Capture Business Opportunities from
Systems of Record and Systems of Innovation

Amit Satoor, SAP
March Hartz, SAP

PUBLIC

SAP Run Simple

Big Data transformation powers digital innovation system
Relevant nuggets of information are still hard to find

Business expectations

- Deep data insights
- Broad data scope
- Access to most recent data
- Answers in real time
- Not delayed by constant data preparation
- Predictive capabilities

Business transformation

Data transformation

Analytics

Blockchain

Internet of Things

Machine Learning
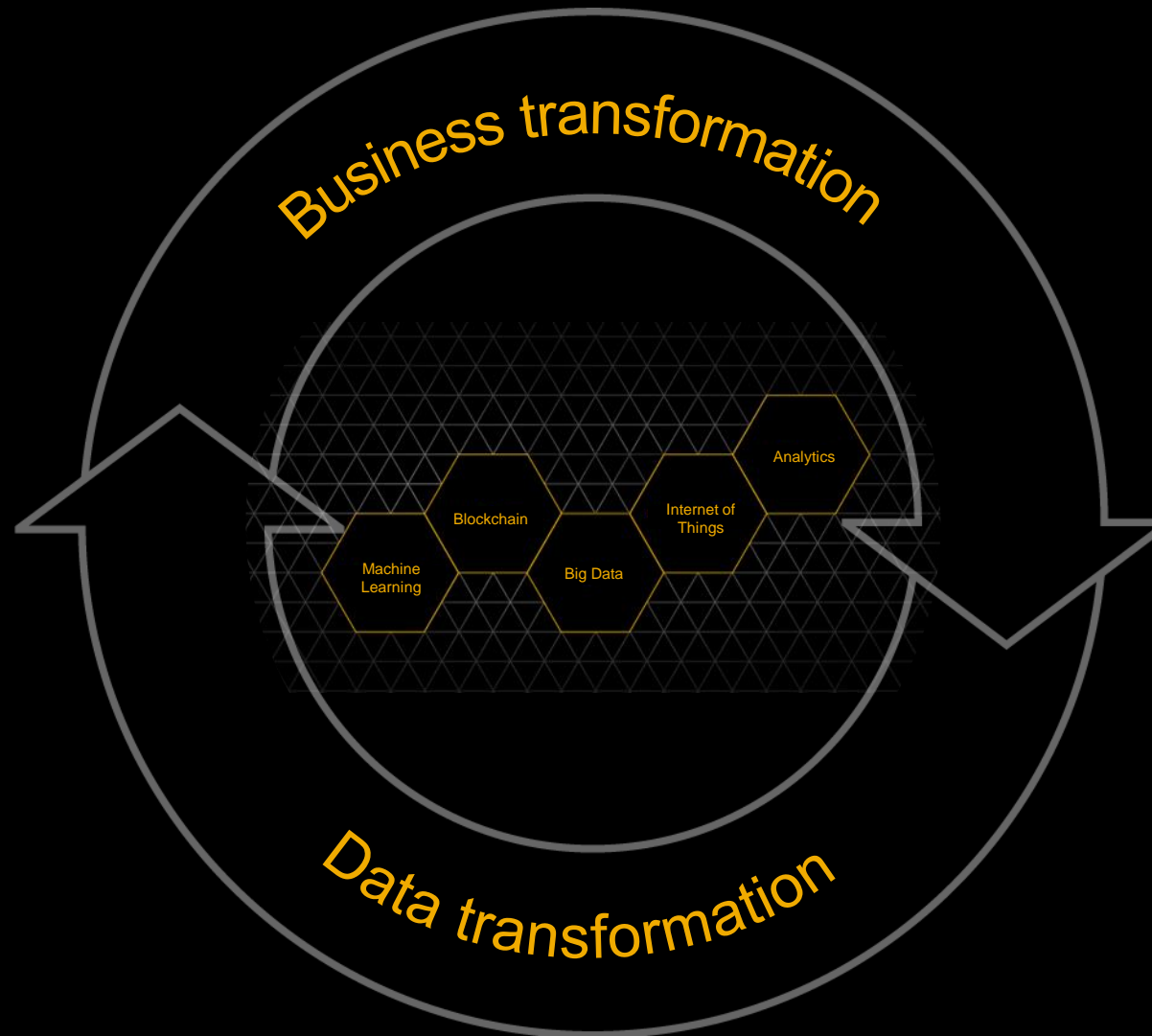
Big Data

SAP Leonardo technologies

- Internet of Things
- Machine learning and artificial intelligence
- Big Data
- Analytics
- Blockchain

Why is it difficult to capture new opportunities?

