SAP Data Hub:
Data Integration with Enterprise Applications

DAT202

PUBLIC

SAP TECHED

Speakers

Las Vegas

September 24–27, 2019

Marc Hartz

Barcelona

October 8-10, 2019

Jens Rannacher

Bangalore

November 13-15, 2019

Lalitendu Samantray

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Take the session survey.

We want to hear from you!

Complete the session evaluation for this session **DAT202** on the SAP TechEd mobile app.

Download the app from
iPhone App Store or Google Play.

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Disclaimer

The information in this presentation is confidential and proprietary to SAP and may not be disclosed without the permission of SAP. Except for your obligation to protect confidential information, this presentation is not subject to your license agreement or any other service or subscription agreement with SAP. SAP has no obligation to pursue any course of business outlined in this presentation or any related document, or to develop or release any functionality mentioned therein.

This presentation, or any related document and SAP's strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information in this presentation is not a commitment, promise or legal obligation to deliver any material, code or functionality. This presentation is provided without a warranty of any kind, either express or implied, including but not limited to, the implied warranties of merchantability, fitness for a particular purpose, or non-infringement. This presentation is for informational purposes and may not be incorporated into a contract. SAP assumes no responsibility for errors or omissions in this presentation, except if such damages were caused by SAP’s intentional or gross negligence.

All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, which speak only as of their dates, and they should not be relied upon in making purchasing decisions.

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Agenda

SAP Data Hub

- Introduction
- Capability Overview

Data Integration

- With SAP Applications
- With non-SAP Applications
- Leveraging SAP Cloud Platform

End-2-end Demo

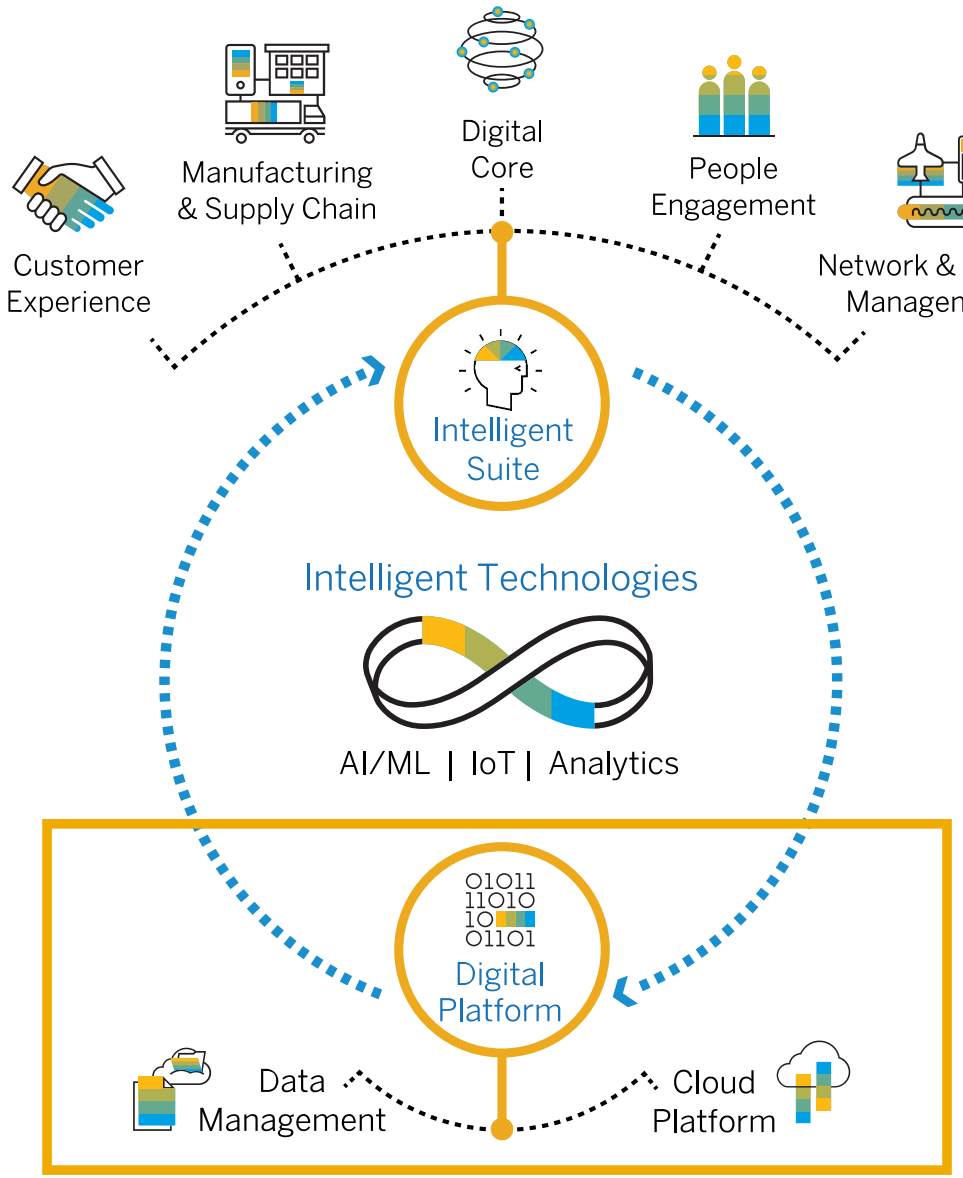
Q&A

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub Introduction

SAP Strategy – Deliver the Intelligent Enterprise

THE INTELLIGENT ENTERPRISE features 3 KEY COMPONENTS:



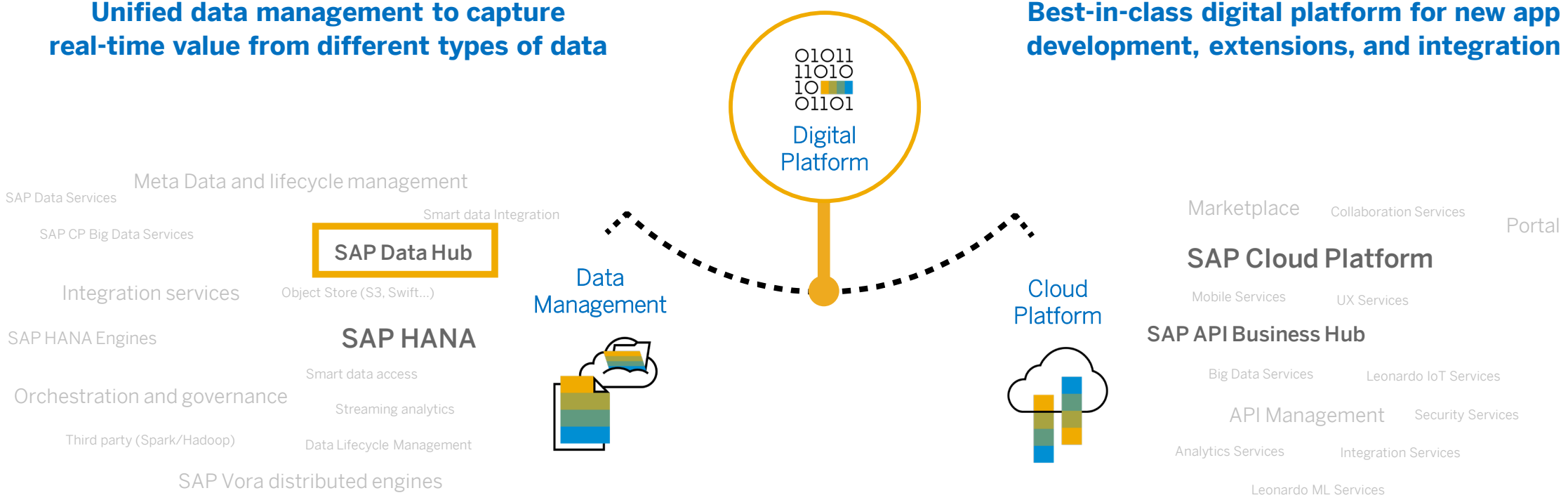
1. Intelligent Suite
2. Digital Platform
3. Intelligent Technologies

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

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

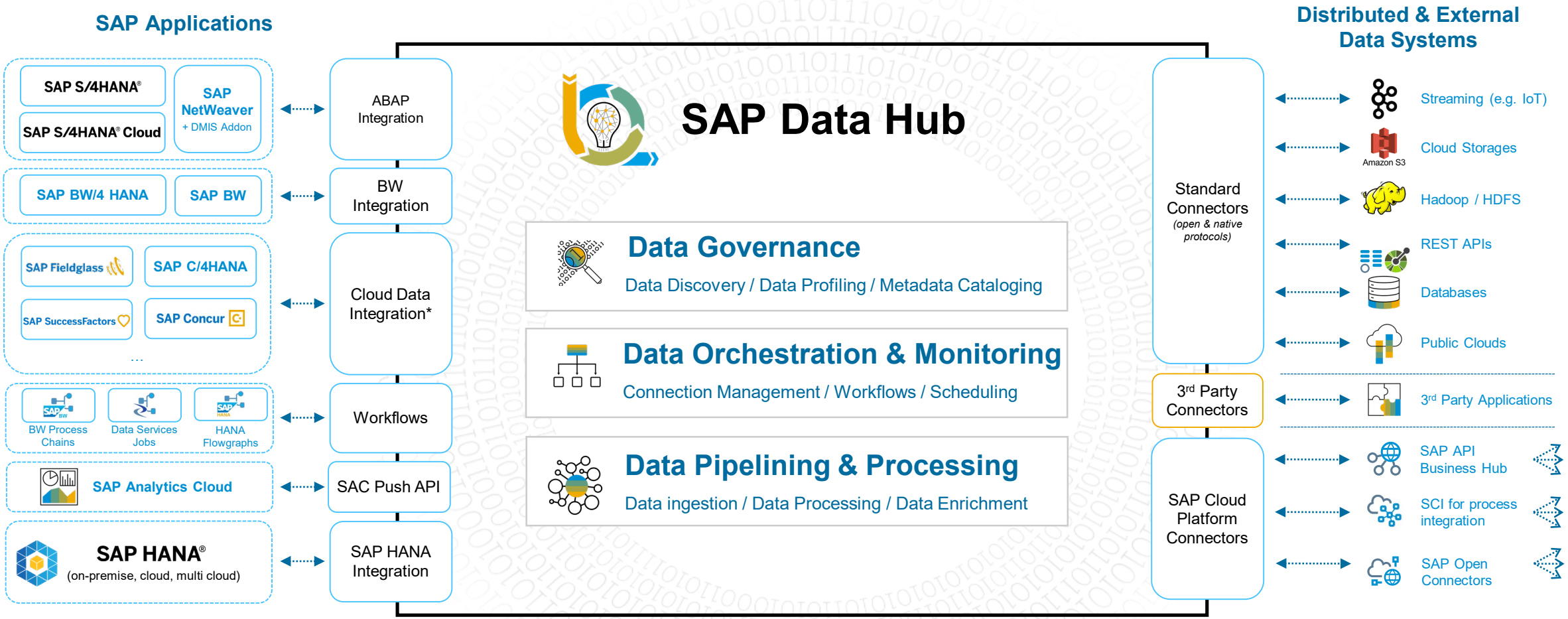
Platform for **extending the business processes** of the Intelligent Suite and enabling new innovations

Delivering deep data and process integrations through **APIs and microservices**

Marketplace for ecosystem to build new innovations leveraging APIs and business services

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub – Unified Data Integration for the Intelligent Enterprise



