



SAP Data Intelligence - Turning disparate data into valuable assets for business innovation

Silvio Arcangeli,
Product Management, SAP Data Intelligence
November, 2020
PUBLIC

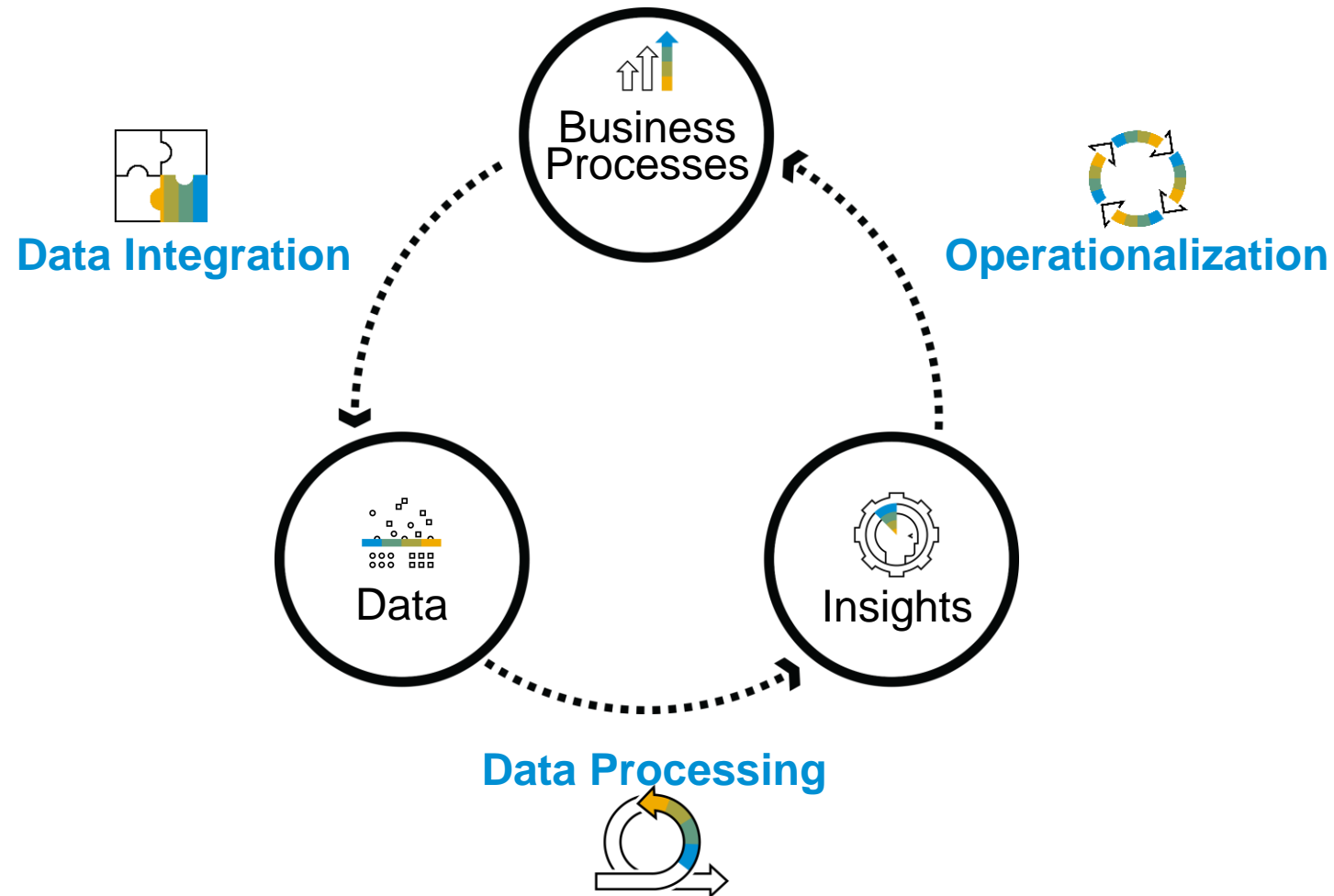
Disclaimer

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.

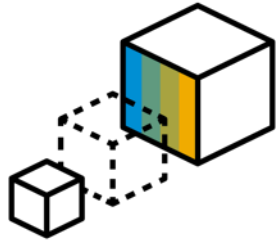
This presentation, or any related document and SAP's strategy and possible future developments, products and or platforms directions and functionality are all subject to change and may be changed by SAP at any time for any reason without notice. The information in this presentation is not a commitment, 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.

The Data-Driven Journey Towards the Intelligent Enterprise



But most enterprises get stuck in the journey!



85%

Innovation initiatives fail to make it to production¹



53%

Enterprises are not yet treating data as a business asset²



82%

Recognize that orchestrating data and processes is critical³

¹ [Gartner research](#)

² [NewVantage Partners' 2019 Big Data and AI Executive Survey](#)

³ [Forrester Consulting, 2019](#)

Why is Data Management so difficult?

