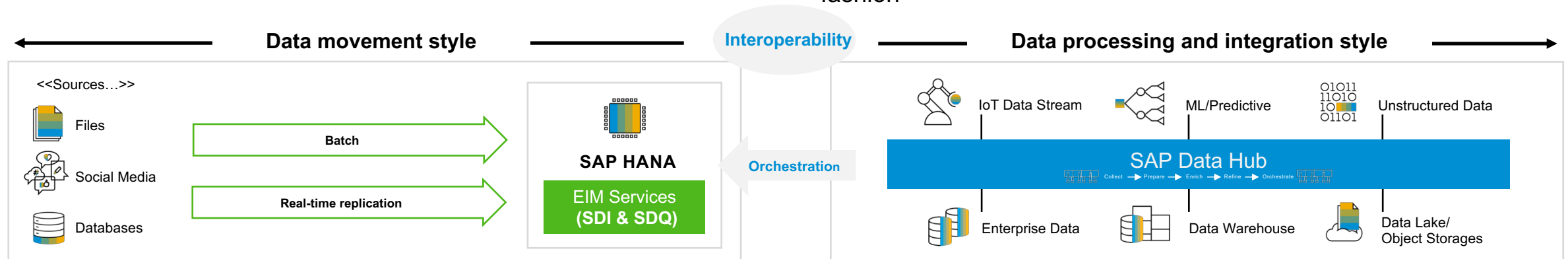
# Complementary solutions: SAP HANA SDI/SDQ and SAP Data Hub

## SAP HANA SDI/SDQ

Are features of the SAP HANA Platform that provides data integration and data quality capabilities

## SAP Data Hub

A pipeline-driven data integration, operations and governance for disparate kinds of data (structured, unstructured, streaming, cloud etc.), supporting both integration and processing in a distributed fashion



### Key use cases

- Provision, cleanse and load data from different sources into SAP HANA in memory platform

### Characteristics

- HANA-centric (one environment in which to provision and consume data)
- In-memory performance
- Centralized, server-based infrastructure
- Relational data focused
- Advanced data transformations & processing (e.g. Join, SQL, DQ,...)

### Key use cases

- Data Science & Machine Learning, Big Data Warehousing, IoT, Data Tiering, Data Network

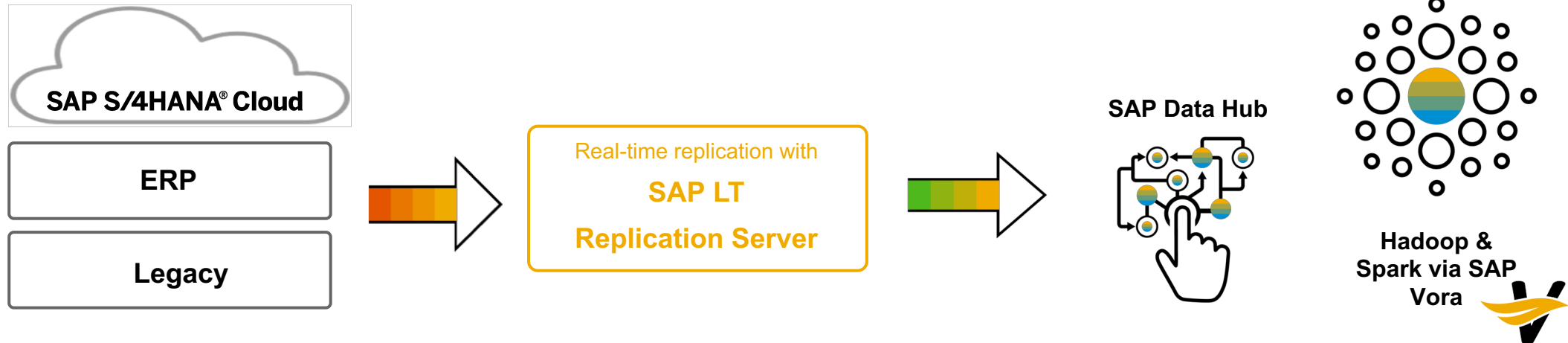
### Characteristics

- Support multiple data ingestion methods
- Pipelining and orchestration of data processes in complex landscapes
- Cloud & on-premise deployment, distributed data processing & serverless computing via Kubernetes
- Big Data focused (tables/views, object storages, any data formats), for both data at rest and data in motion
- Complex data refinery & processing (e.g. ML, Predictive, Image, custom code)