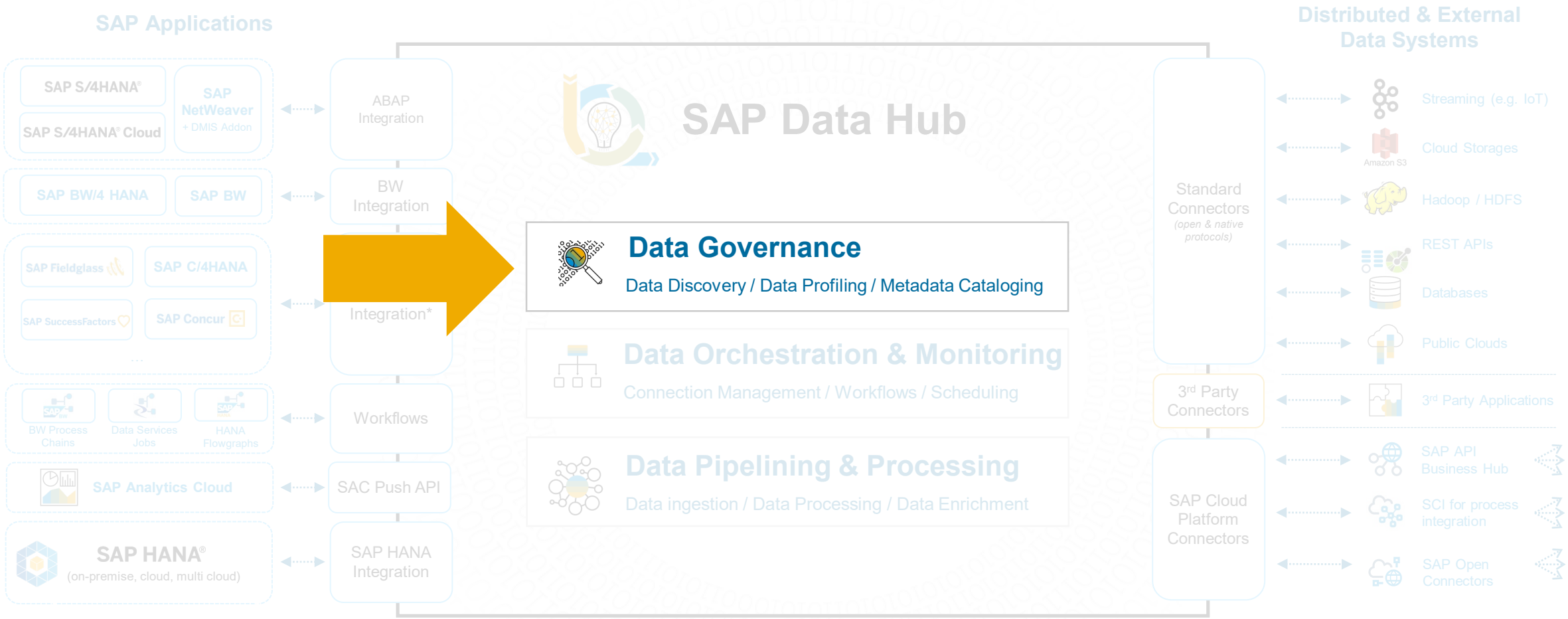
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

*This is the current state of planning and may be changed by SAP at any time without notice.

SAP Data Hub

Capabilities Overview

SAP Data Hub Capabilities: Data Governance



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Data Governance

Metadata Management

Build up catalog to get insight into your company's metadata

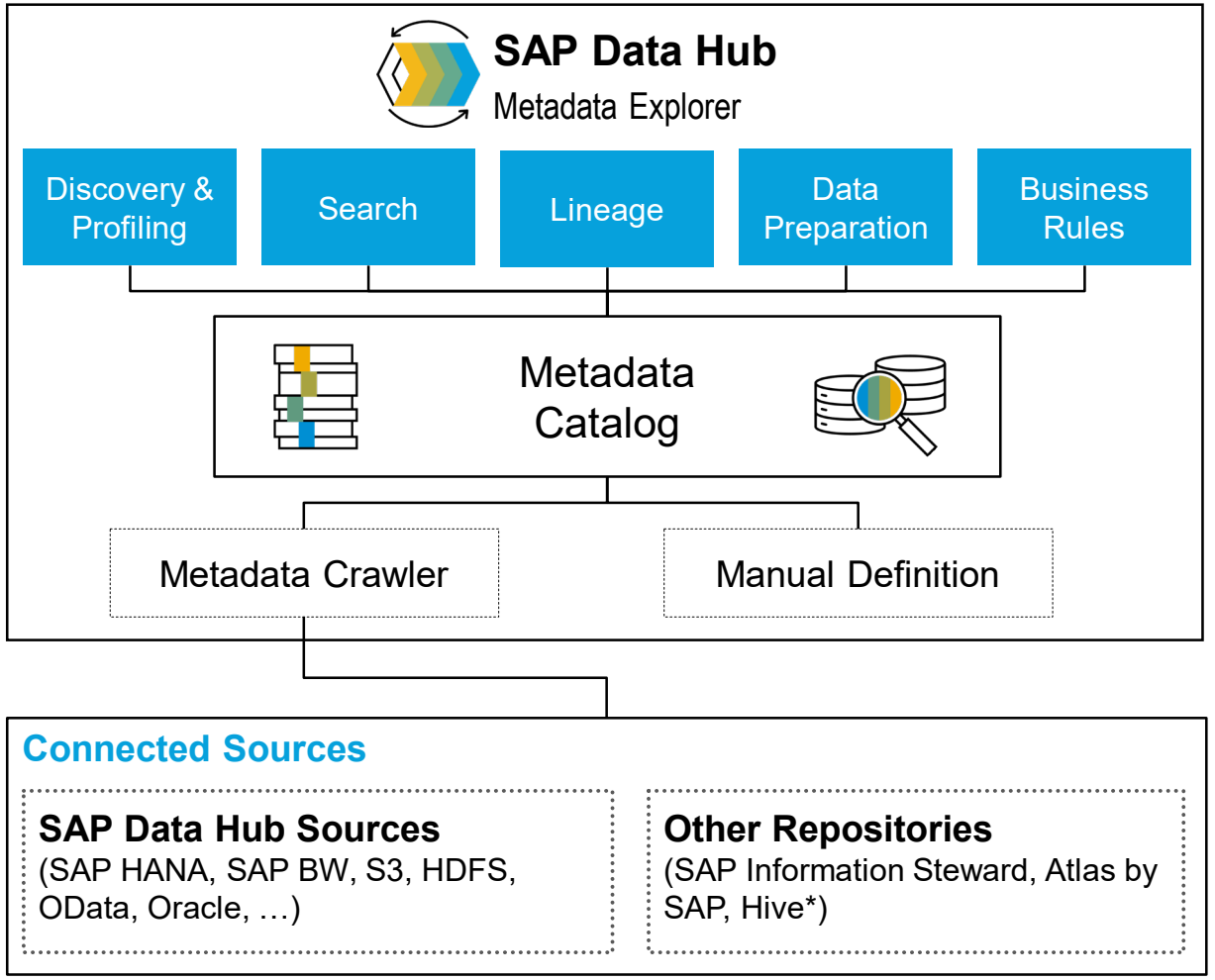
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Data Governance

Metadata Management using the Meta Data Explorer

DAT162

- **Explore** (Discover, Preview, Profile) data assets in connected systems
- System-wide **Metadata Catalog**
  - Metadata Crawler: Automatic indexing und publication of metadata
  - Manual Definition of metadata (datasets)
- **Classification** of data using labels (e.g. regarding quality and sensitivity)
- **Data Lineage** provides a clear end-to-end understanding of the data's origins and usage
- Define and run **Business Rules** to understand the data quality of your data
- **Data Preparation** allows business users to acquire, enrich, and enhance data in a spreadsheet like UI
- Tight integration with modeling environments for Data Pipelines and Workflows

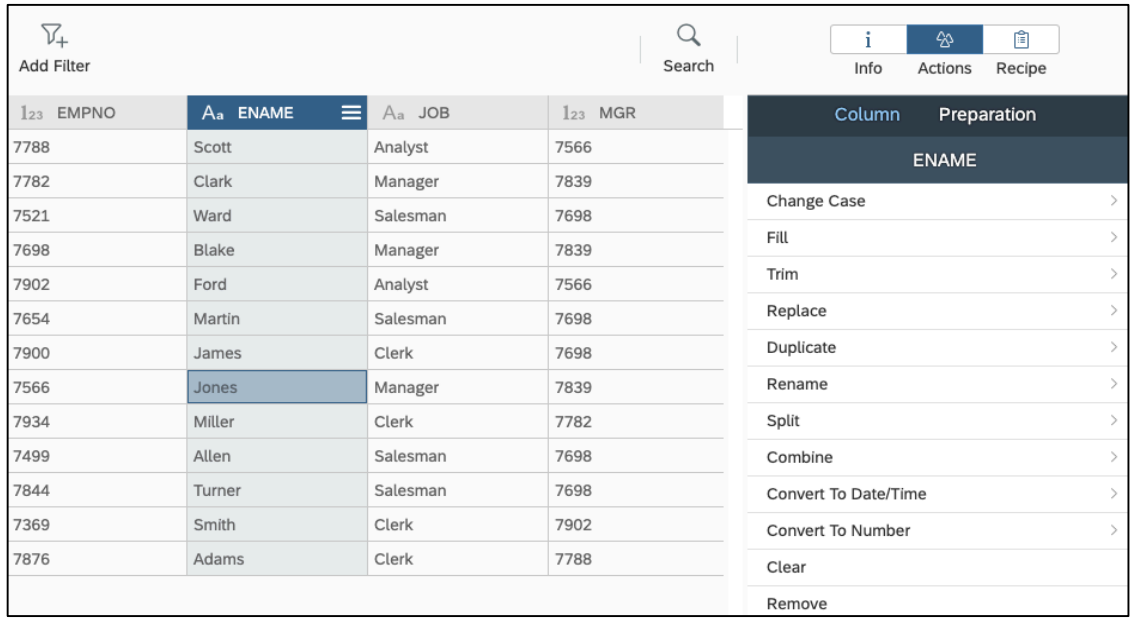


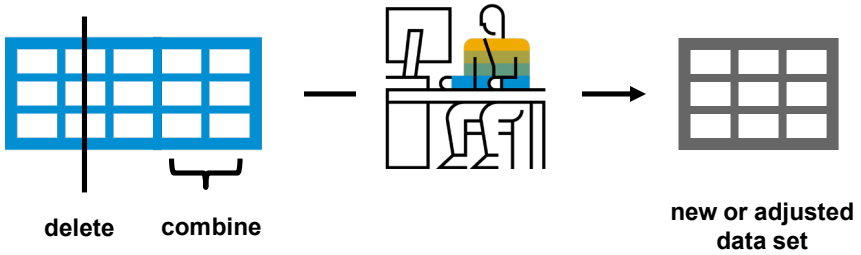
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

*This is the current state of planning and may be changed by SAP at any time without notice.

Data Governance

Data Preparation Self-service for Non-Technical Users





Capabilities

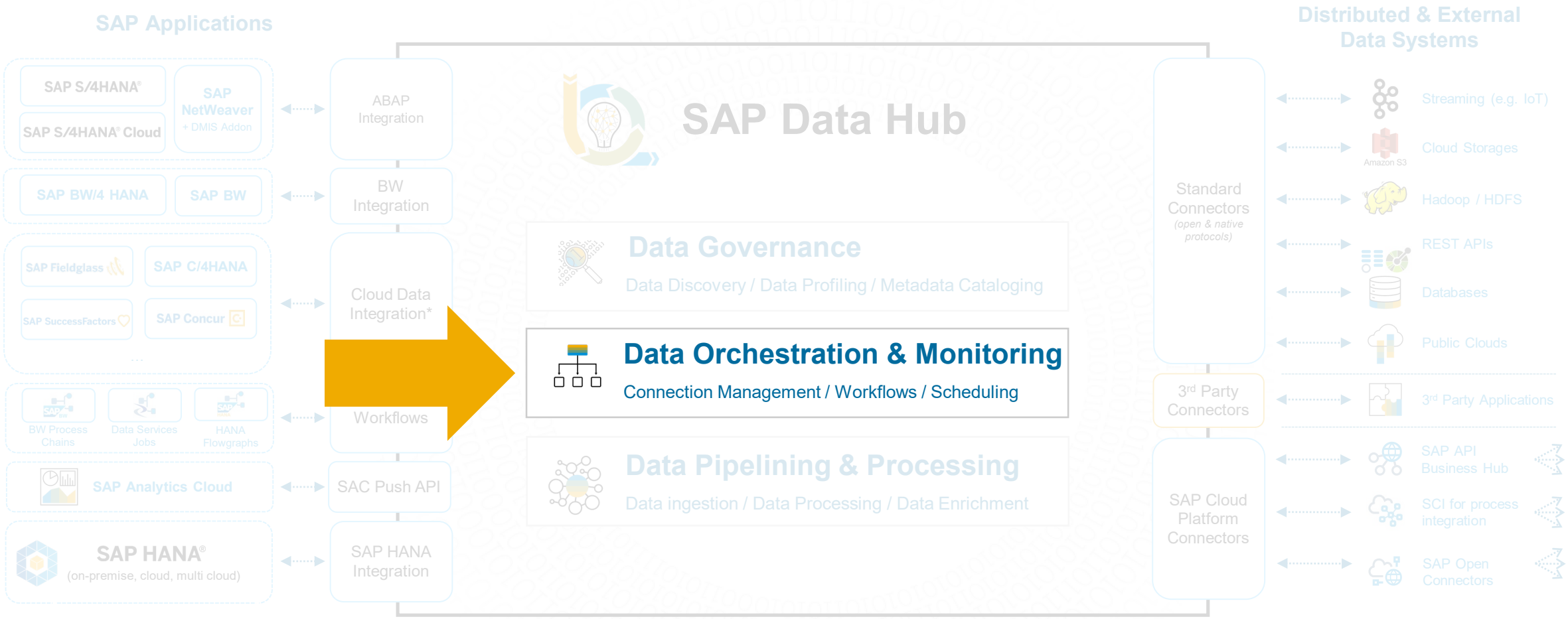
- Prepare the data based on a sample dataset
- Transform, shape, harmonize, curate, enrich the data via simple click actions
- View, present and report the outcome immediately
- Apply the transformations to the full dataset at any time