Enterprise IT is challenged to a whole new degree!



Integration

“I need to join my sensor data with some image bitmap and with the related master data”



Processing

“I’m clustering the data in Python to feed a predictive model in HANA”



Discovery

“What is this data and what does it mean?”



Quality

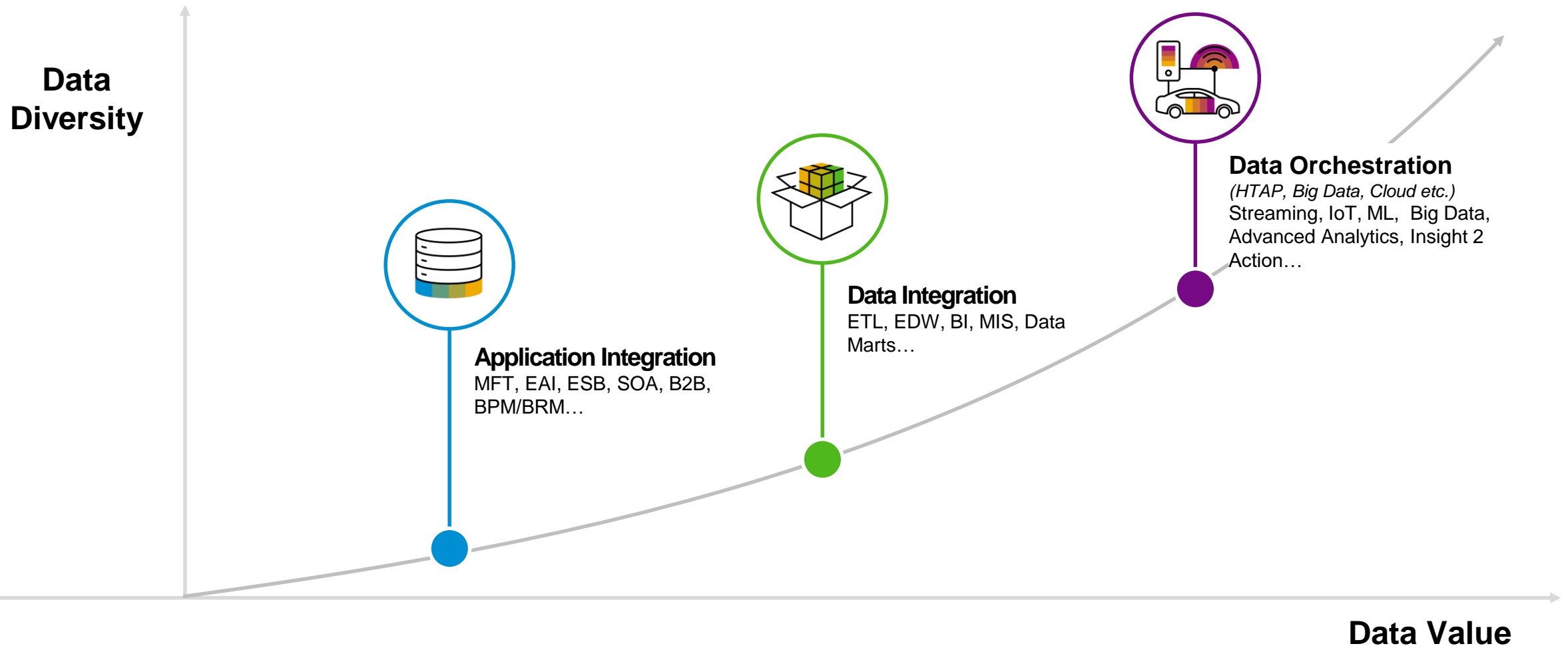
“Can I rely on this data to drive decisions in a critical process? How was it gathered?”