Silos, data movement, and complexity in understanding raw data hinder business agility and innovation
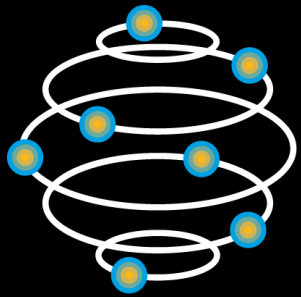
- Exploding data volumes
- Accelerated data velocity
- Increasing data variety

Cloud storages

Proliferating data lakes

Multiple EDWs, data marts
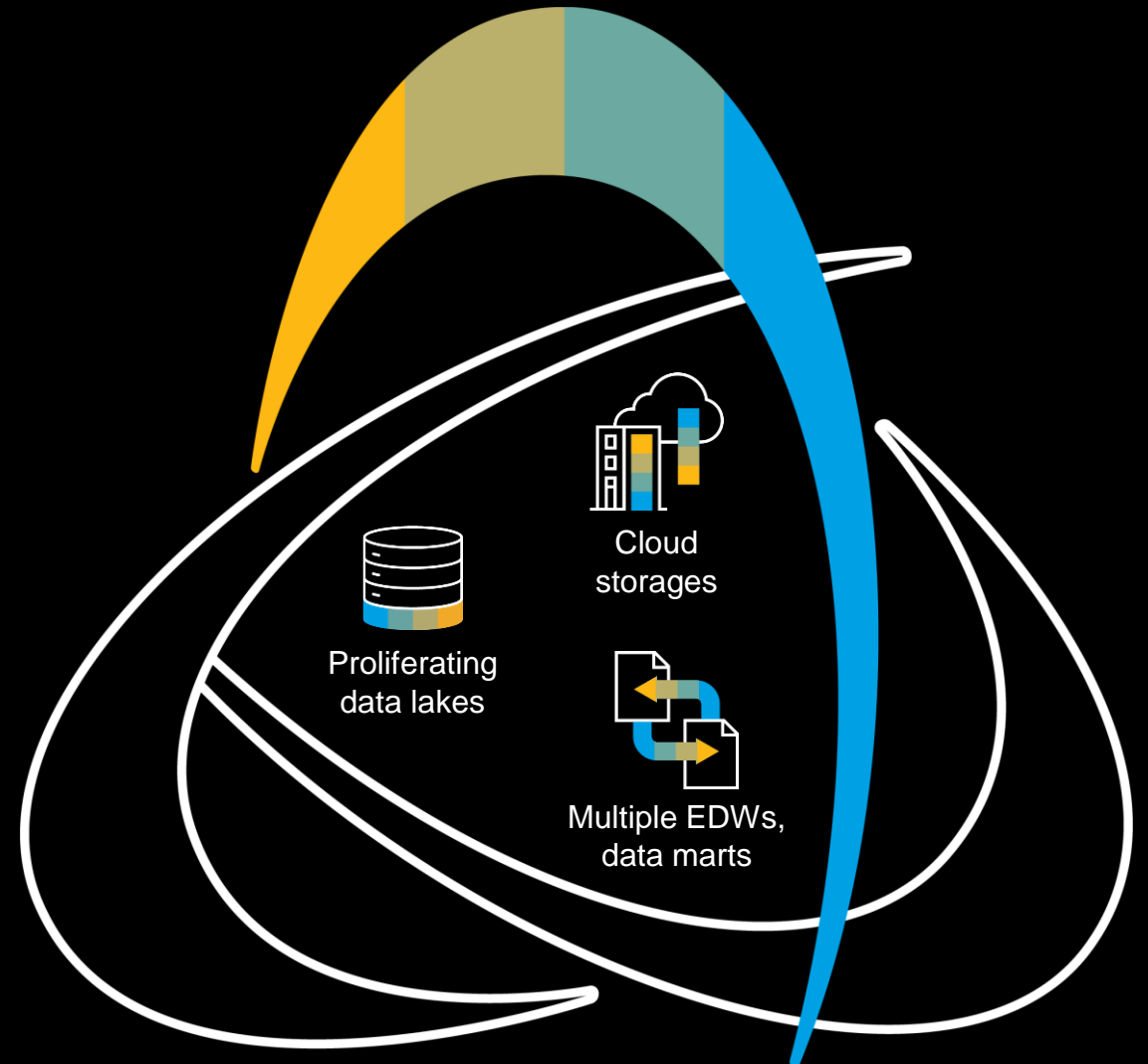
7 LAWS
OF UNIVERSAL DATA