Main Use Cases

- End-to-end self-service data preparation
- Improve data quality to achieve data excellence
- Create new data sets based for scenario and project requirements

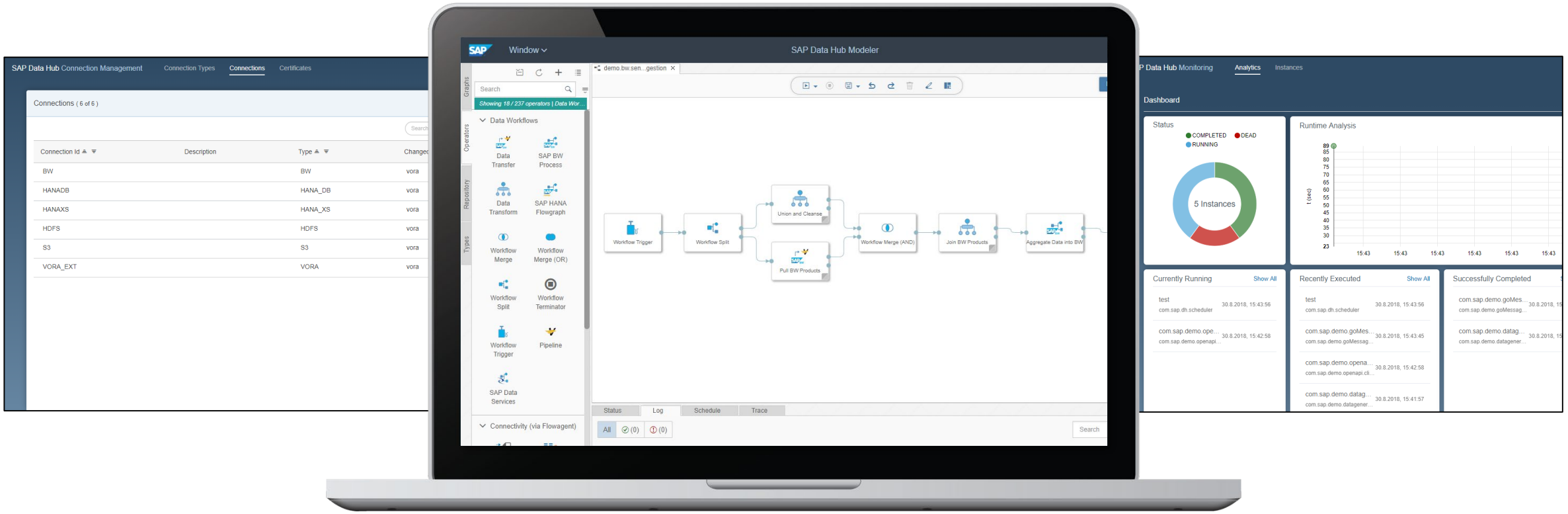
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub Capabilities: Data Orchestration & Monitoring



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Data Orchestration and Monitoring



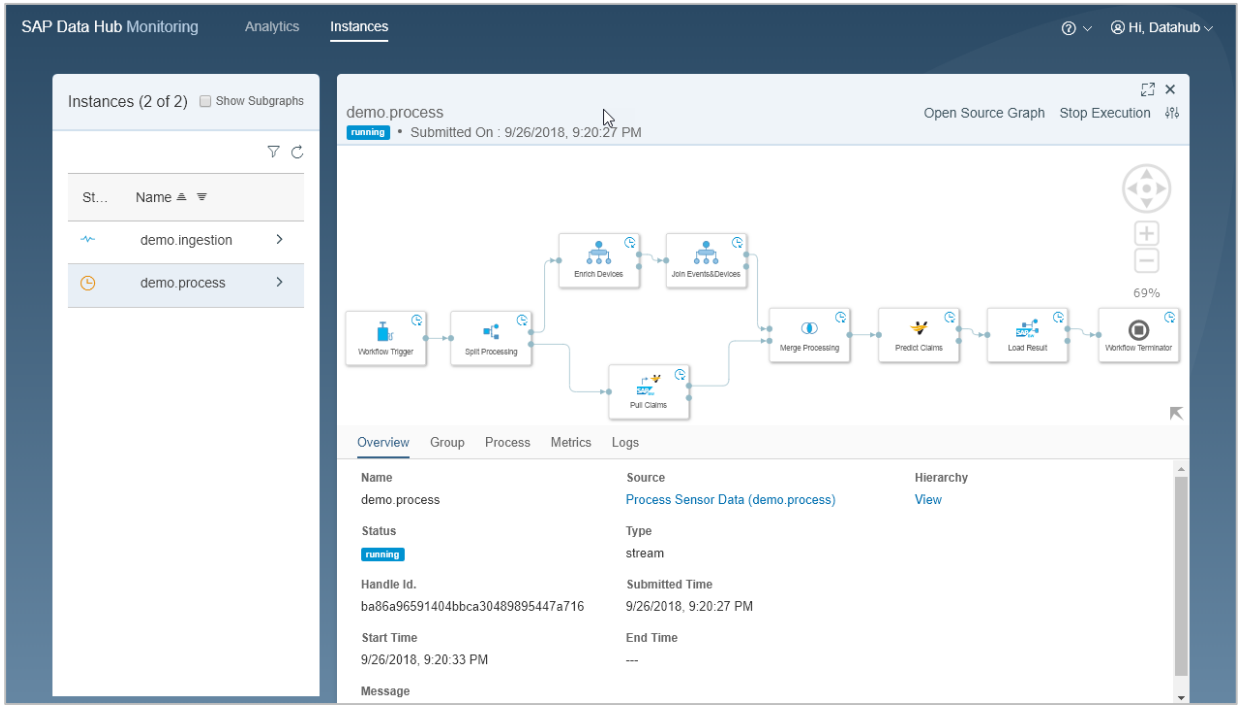
Connect, orchestrate and monitor processes across systems

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Data Orchestration and Monitoring

Connection Management, Data Workflows, Scheduling

- Central management of all system connections
  - Connect to diverse systems natively and remotely
  - Support user access management
- **Data Workflows** for cross-system data orchestration, e.g.
  - Trigger Execution of SAP BW process chains
  - Transfer data from SAP BW and SAP HANA
  - Execute remote SAP Data Services jobs
  - Trigger iFlows in SAP Cloud Platform Integration
  - Submit Spark jobs to Hadoop clusters
- **Scheduling** of Data Workflows
- Extensive **Monitoring**
  - Support view, control, and auditing of data operation in the connected data landscape

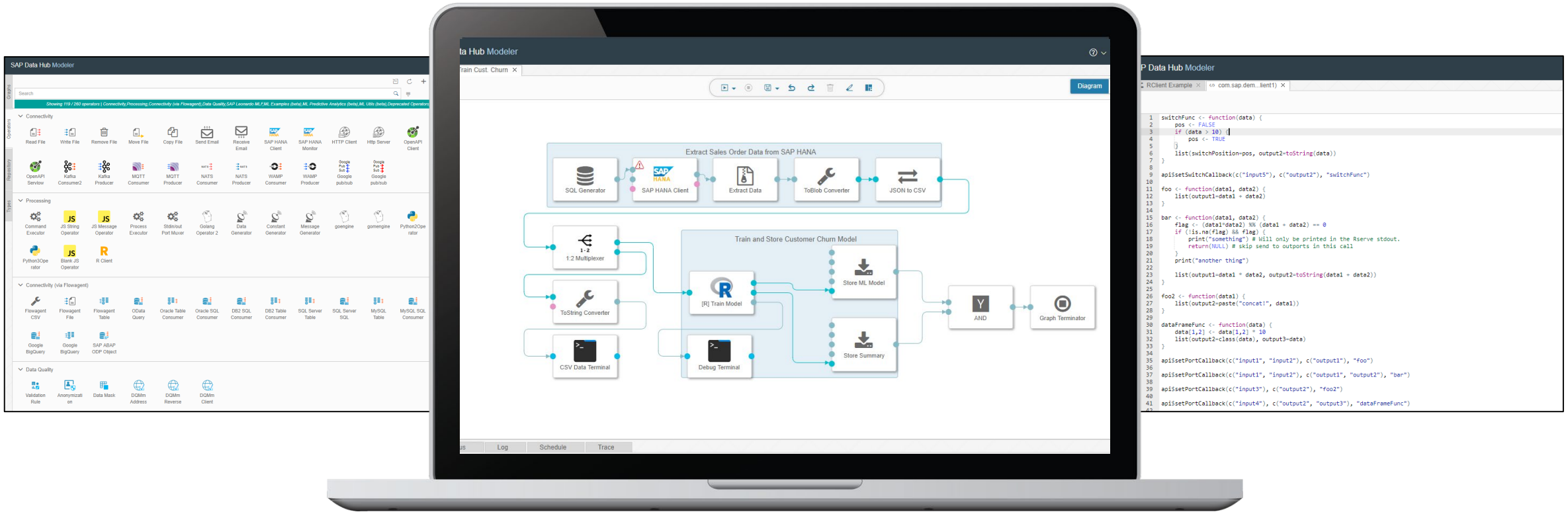


© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub Capabilities: Data Pipelining & Processing

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Data Pipelining & Processing



Build scalable and flexible flow-based applications to process, refine, and enrich data at the source.

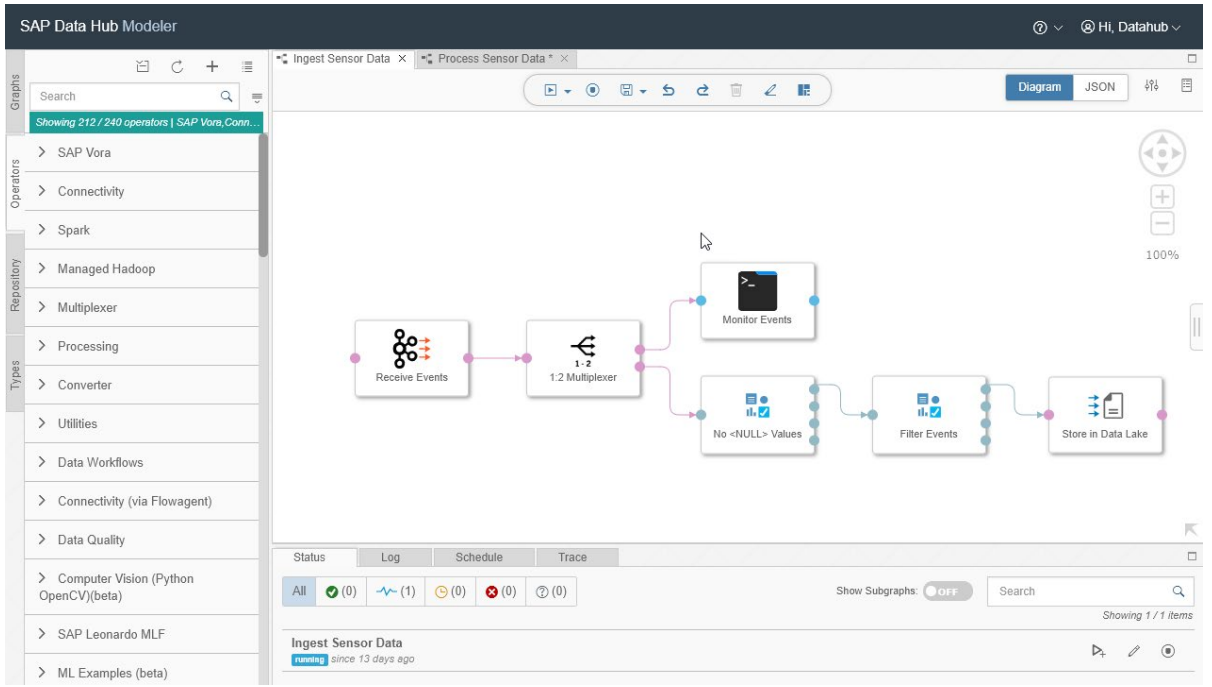
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Data Pipelining & Processing