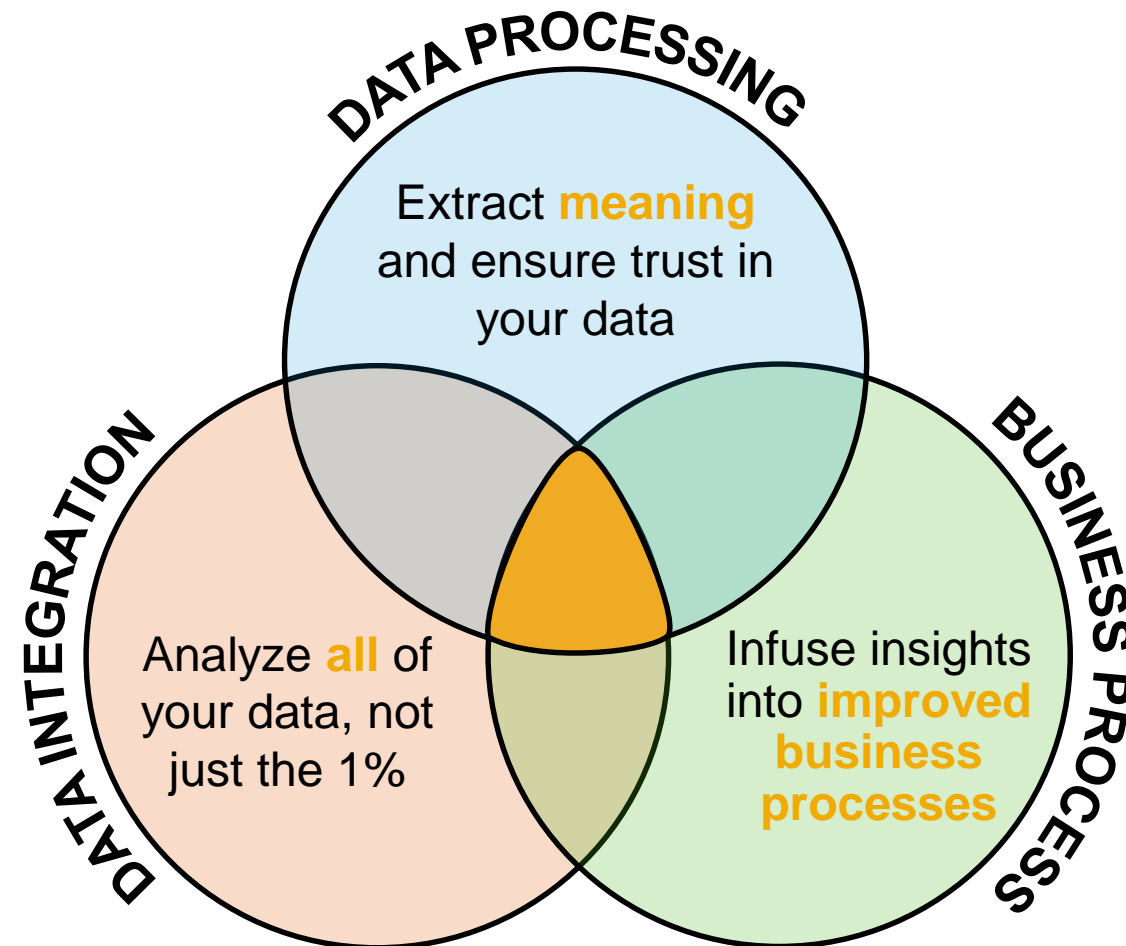
Operations

“My data and engines are spread between my on premise data centre and my hyperscaler cloud”

The evolution of Data Management in enterprise landscapes



SAP Data Intelligence – From Data Integration to **Data Orchestration**



A real-life example: HR Best Fit with SuccessFactors

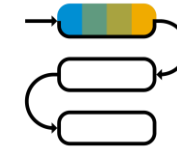
The business objective: find best fits within workforce to cover job postings