Logical model
Integrated for 360-degree business view

Low latency
Immediate response

Performance
To handle all types of needs

Independence
Of data and compute for right-place decisions

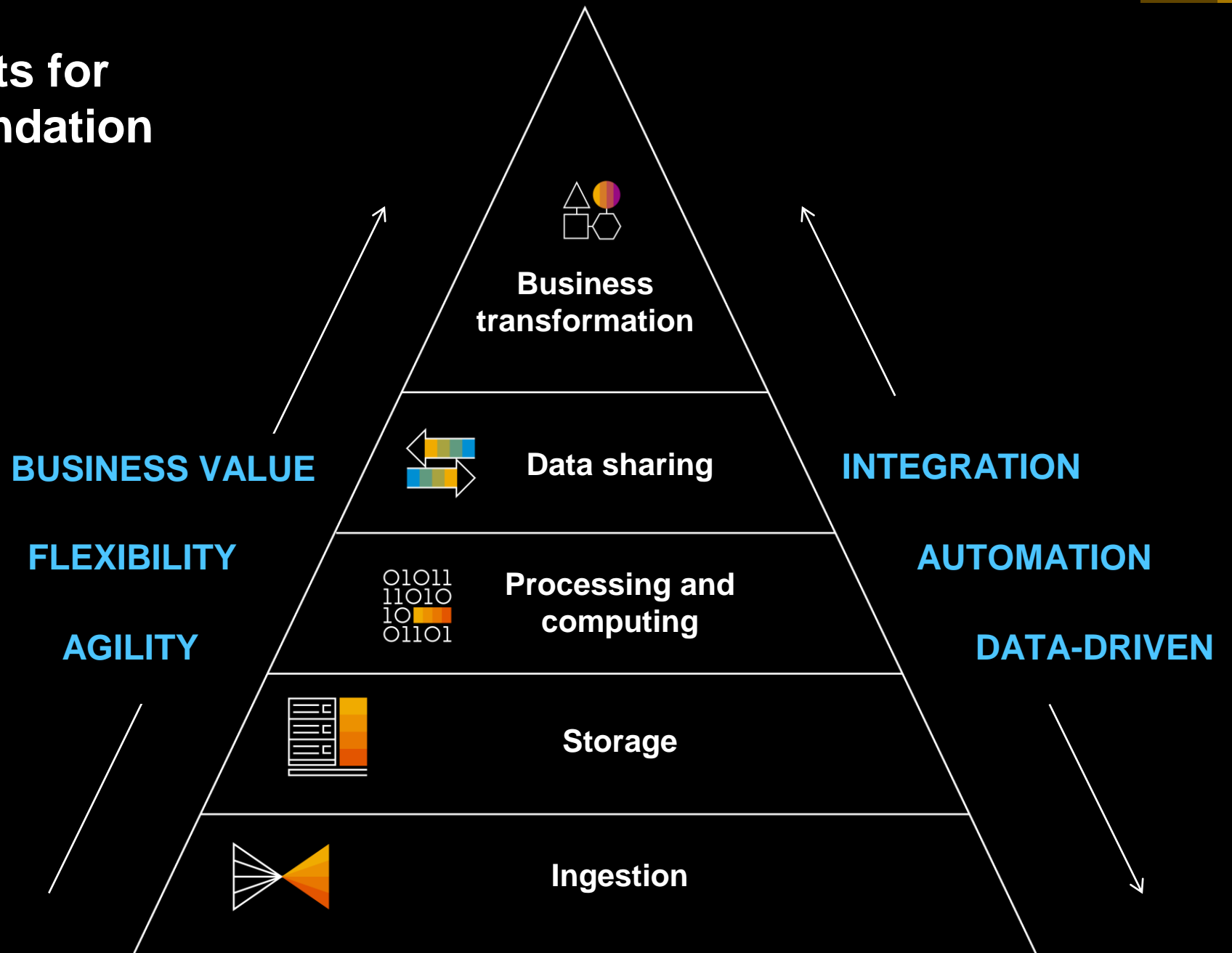**Frictionless**
Sharing and enrichment for thorough understanding