Build Flow-based Applications using the Pipeline Modeler

DAT363

- **Data Pipelines** = Flow-based applications
  - Operators (independent computation units)
  - Data (messages) flows between operators
- **Extensible**
  - Over 250 pre-defined operators (Connectivity, Processing, Data Quality, CV, ML, etc.)
  - Custom / Partner operators
  - Wrap any custom code
- **Scalable**
  - Containerized – Docker containers constitute the operators' execution environments
  - Distributed – Easy horizontal scaling
- **Re-Usability**
  - Create complex, multistep, reusable data pipelines and operators

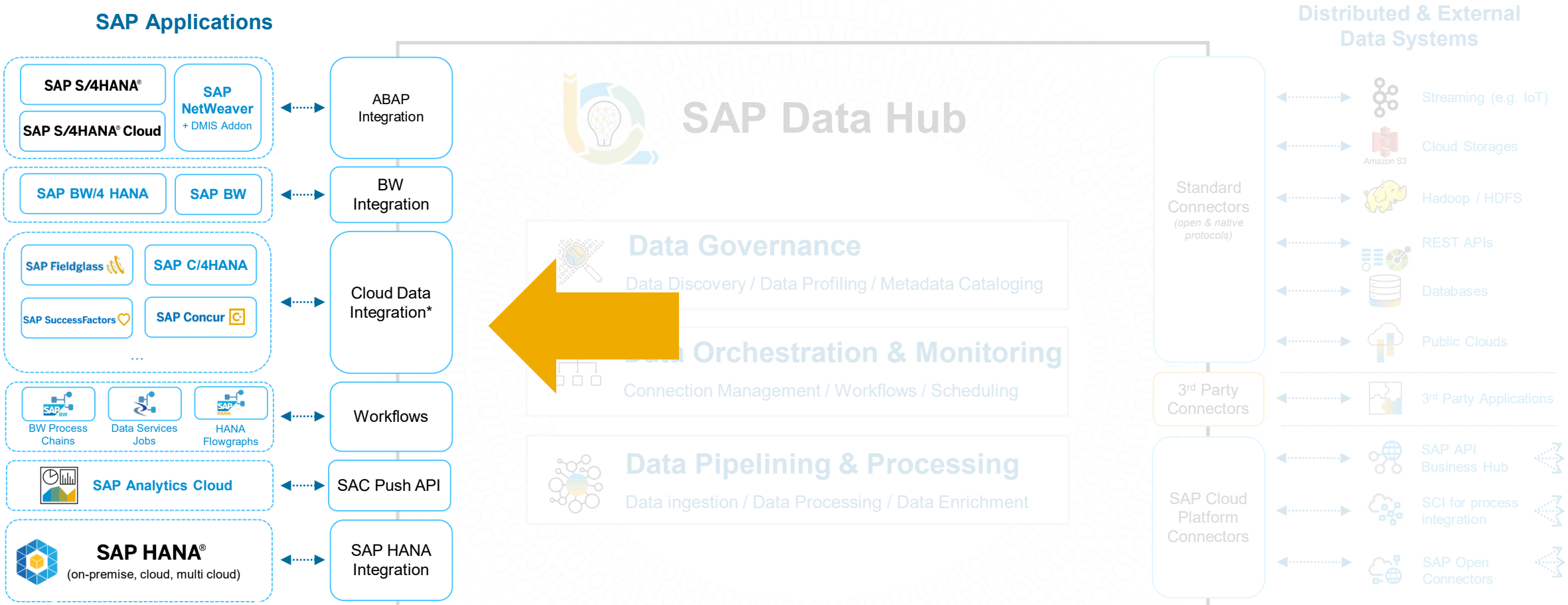


© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub

Data Integration with SAP Applications

SAP Data Hub: Data Integration with SAP Applications

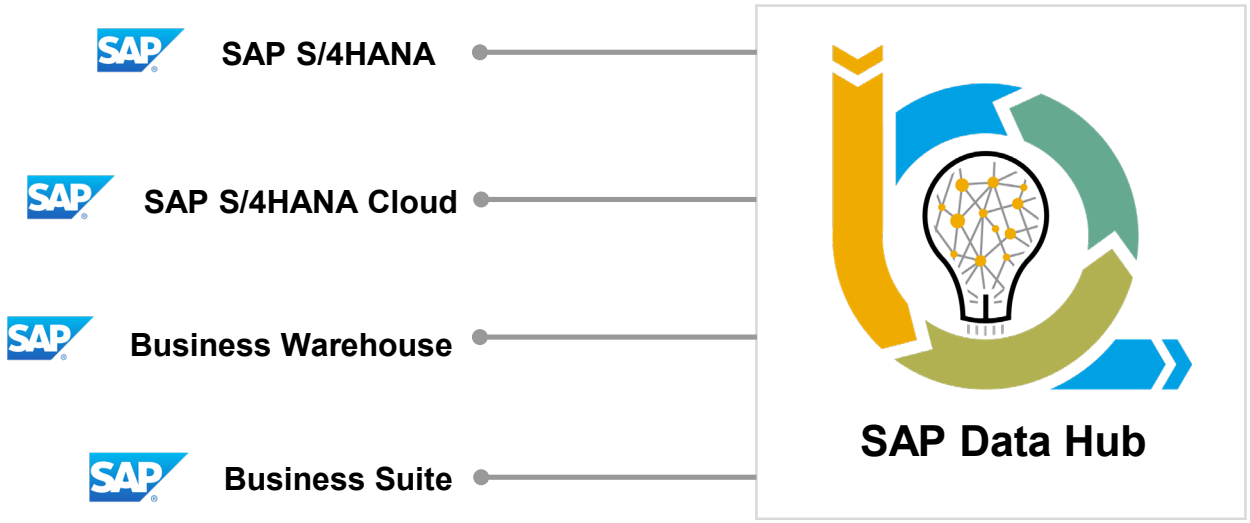


© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Integration with ABAP-based SAP Systems

DAT108

One model to consolidate all interaction scenarios between SAP Data Hub and an ABAP-based SAP systems directional and bi-directional



Capabilities

Provide **ABAP METADATA** to the SAP Data Hub Metadata Explorer

ABAP DATA PROVISIONING to transfer data into SAP Data Hub

ABAP FUNCTIONAL EXECUTION that is triggerable as a SAP Data Hub Operator

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

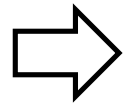
Integration requires certain system level, planned at least SAP S/4HANA 1909, SAP S/4HANA cloud 1911, SAP NetWeaver 7.00 with DMIS 2011/2018 Q4/2019 version. Certain functionality can only be made available for certain release levels.

Integration with ABAP-based SAP Systems

Good to know

ABAP CDS READER

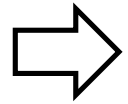
Operator that supports initial load and delta replication of ABAP CDS views.



ABAP CDS Views are only available in SAP S/4HANA

CUSTOM ABAP OPERATOR

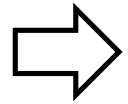
Operator to implement your ABAP custom code that will be executed as part of a pipeline in the connected ABAP system. For example to call a function module.



SAP S/4HANA cloud does not support custom code

SLT CONNECTOR

Operator to communicate with SAP LT Replication Server (SLT). Leverage existing replication scenarios to bring data into a SAP Data Hub pipelines.



Allows data consumption of lower SAP NetWeaver (<7.00) and S/4HANA (<1909) releases

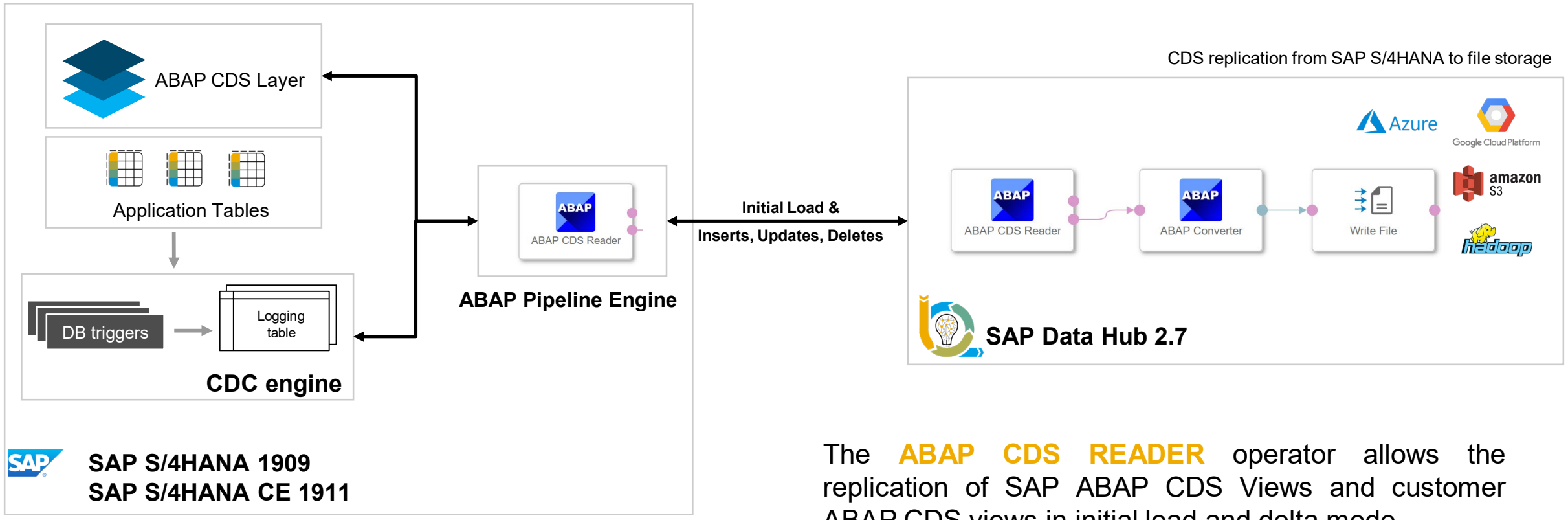
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Integration requires certain system level, planned at least SAP S/4HANA 1909, SAP S/4HANA cloud 1911, SAP NetWeaver 7.00 with DMIS 2011/2018 Q4/2019 version. Certain functionality can only be made available for certain release levels.