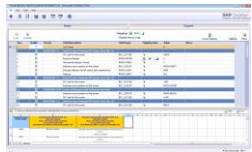
Data Integration



Data Processing



Business Process



SAP HR data



SFSF Seniority data



Knowledge Scoring Data



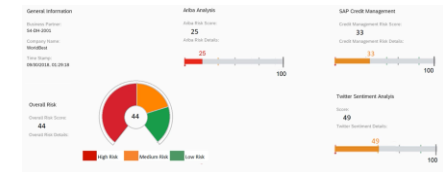
Location Data



Areas & Knowledge Data



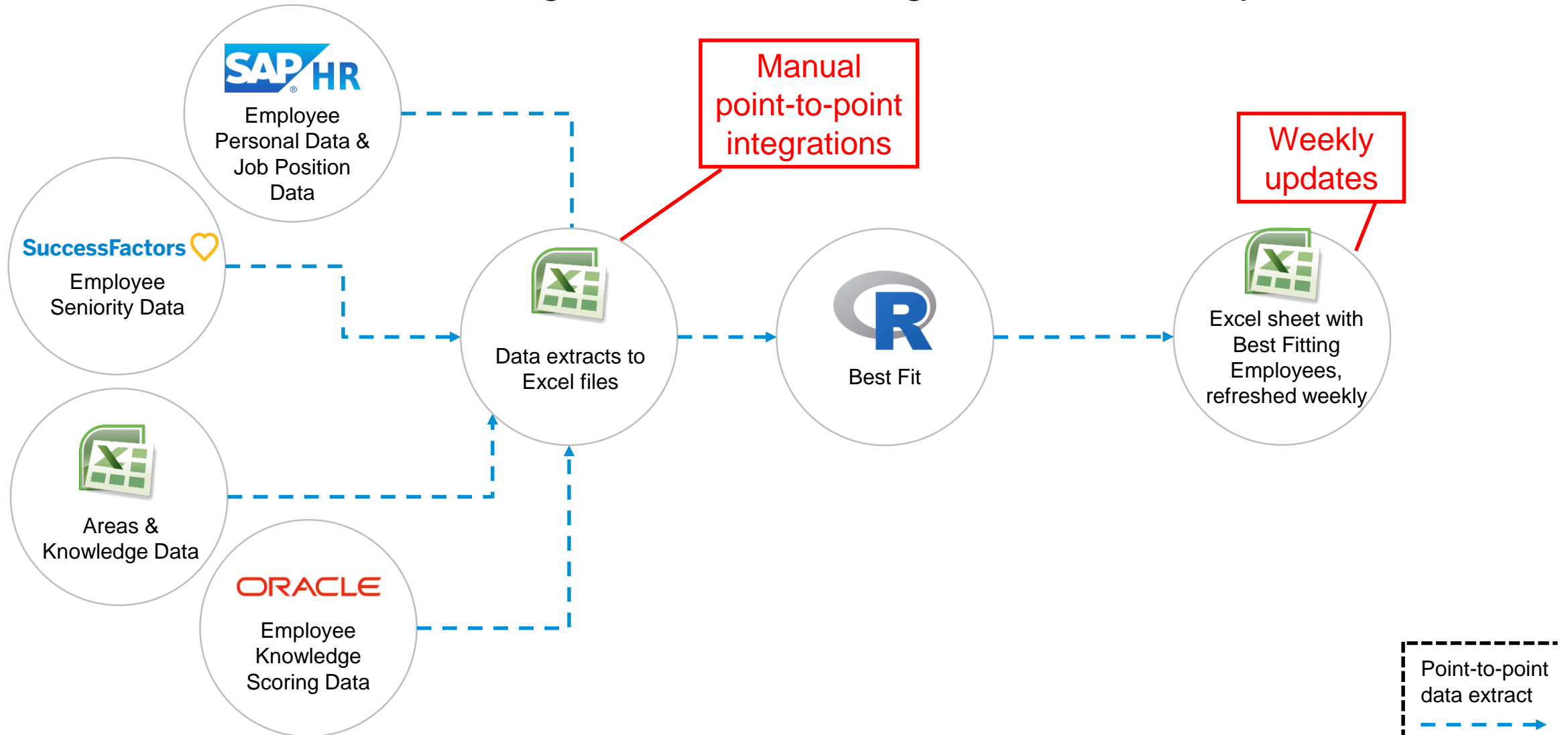
R Best Fit Model



Job posting
vs.
Employees
Score

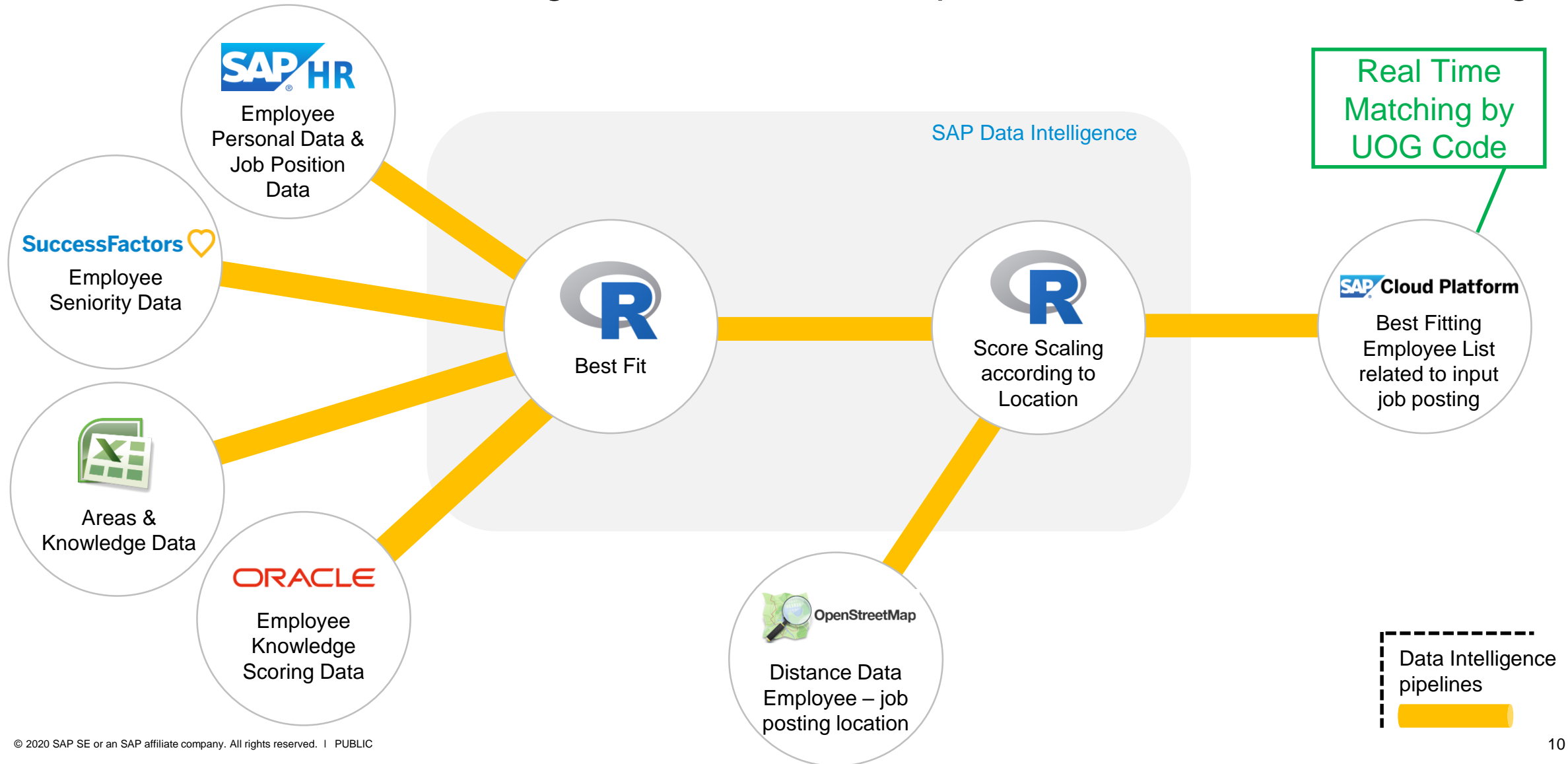
A real-life example: HR Best Fit with SuccessFactors