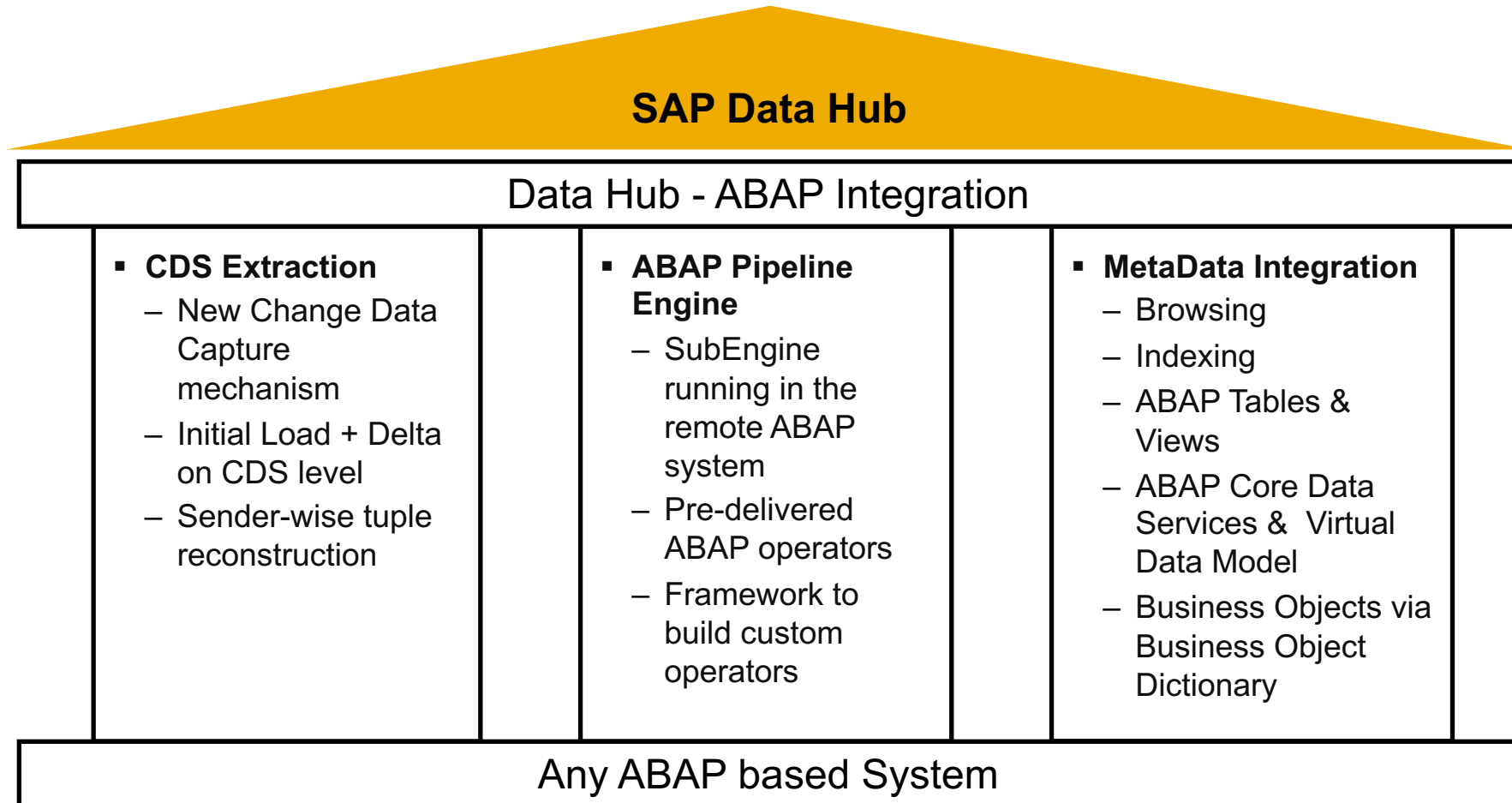
# Data Integration & Connectivity

Replication via SAP LT Replication Server into SAP Vora

## The direct replication of data to SAP Vora with SAP LT Replication Server



# Integration with ABAP-based systems





# Enterprise Application Integration

Orchestration of SAP Cloud Platform integration to interact with processes