Integration with ABAP-based SAP Systems

Example Data provisioning for SAP S/4HANA with ABAP CDS Reader

ABAP CDS is the semantically rich data model in SAP S/4 HANA and allows the consistent representation of a business object like a Business Partner

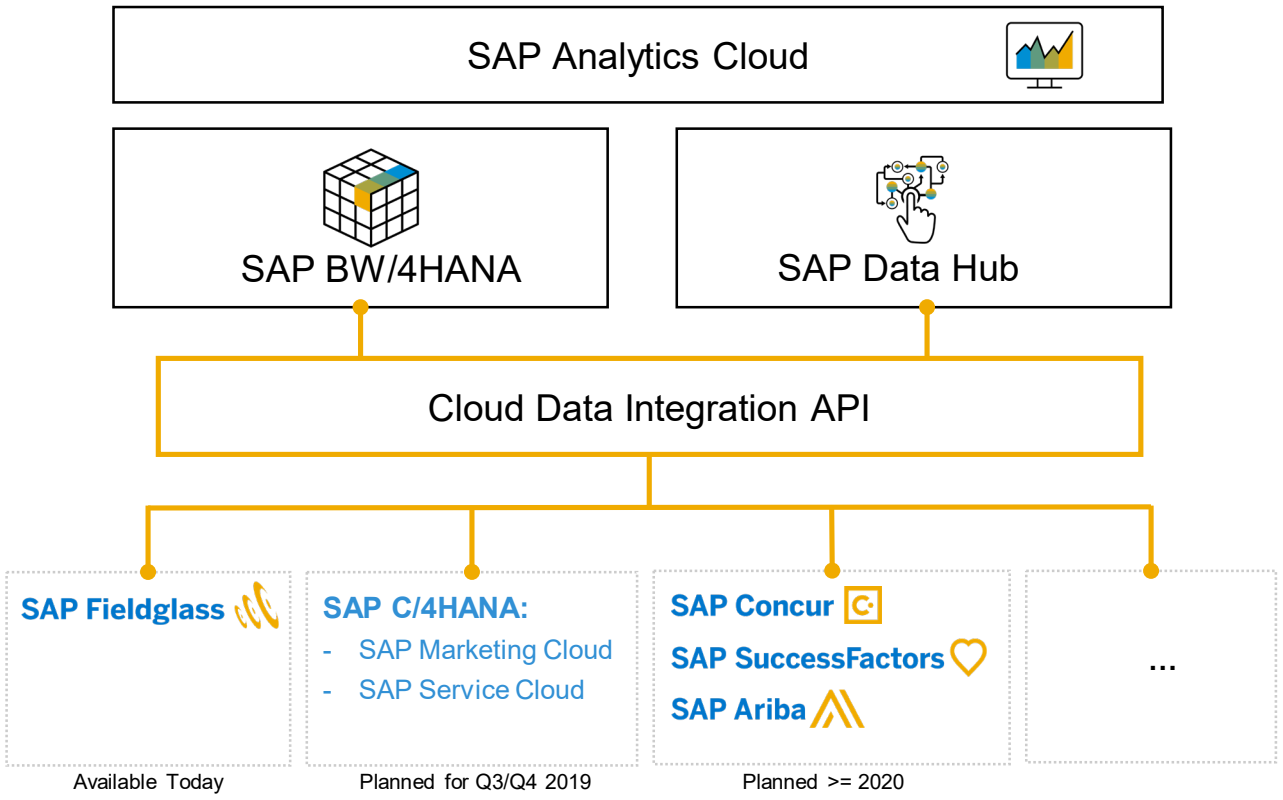


The **ABAP CDS READER** operator allows the replication of SAP ABAP CDS Views and customer ABAP CDS views in initial load and delta mode.

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Integration with SAP Cloud solutions

One SAP Cloud Data Integration (CDI) for integration into all SAP Cloud solutions



Main Use Cases

- Holistic Data Management with SAP Data Hub
- Building Data Warehouse Analytics with SAP BW/4HANA
- Advanced Analytics and Planning with SAP Analytics Cloud

Capabilities

- Support both full and delta requests
- Seamless integration of data and metadata
- OData v4 as communication protocol

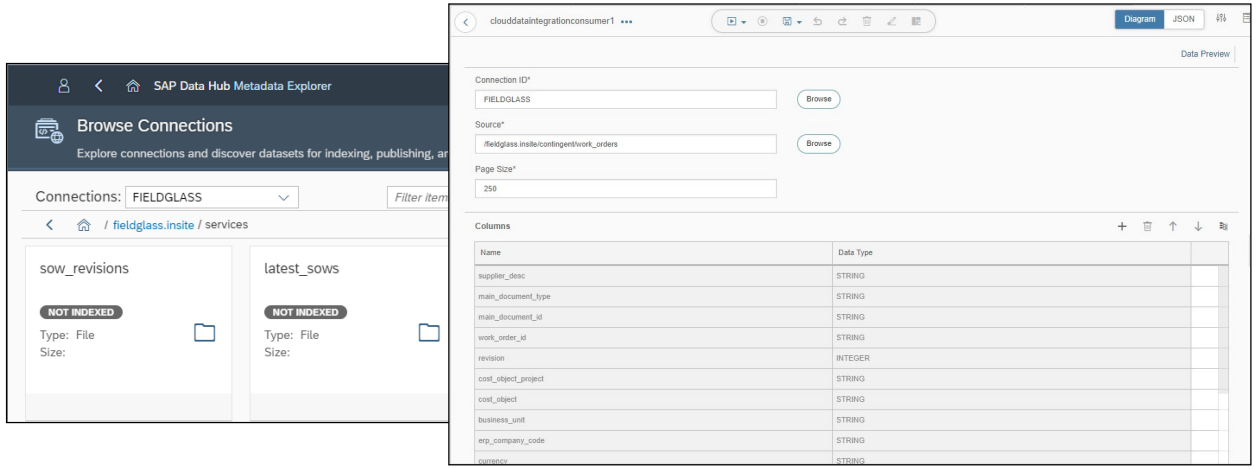
SAP Cloud Data Integration (CDI) is a core concept for integrating all SAP solutions in the cloud using **ONE API** based on open standards

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

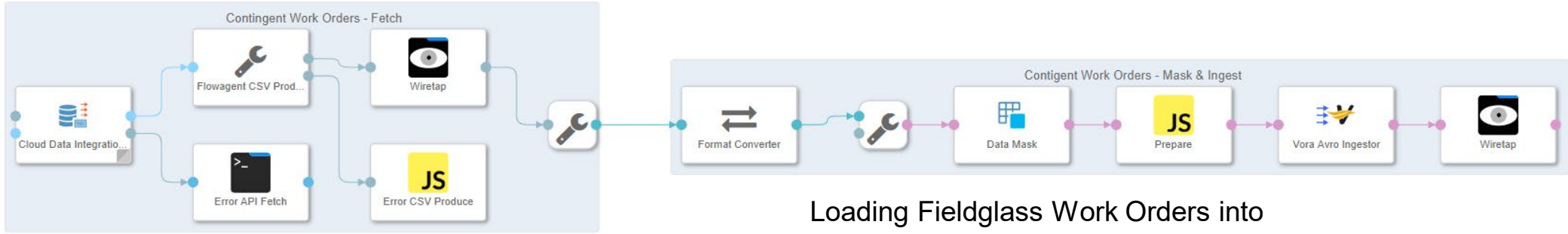
Integration with SAP Cloud solutions

Example Cloud Data Integration scenario with SAP Fieldglass & SAP Data Hub

Understanding metadata of SAP Fieldglass within the SAP Data Hub Metadata Explorer



SAP Fieldglass



Loading Fieldglass Work Orders into SAP Vora tables with data protection

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

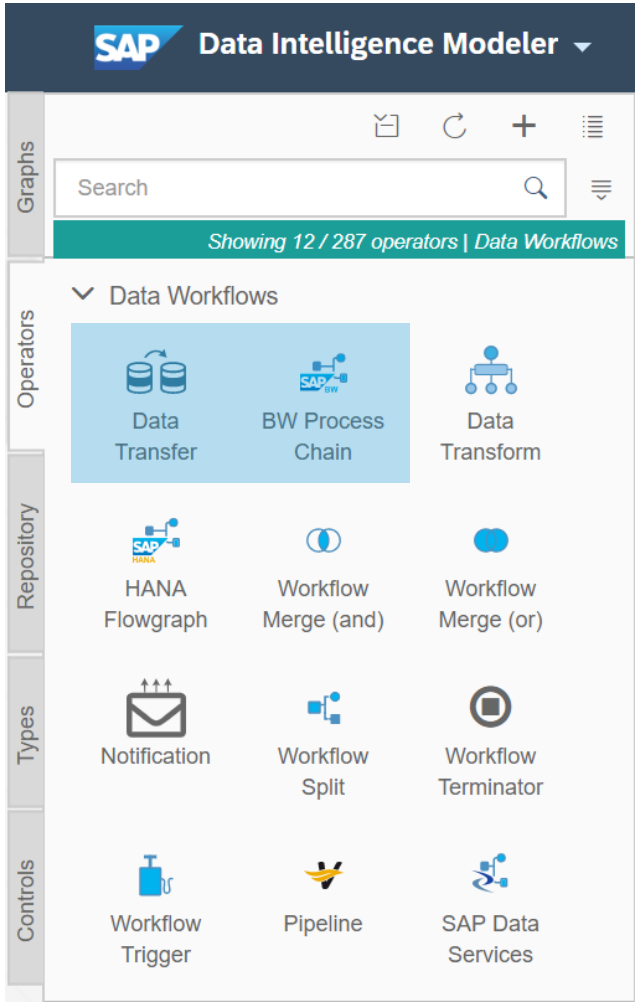
Integration with SAP Business Warehouse

Available Pipeline Operators:

- SAP BW Process Chain
  - Trigger execution of a process chain on a SAP BW system
- Data Transfer (SAP BW & SAP HANA)
  - Transfer data (query, infoprovider / tables, views) from SAP BW / SAP HANA into big data stores or SAP Vora tables
  - Via INA interface or direct access to HANA (Calculation Views)

Typical Scenarios:

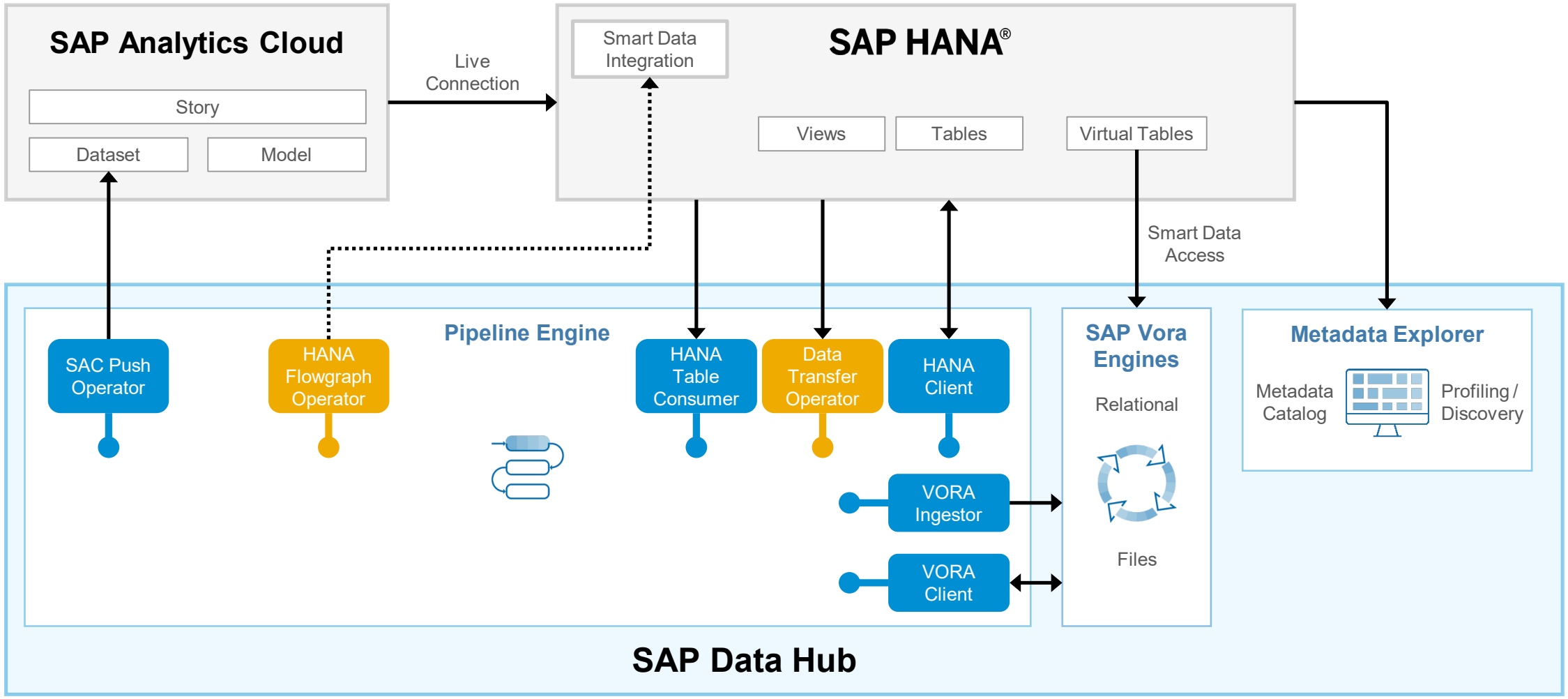
- Load Data from Data Lake into BW
- Data Tiering from BW into Data Lake



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Integration with SAP HANA and SAP Analytics Cloud