The as-is before Data Intelligence: manual integration and weekly batches



A real-life example: HR Best Fit with SuccessFactors

The solution with Data Intelligence: streamlined operationalization, real-time insights



SAP Data Intelligence – Delivering modern Data Management to the Enterprise

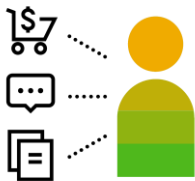


Data Pipelining

Flexible pipelines spanning structured / unstructured / streaming, batch / (near-)real-time, OLTP and OLAP, across any cloud and on-prem deployment.



Orchestration of Heterogeneous Data Processing



Enterprise Data Catalog

Metadata crawling, browsing and profiling; data lineage, data preparation, data anonymization and privacy enforcement

SAP Data Intelligence

Data pipelines integrate disparate data across distributed infrastructures



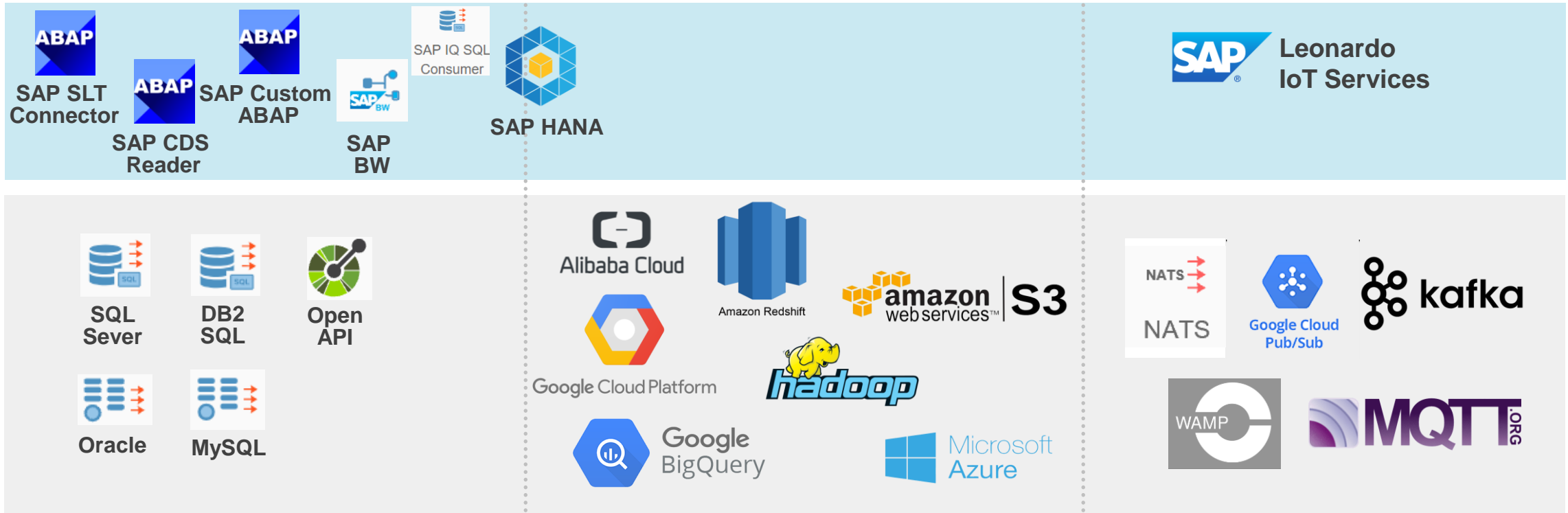
Structured Data



Unstructured Data

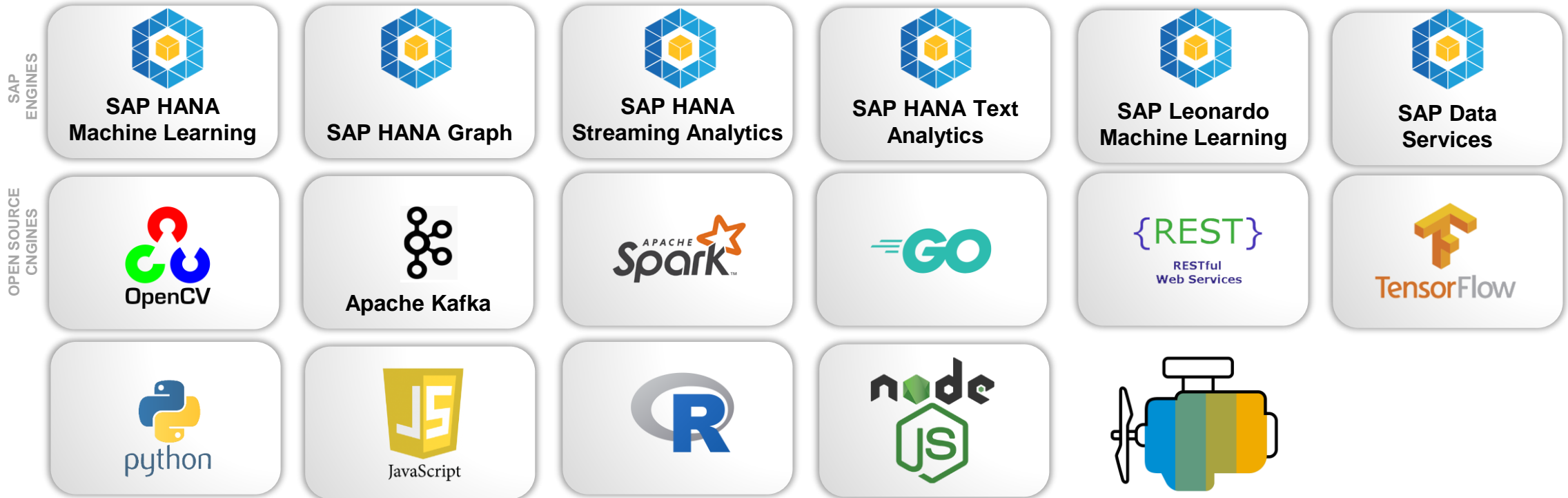


Streaming Data



SAP Data Intelligence

Orchestrate several heterogeneous processing engines



... across distributed infrastructures, cloud and on-prem, minimizing data movements



SAP Data Intelligence

A centralized Data Catalog for your enterprise data governance

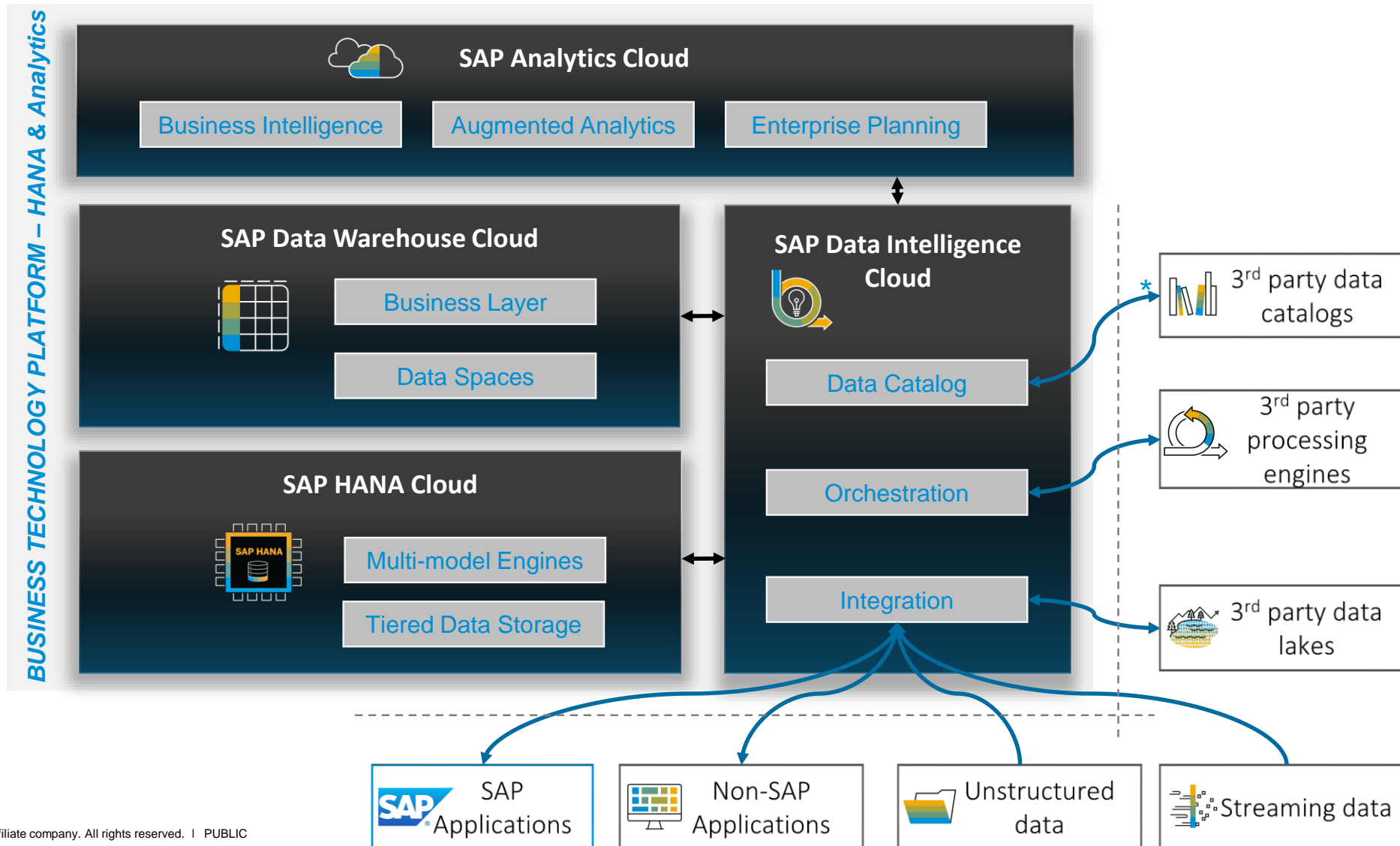
- Data discovery, browsing and profiling
- Self-service Data Preparation
- Data lineage
- Data quality rules and scorecards