## Enrichment of existing Sap Cloud Platform Integration scenarios with SAP Data Hub Modeler functionality

### Main Use Cases

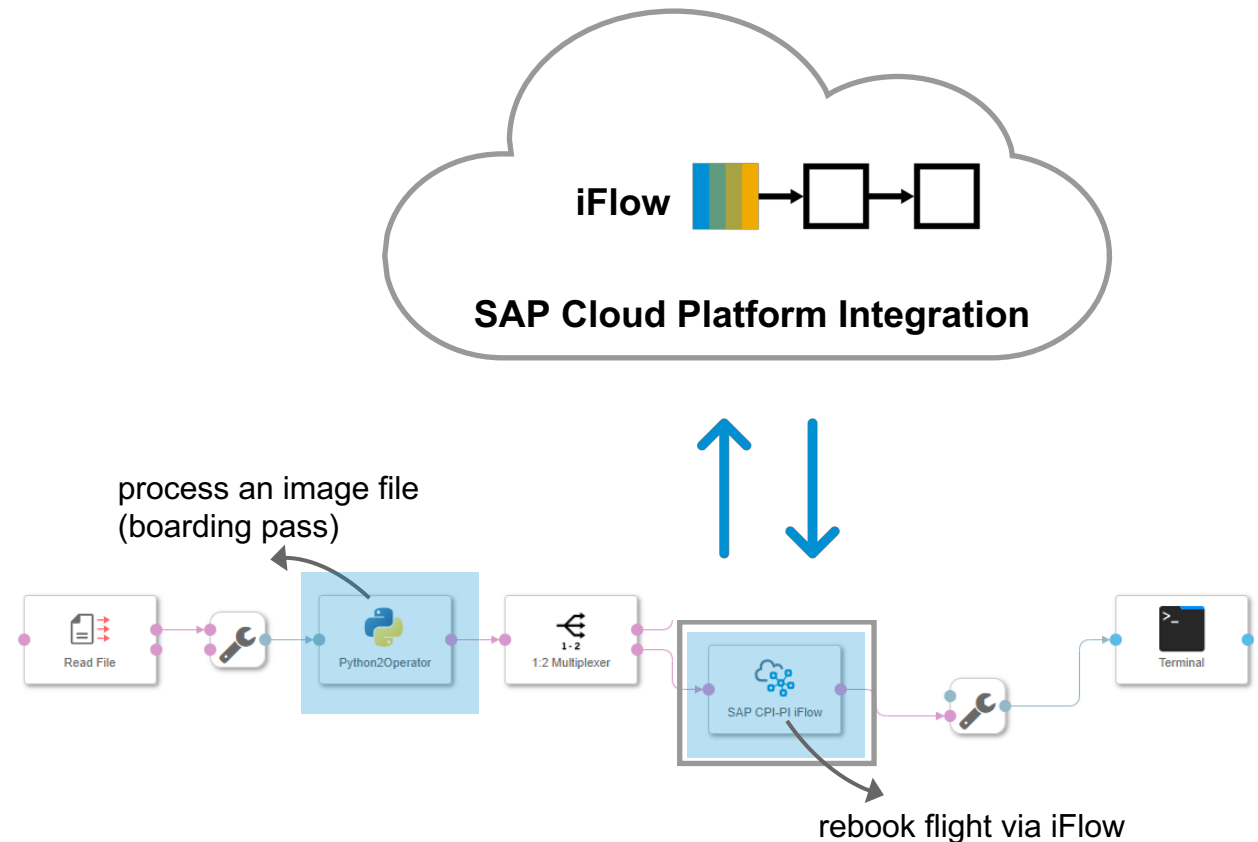
- Enable SAP Cloud Platform Integration – Process Integration customers with the ability to reuse existing assets (iFlow)
- Broaden scenarios by blend in complementing capabilities for example machine learning or image processing