SAP Data Hub integrates with SAP HANA and SAP Analytics Cloud on multiple levels



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub

Data Integration with Non-SAP Applications

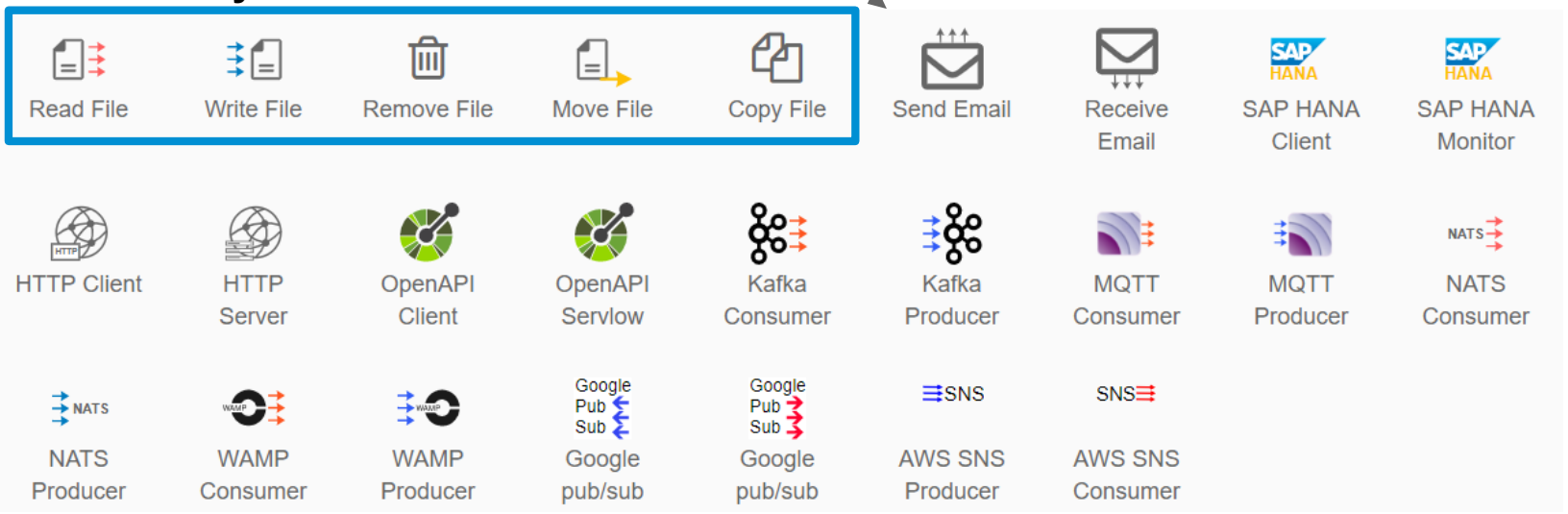
SAP Data Hub: Non-SAP Data Integration

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Non-SAP Data Integration

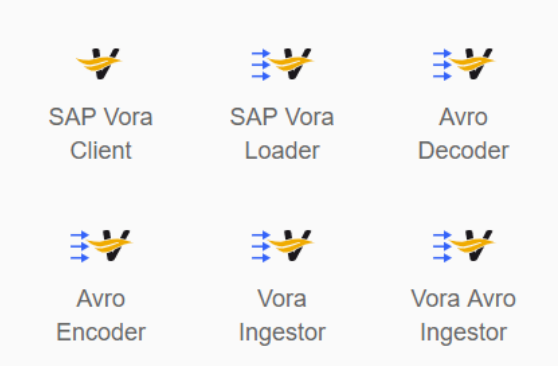
Built-in Standard Connectors

Connectivity:

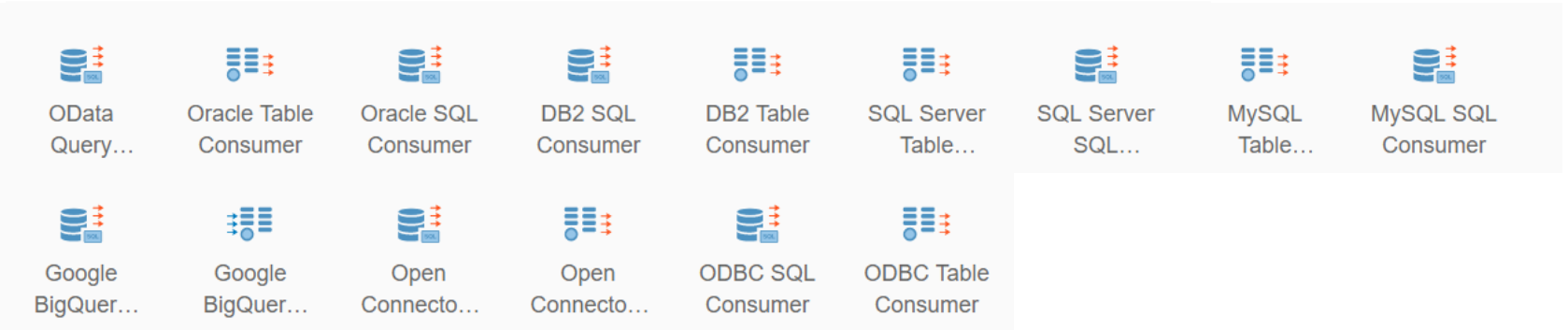


- Azure Data Lake (ADL)
- Google Cloud Storage (GCS)
- HDFS
- Amazon S3
- Azure Storage Blob (WASB)
- Local File System (file)
- SAP Semantic Data Lake
- WebHDFS

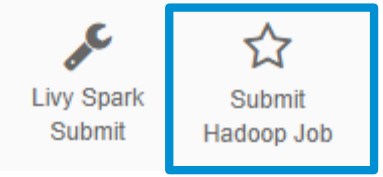
SAP Vora:



Connectivity (via Flowagent):



Spark / Hadoop:



- Spark
- Spark SQL
- PySpark
- Hive

Data Quality:

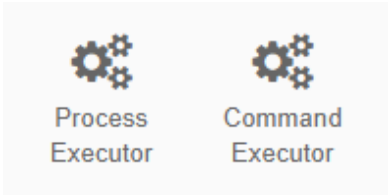
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Non-SAP Data Integration

Operators for Data Processing

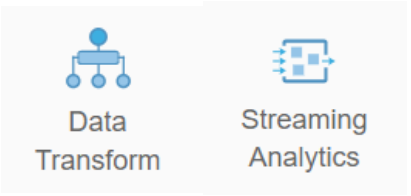
Process / Command Executors:

- Run a process within a pipeline and give contiguous stream to it
- Run a shell command for each arrival of a message within a pipeline



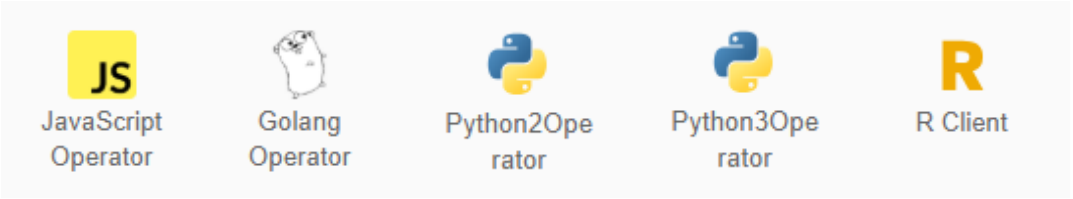
Transformation Operators:

- Run on-the fly transformations and do event stream processing using continous query language (CQL) on data within a pipeline



Scripting Operators:

- Write and run custom scripts for data manipulation within a pipeline
- Build re-usable operators in different programming languages



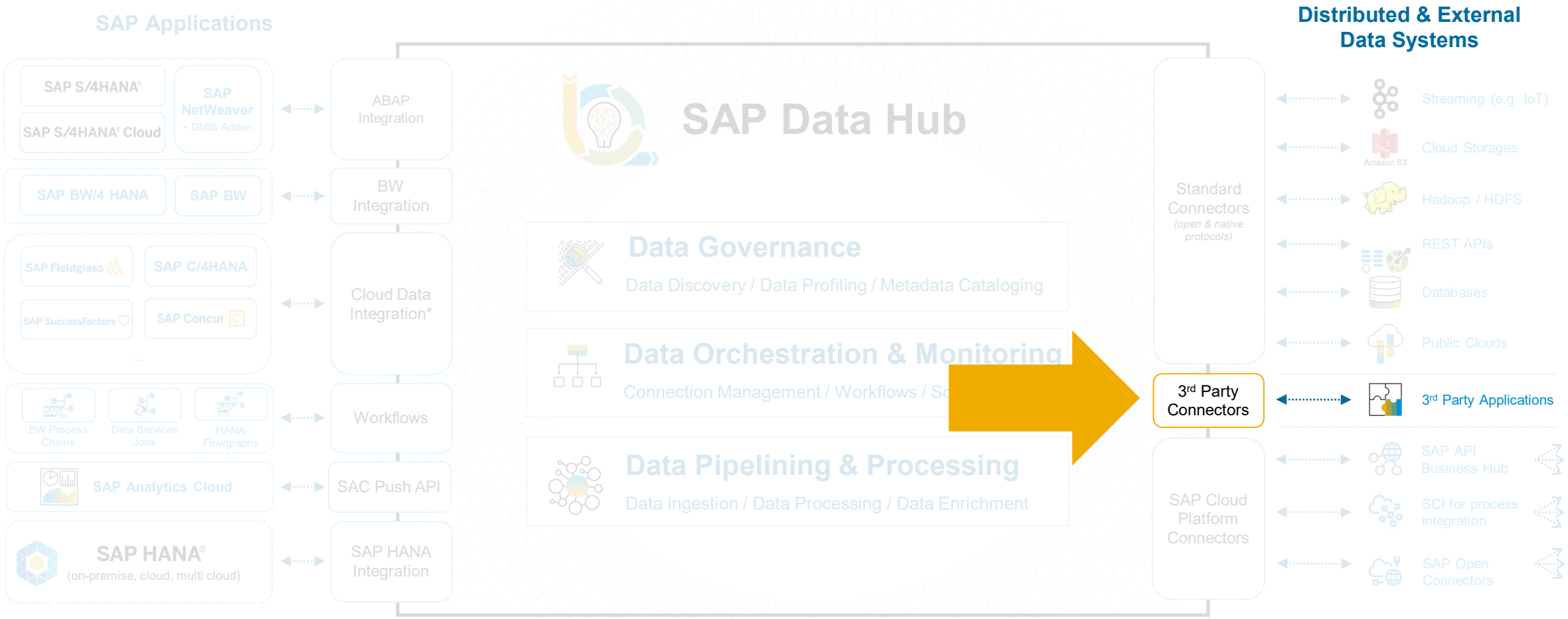
Subengines:

- Develop and compile new operators locally using SDK
- Register and run custom operators in available pipeline subengine

*This is the current state of planning and may be changed by SAP at any time without notice.

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

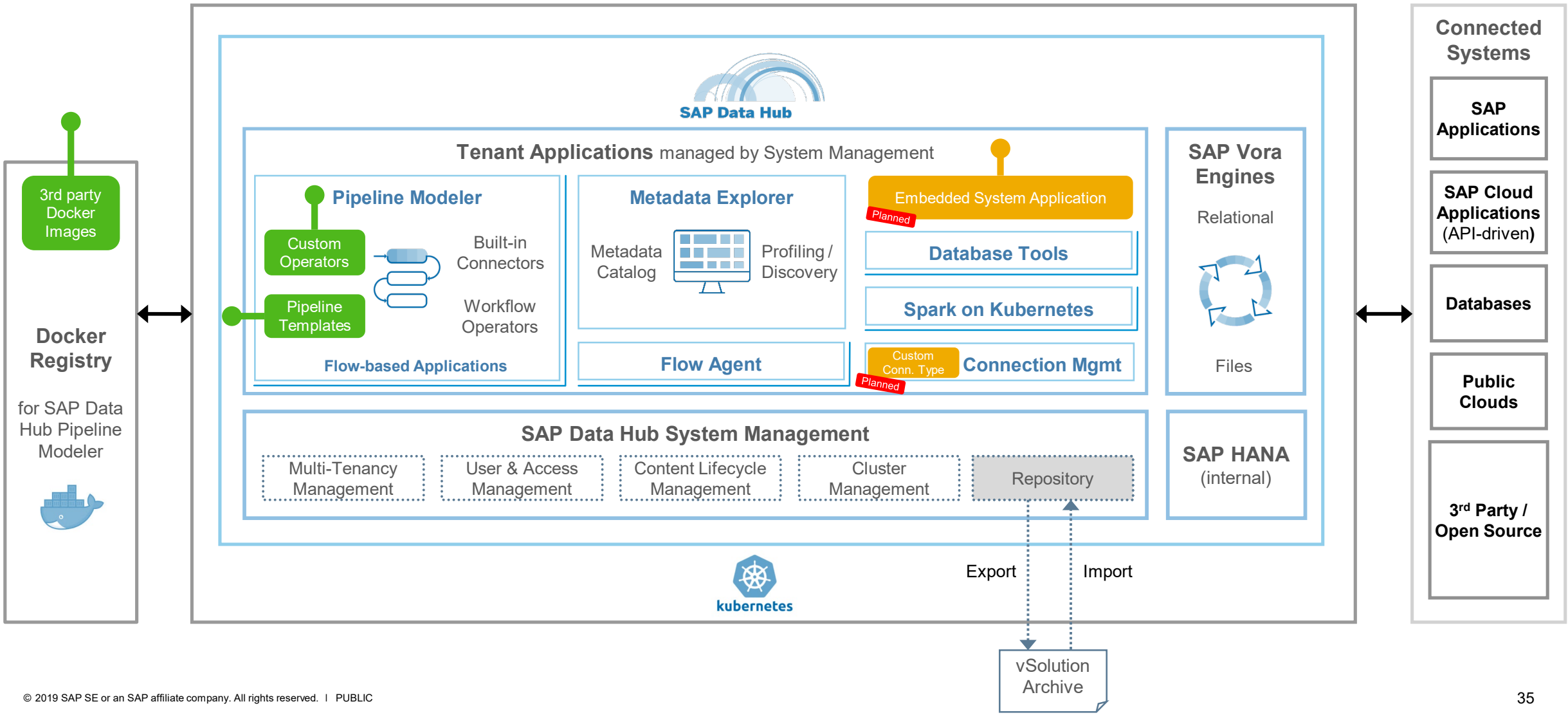
SAP Data Hub: Extensibility for 3rd Party Connectivity