The image displays three main screenshots from the SAP Data Intelligence interface:

- Top Screenshot: SAP Data Intelligence Metadata Explorer**
Connection: GOOGLE_CLOUD_PLATFORM | Fact Sheet: /Flight Data/Flight_Statistics.csv | INDEXED | PROFILED
Metadata Summary: Type: CSV, Size (bytes): 363,245, Owner: root, Rows: 5,083, Last Modified: 4/10/2019, 2:43 PM
Column Types: Donut chart showing distribution of data types (Character, Numeric, DateTime, Other).
Trend by Row Count: Bar chart showing a trend of 5083 and 1.
- Middle Screenshot: Dashboard - Address Quality Dashboard**
Dashboard highlighting customer address quality.
High Level Results: Customer Rulebook Rulebook Pass % (95.6%), USA Customer File 1 Dataset Pass % (9.24), USA Customer File 2 Dataset Pass % (9.94), MDG Customers Dataset Pass % (9).
- Bottom Screenshot: Dataset Lineage and Column Preparation**
Dataset Lineage: /GCS_CONN/output/revenue_BY_COUNTRY_AND_REGION.csv
Lineage Overview: Flow diagram showing data lineage from source datasets to target datasets.
Dataset Summary: Name: revenue_BY_COUNTRY_AND_REGION.csv, Qualified Name: output/revenue_BY_COUNTRY_AND_REGION.csv, Status: LINEAGE (NOT PROFILED).
Columns: Table with columns Name, Type, Description, Tags. Row: N_NAME, STRING, Description, Tags.
Column Preparation: Menu for ENAME column with options: Change Case, Fill, Trim, Replace, Duplicate, Rename, Split, Combine, Convert To Date/Time, Convert To Number, Clear, Remove.

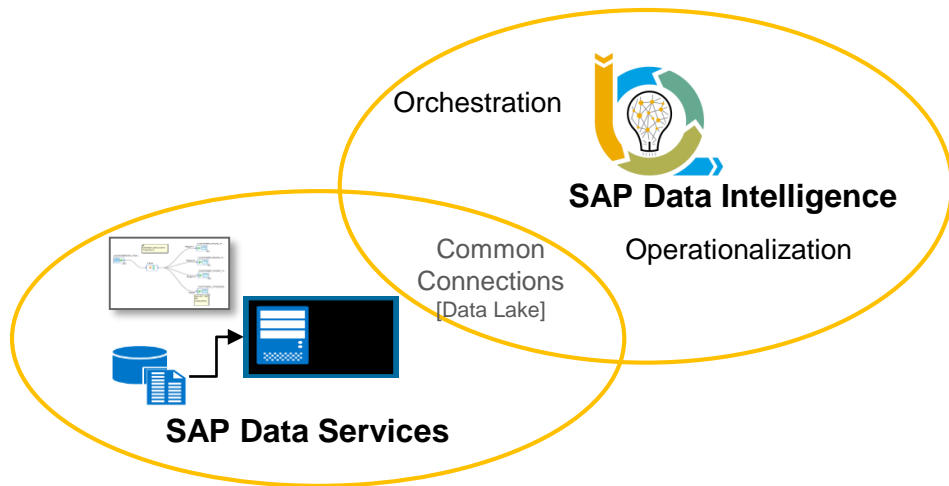
SAP Data Intelligence + SAP HANA Cloud + SAP Data Warehouse Cloud