### Capabilities

Pre-defined operator for triggering iFlows

Smooth connection via HTTP Basic authentication

OAUTH V2 is planned for future release

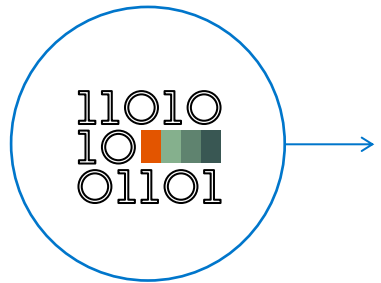
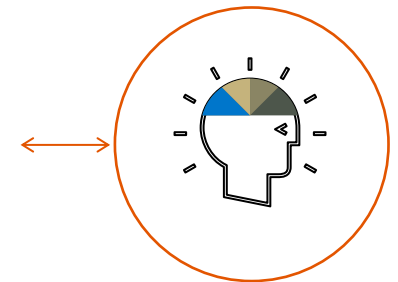
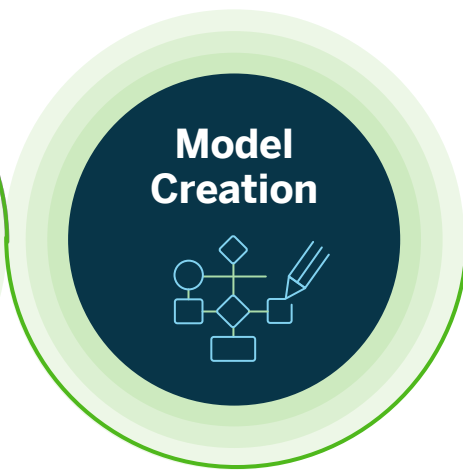