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Extensibility for 3rd Party Connectivity

SAP Data Hub Extension Points



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Connecting data sources with SAP Data Hub

Three ways of connecting systems, engines and repositories

01 Managed system connectors

Available in the SAP Data Hub launchpad, these are the richest and most integrated connectors, providing metadata and governance services, profiling, source crawlers, etc.

- Includes connectors to ETL tools like SAP Data Services and SAP HANA SDI, where SAP Data Hub leverages brokered connections through an intermediary, allowing connection with many additional systems.
- All currently supported connectors can be found here.

02 Standard pipeline operators

Additional systems can be accessed within the pipeline modeler via specific operators, e.g. Kafka.

- All currently supported standard operators can be found here.

03 Custom operators

Custom operators can be created based on SAP provided base operators like JavaScript, Python, Spark, RestAPI, etc. Those are very powerful, and allow integration with a wide range of systems, applications, and data stores. They require an engineering skillset.

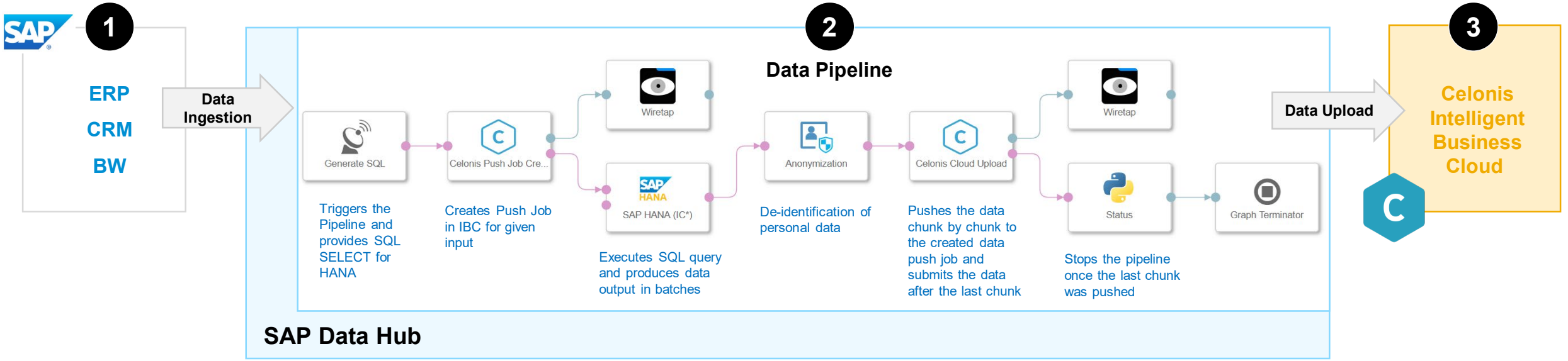
- How to build custom operators is described in this blog series.

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Extensibility for 3rd Party Connectivity

Example with Celonis Intelligent Business Cloud

- **SAP internal Lead-to-Cash Scenario:** Understand internal process inefficiencies all the way from demand management to maintaining a deal through process mining in Celonis leveraging SAP Data Hub
- Involved Steps
  1. Collect purchase related activities in ERP system (SAP S/4HANA)
  2. Anonymize personal data, transform activities into required event log structure to do process mining and upload event log to Celonis Intelligent Business Cloud (SAP Data Hub + 3rd Party Adapter from Celonis)
  3. Run Process Mining (Celonis Intelligent Business Cloud)



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Extensibility for 3rd Party Connectivity

Open for Partner and Customer Content Delivery

Data Hub‘s Extensibility Options

Partners and customers can add own content based on different layers:

- **Pipelines Templates:**
  - Parameterizable scenario templates
  - (Sub)graphs called in other graphs
- **Operators:**
  - Extend pre-defined base operators using powerful runtimes, e.g. Python, Go, Javascript, and R
  - Develop new operators with own runtimes, e.g. Process Executor
- **Dockerfiles:**
  - Provide execution environments for operators
  - Incorporate 3rd party libraries and executables

Data Hub Content Delivery

- Custom Data Hub Content can be packed into a **SAP Data Hub Solution** Archive
- SAP Data Hub Solutions can be defined, exported and imported via
  - SAP Data Hub System Management UI
  - vCTL Command Line Tool*
- SAP Data Hub Solution Structure:

```
vsolution
├── vrep
│   └── vflow
│       ├── dockerfiles
│       ├── graphs
│       └── operators
│           └── path
│               └── to
│                   └──package
│                       └── operator.json
└── vsolution.json
```

SAP Data Hub Content Marketplace:

- Planned Marketplace and Delivery channel for SAP Data Hub Partner Content via **SAP API Business Hub** in combination with **SAP App Center** for monetization

A win-win partnership:

- Partners can reach more customers with their services, while SAP Datahub Hub expands its reach with relevant content
- Our partners understand the customer use cases and can better support with domain-specific content by generalizing the specific issue with a broad solution

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub Content Delivery

Publishing Content via SAP API Business Hub

SAP API Business Hub: Search, discover and consume content packages with pre-packaged integrations and sample applications from SAP and partners

List

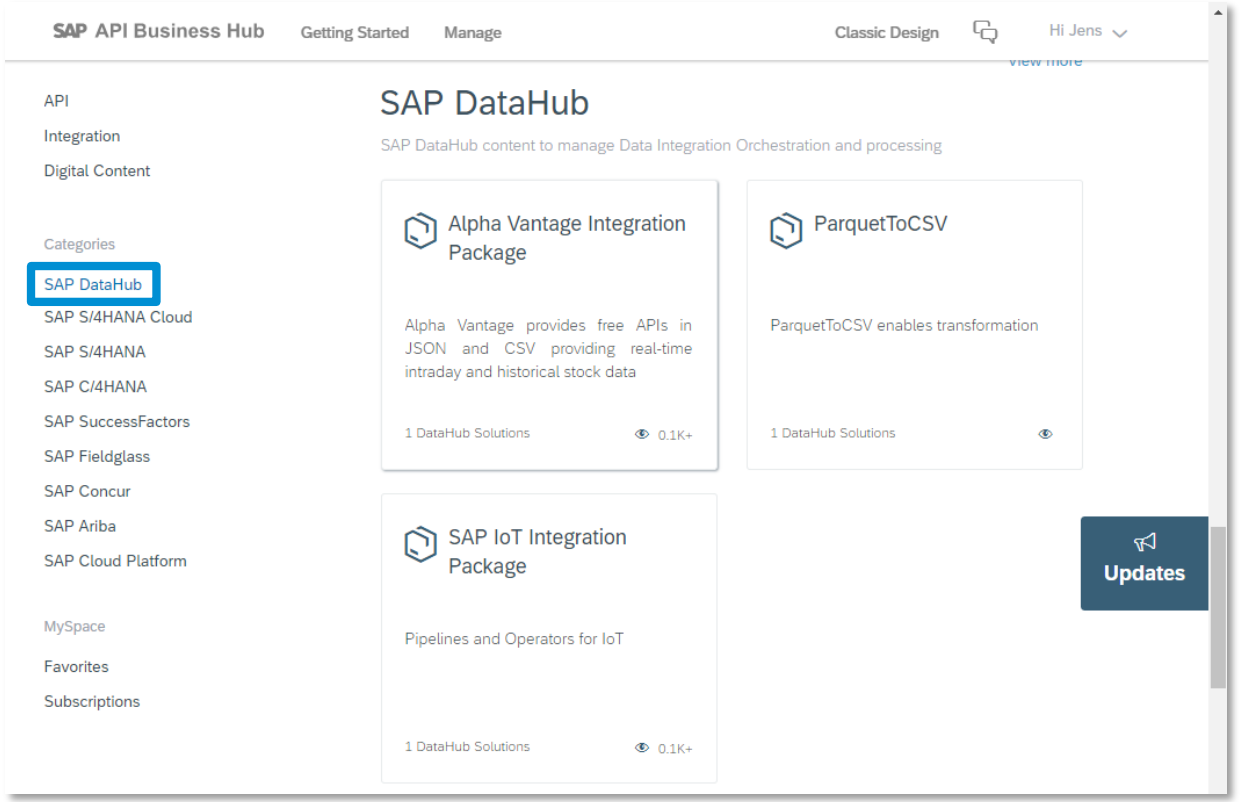
Digital Content (Integration)

Engage

With your developer ecosystem

Monetize

Via App Center, or by OEM-ing with SAP



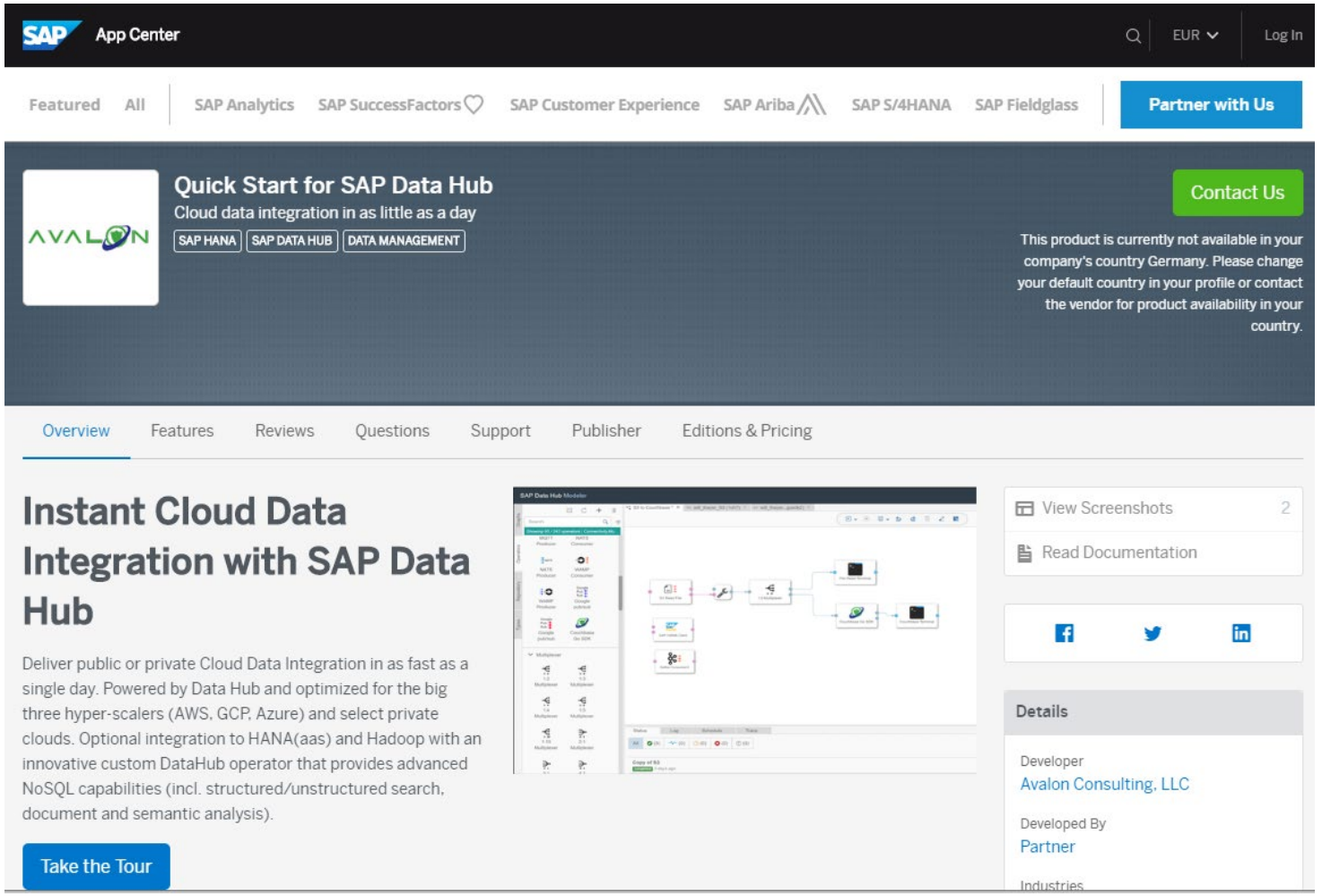
© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub Content Delivery