Enabling and end-to-end data fabric

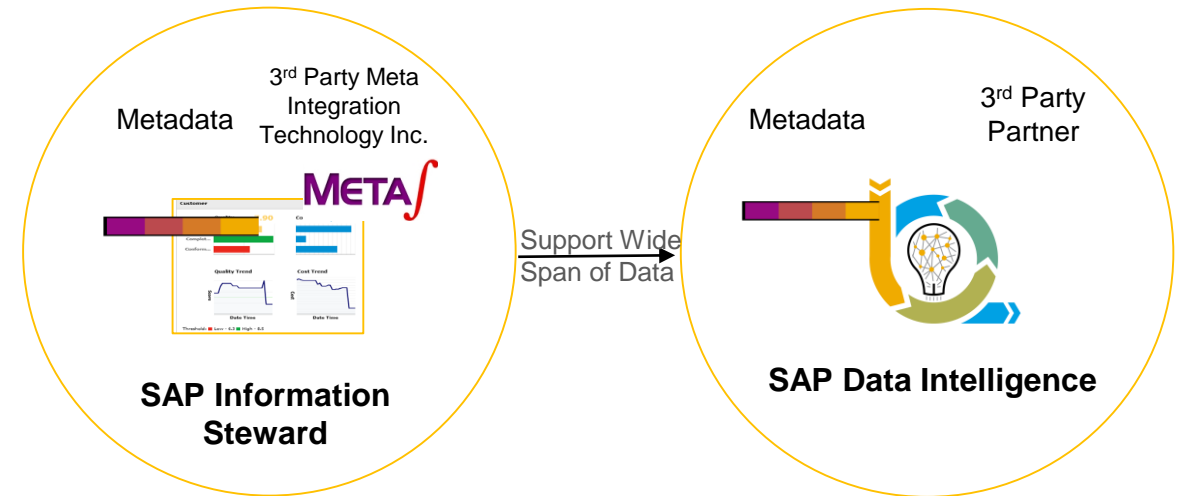


SAP Data Intelligence and SAP on-prem EIM tools

Enabling and end-to-end hybrid data management



- Extend reach of Data Services assets across diverse system landscapes
- Orchestrate Data Services jobs within Data Intelligence pipelines
- Enhance enterprise business processes with high value data and insights from external systems



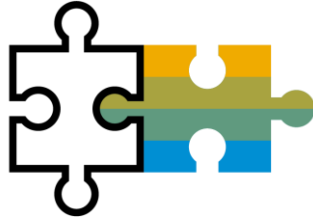
- Reuse Information Steward rules within Data Intelligence
- Shared glossary of terms
- Metadata shared across applications, augmented by ML learning models

Upcoming Innovations



Operations & Administration

Workload Management, Debugging, Back Up & Recovery, Cloud Connector Support, Versioning, Lifecycle Mngt



Deeper Application Integration

Align integration model to consume and interact with Apps



Content & Templates

Offering out-of-the-box templates and modeling wizards



Extend Span of Connectivity

Enhanced connectivity to Cloud Services (e.g. DWHs), improved file handling



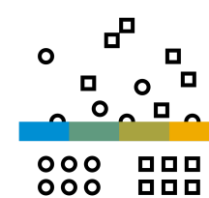
Integration with HANA Cloud / DWC

End to end data pipelining for HANA Cloud & Analytics Cloud



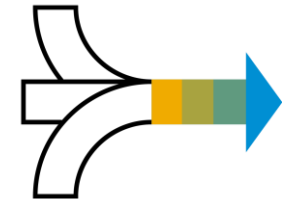
Intelligent Processing

Apply ML and intelligence in modelling and cataloging, simplification of workflow



Data Governance

Predefined glossary terms, Support lineage at column level, Auto extraction of lineag ABAP Lineage object level extraction



Hybrid Capabilities

Consolidate functions and foster hybrid on-premise / cloud scenarios with EIM portfolio

SAP Data Intelligence – Why it is different



Single end-to-end tool

Discover and catalog your data, integrate it and orchestrate, manage ML operations, all in one single product.



Multi-cloud, agnostic and open

Adaptable architecture built on open technologies, available as-a-service, or BYOL in the cloud on any hyperscaler or on premise.



Re-use existing on-prem investments

Natively reuse and connect all SAP on prem technologies (HANA, BW, Data Services, SLT etc.)

Thank you.

Contact information:

Silvio Arcangeli

Product Management

SAP Data Intelligence

silvio.arcangeli@sap.com