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

Unified data management to capture real-time value from different types of data

Best-in-class digital platform for new app development, extensions, and integration

Next-generation data management expands SAP HANA in-memory database to address structured and unstructured data use cases and external data

SAP HANA powers SAP applications as the foundation of high-performance data warehousing and analytics

SAP Data Hub provides data orchestration and metadata management across heterogeneous data sources

Platform for extending the business processes of the Intelligent Suite and enabling new innovations

Delivering deep data and process integrations through APIs and microservices

Marketplace for ecosystem to build new innovations leveraging APIs and business services
SAP Data Management Portfolio

- S/4 Data Migration
- SAP LT Replication Server
- Data Quality Management (DQaaS)
- Information Steward (IS)
- Agile Data Preparation (ADP)
- Smart Data Integration (SDI)
- Smart Data Quality (SDQ)
- Data Services (DS)
- Big Data Services (BDS)

SAP Data Hub

SAP Data Hub as a Service (on SAP Cloud Platform)
Data Integration: When to use what?
The main use case scenarios for SAP technologies

**Data Mart/Warehouse | Data Migration | Data Quality**

- On Premise & Cloud Sources
  - Extract → SAP Data Services
  - Transform → Load
- On Premise & Cloud Sources
  - Extract → SAP HANA SDI
  - Transform → Load

**Governance | Pipelines | Orchestration**

Intelligent Enterprises
(SAP Leonardo IoT, SAP Hybris, SAP Concur)

- SAP Data Hub
- Unstructured Data
- ML
- Data Lakes
- IoT
- Cloud Storage

- SAP
- BW
- SLT
- SAP DS
- Open Source

**Real-Time Data Mart/Warehouse**

- NetWeaver supported databases
- ABAP-based applications
- SAP HANA SDI
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 ...
The Approach for EIM: Integrate and Innovate
Outlook for data integration and processing portfolio

SAP LT Replication Server
Data Quality Management (DQaaS)
Information Steward (IS)

Agile Data Preparation (ADP)
Smart Data Integration (SDI)
Smart Data Quality (SDQ)
Data Services (DS)

SAP Data Hub
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

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…)

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
• 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

**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,....)

---

**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**
- 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)
The direct replication of data to SAP Vora with SAP LT Replication Server
Integration with ABAP-based systems

SAP Data Hub

Data Hub - ABAP Integration

- **CDS Extraction**
  - New Change Data Capture mechanism
  - Initial Load + Delta on CDS level
  - Sender-wise tuple reconstruction

- **ABAP Pipeline Engine**
  - SubEngine running in the remote ABAP system
  - Pre-delivered ABAP operators
  - Framework to build custom operators

- **MetaData Integration**
  - Browsing
  - Indexing
  - ABAP Tables & Views
  - ABAP Core Data Services & Virtual Data Model
  - Business Objects via Business Object Dictionary

Any ABAP based System
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

process an image file (boarding pass)
rebook flight via iFlow

© 2019 SAP SE or an SAP affiliate company. All rights reserved. | PUBLIC
Machine Learning & Data Science

Orchestration | Integration | Operationalization

Machine Learning IDE
ML research | Model publishing | Lifecycle management

Data Preparation

Model Creation

Service Deployment & Operations

ML & Data Science Repository

Enterprise Data Sources

Intelligent Suite

© 2019 SAP SE or an SAP affiliate company. All rights reserved.
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.
Let’s Be Social.

Stay connected. Share your SAP experiences anytime, anywhere.
Join the ASUG conversation on social media: @ASUG365 #ASUG