Publishing Content via SAP App Center

Publish SAP Data Hub content to be monetized on **SAP App Center** in 3 easy steps:

- Become an SAP Partner
- Submit for SAP Data Hub Readiness Check
- Publish and begin Transacting



© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

SAP Data Hub

Integration with SAP Cloud Platform

SAP Data Hub: Integration with SAP Cloud Platform

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Integration with SAP Cloud Platform

Orchestration of SAP Cloud Platform integration to interact with processes

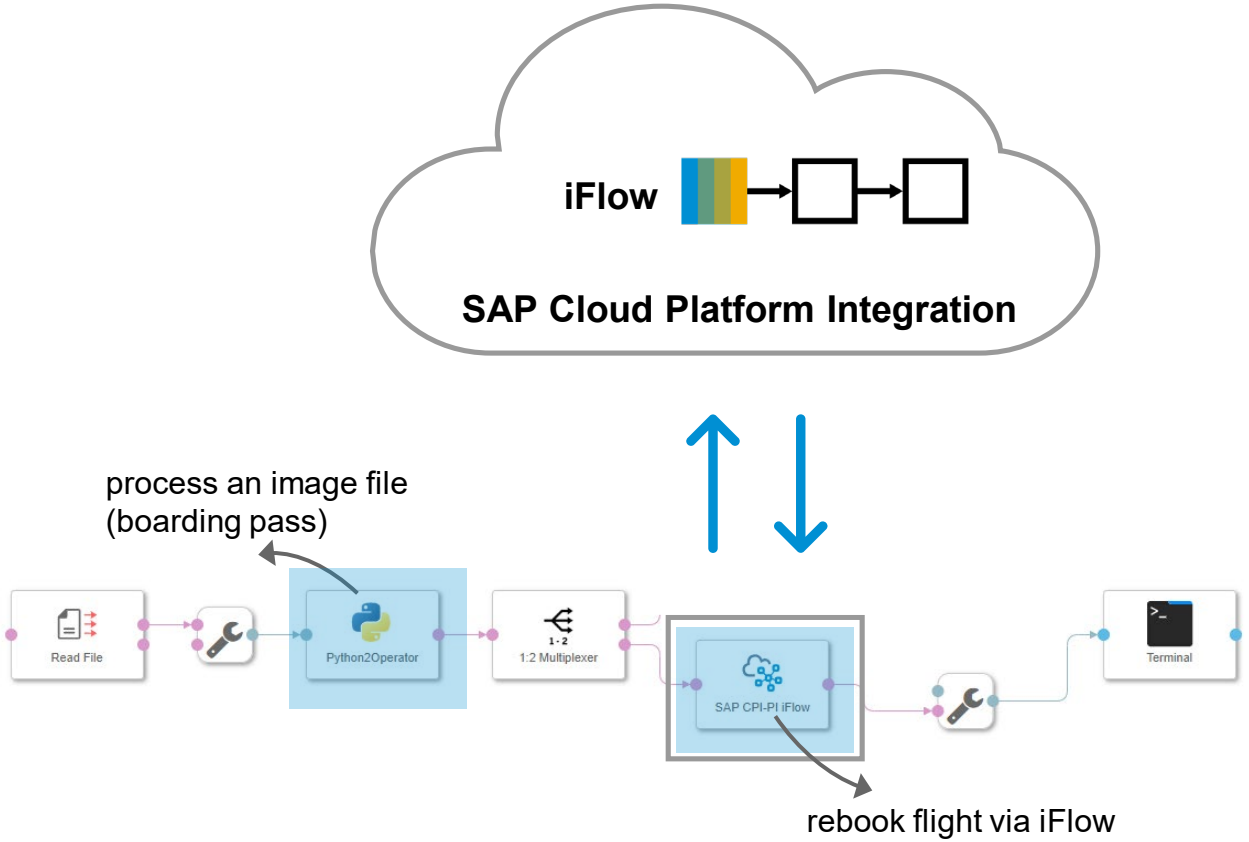
Enrichment of existing SAP Cloud Platform Integration scenarios with SAP Data Hub Modeler functionality

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

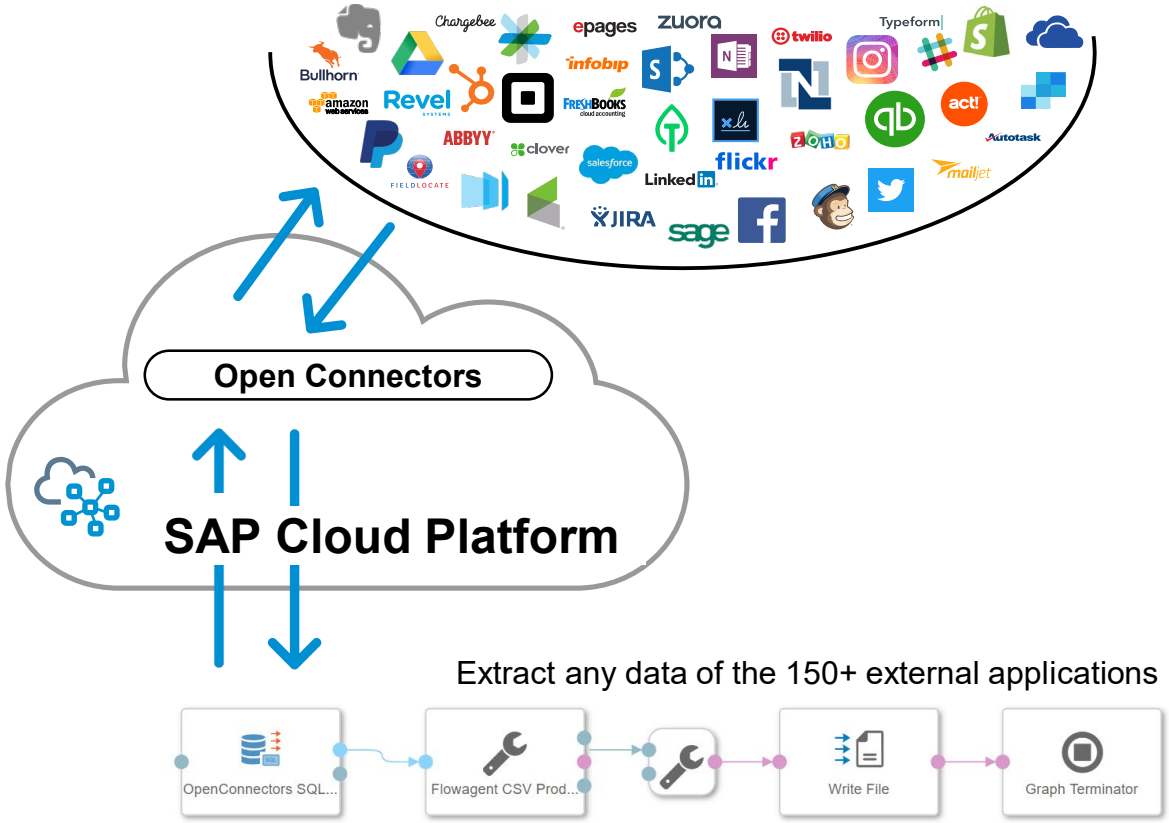


© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Integration with SAP Cloud Platform

Leveraging SAP Cloud Platform Open Connectors to consume external sources

Accelerate integration with external applications via SAP CP Open Connectors



Main Use Cases

- Seamless integration to over 150 non-SAP cloud application
- Embed external application and data into your process to broaden the scenario
- Develop and map canonical data model to extend pre-build connectors

Capabilities

- Easy connection through one API connector to reduce cost of third-party integration
- Normalized authentication, error handling, search, pagination, bulk support
- Support metadata extraction and data preview

Find all supported connectors here

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Demo

End-2-End Demo of SAP Data Hub Integration with SAP and non-SAP leveraging SAP Cloud Platform connectivity and adding value using SAP Data Hub’s processing and machine learning capabilities.

Continue your SAP TechEd 2019 Learning Experience

Join the digital SAP TechEd Learning Room 2019 in SAP Learning Hub

- Access SAP TechEd **Learning Journeys**
- Discover **related** learning content
- Watch **webinars** of SAP TechEd lectures
- Learn about SAP's latest innovations with **openSAP**
- Collaborate with **SAP experts**
- **Self-test** your knowledge
- Earn a SAP TechEd **knowledge badge**

SAP Learning Hub

openSAP

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

Engage with the SAP TechEd Community

Access replays and continue your SAP TechEd discussion after the event within the **SAP Community**

Access replays

- Keynotes
- Live interviews
- Select lecture sessions

http://sapteched.com/online

Continue the conversation

- Read and reply to blog posts
- Ask questions
- Join discussions

sap.com/community

Check out the latest blogs

- See all SAP TechEd blog posts
- Learn from peers and experts

SAP TechEd blog posts

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC

More information

Related SAP TechEd Learning Journeys

- AIN2 – Transform your business processes with intelligent technologies
- DAT3 – Tame data challenges

Related SAP TechEd sessions

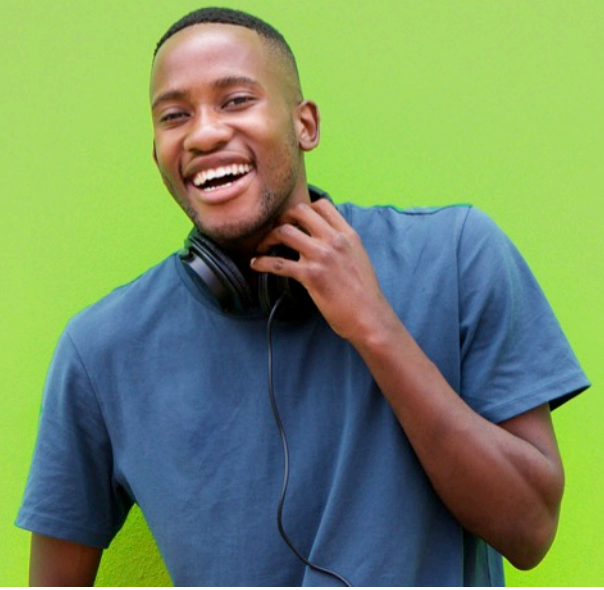
- DAT302 – Introduction to SAP Data Intelligence and SAP Data Hub
- DAT103 – The Data Management Strategy from SAP
- DAT363 – Create and Manage End-to-End Data Pipelines with SAP Data Hub
- DAT162 – Metadata Exploration and Self-Service Data Preparation with SAP Data Hub
- DAT108 – Integrating SAP S/4HANA into SAP Data Hub: Overview and Use Cases

---

Public SAP Web sites

- SAP Community: www.sap.com/community
- SAP products: www.sap.com/products

© 2019 SAP SE or an SAP affiliate company. All rights reserved. I PUBLIC



Thanks for attending this session.

Feedback

Please complete your session evaluation for DAT202.

Contact for further topic inquiries

SAP Data Hub
Product Management
datahub@sap.com

SAP TECHED

Follow us

www.sap.com/contactsap

© 2019 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platforms, directions, and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies.

See www.sap.com/copyright for additional trademark information and notices.

THE BEST RUN SAP