# Machine Learning & Data Science

Orchestration | Integration | Operationalization

## Machine Learning IDE

ML research | Model publishing | Lifecycle management

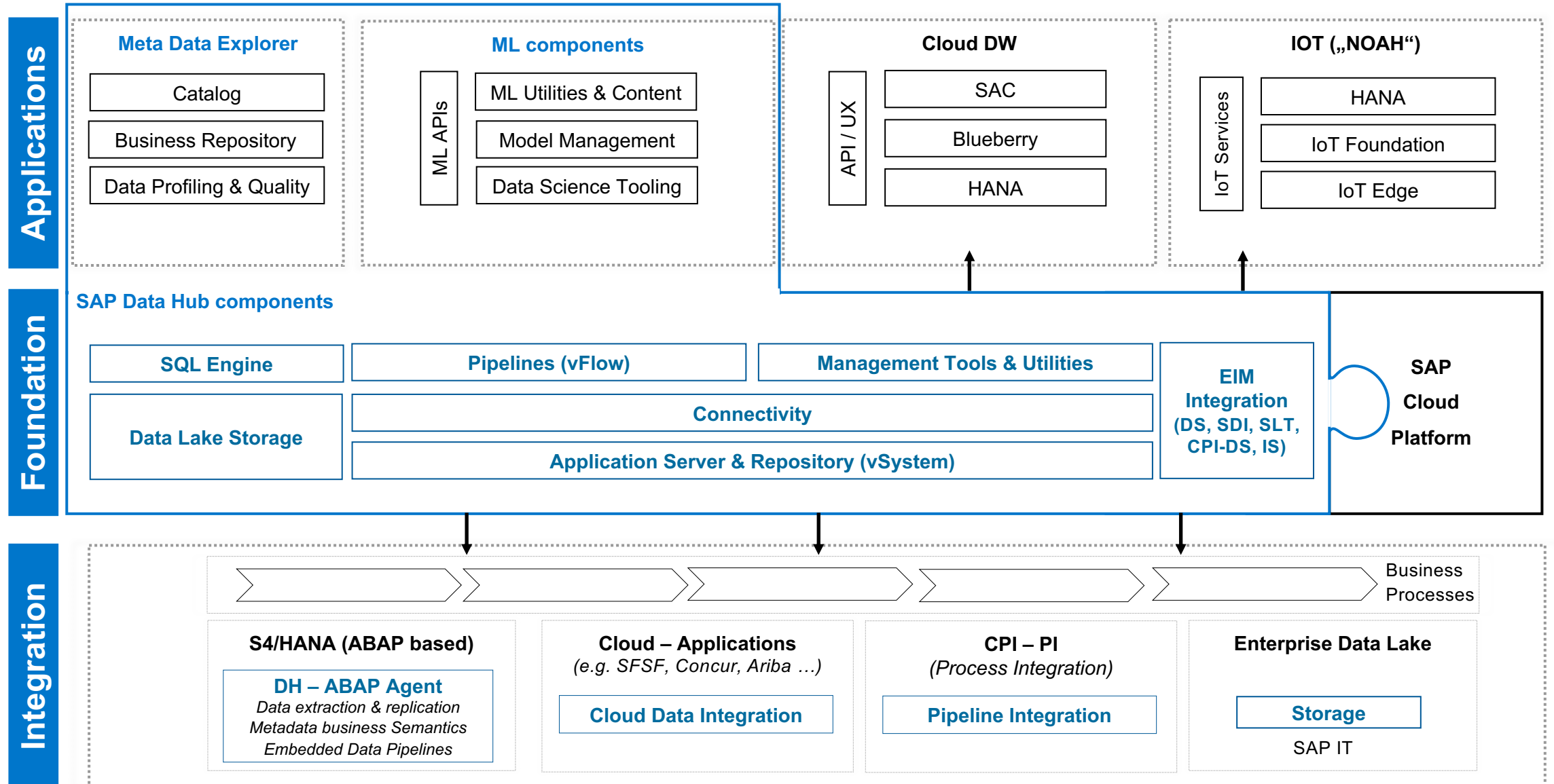


Enterprise Data Sources

Intelligent Suite

## ML & Data Science Repository

# SAP Data Intelligence: SAP Data Hub + Machine Learning



# Summary

SAP offers a complete Data Management Portfolio to solve your business challenges

SAP Data Hub complements and integrates the existing SAP EIM solutions

- Investment protection for your existing solutions

SAP Data Hub (with SAP HANA) forms the Data Foundation of the Intelligent Enterprise

- With best (pre-defined) integration into all SAP solutions

SAP Data Hub is open and uses state-of-art technology to build new, innovative, intelligent and data driven solutions integrated in your business processes

[Start your projects now!](#)

# Take the Session Survey.

We want to hear from you! Be sure to complete the session evaluation on the SAPPHIRE NOW and ASUG Annual Conference mobile app.



# Presentation Materials

Access the slides from 2019 ASUG Annual Conference here:

<http://info.asug.com/2019-ac-slides>

# Q&A

For questions after this session, contact us at  
[andreas.wesselmann@sap.com](mailto:andreas.wesselmann@sap.com).

# Let's Be Social.

Stay connected. Share your SAP experiences anytime, anywhere.

Join the ASUG conversation on social media: **@ASUG365 #ASUG**