Freedom
Cloud, hybrid, or on premise

Governance
Security, auditing, and compliance

Critical components for universal data foundation

Business transformation

BUSINESS VALUE · **Data sharing** · INTEGRATION

FLEXIBILITY · **Processing and computing** · AUTOMATION

AGILITY · DATA-DRIVEN

Storage

Ingestion

A radical new approach to uncover new opportunities

Embrace new innovations to run and differentiate your business

Enterprise data

SAP HANA platform

SAP BusinessObjects solutions

SAP solutions for EIM

SAP Vora

SAP Cloud Platform Big Data Services

Big Data

mahout

Spark

HIVE

SQL

Spark Streaming
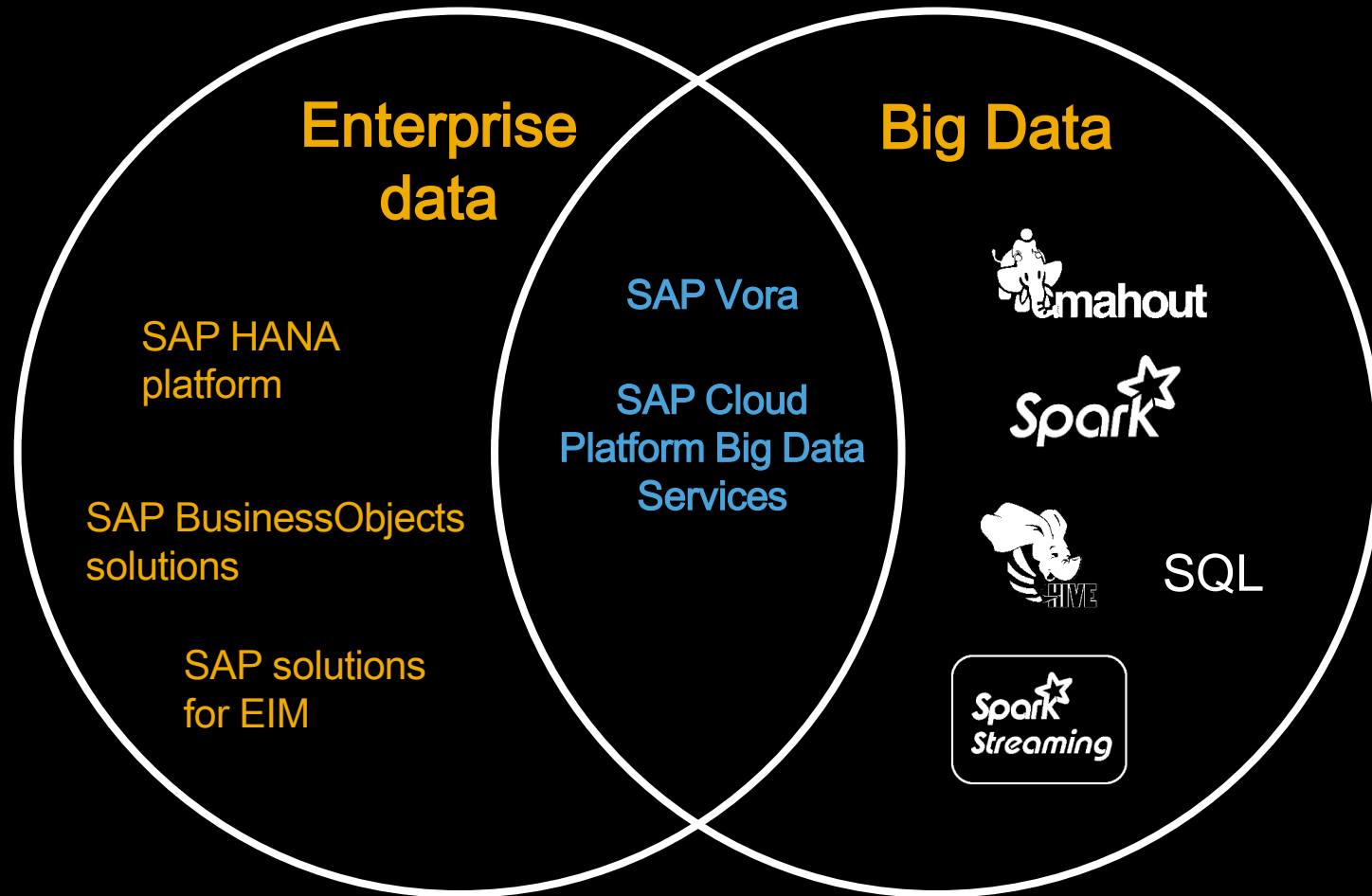
Big Data **Applications**

Big Data **Analytics**

Big Data **Computing framework**

Big Data **Managed services**

SAP HANA platform

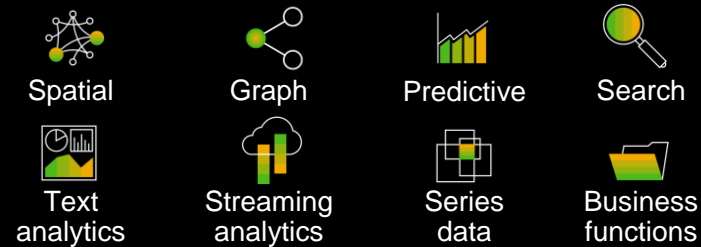The in-memory data platform for digital business

─── All devices ───
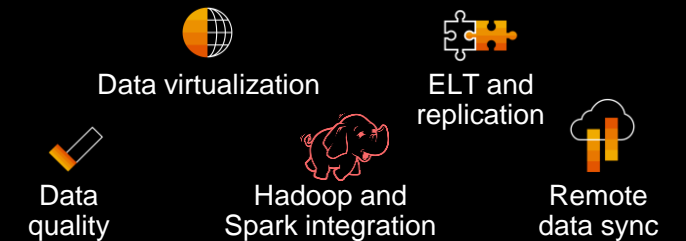
─── SAP, ISV, and custom applications ───

SAP HANA platform

APPLICATION DEVELOPMENT

Web server

JavaScript

SAP Fiori UX

Graphic modeler

Application lifecycle management (ALM)
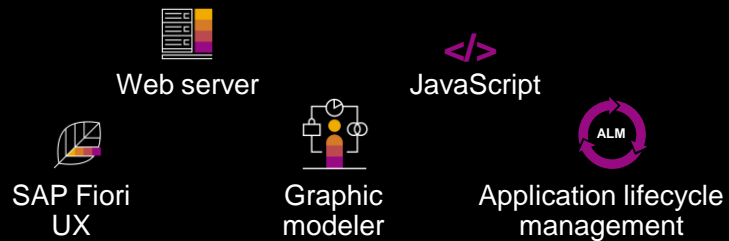
ADVANCED ANALYTICAL PROCESSING

Spatial

Graph

Predictive

Search

Text analytics

Streaming analytics

Series data

Business functions

DATA INTEGRATION AND QUALITY

Data virtualization

ELT and replication

Data quality

Hadoop and Spark integration

Remote data sync

DATABASE MANAGEMENT

Columnar OLTP+OLAP

Multicore and parallelization

Advanced compression

Multitenancy

Multitier storage

Data modeling

Openness

Admin and security
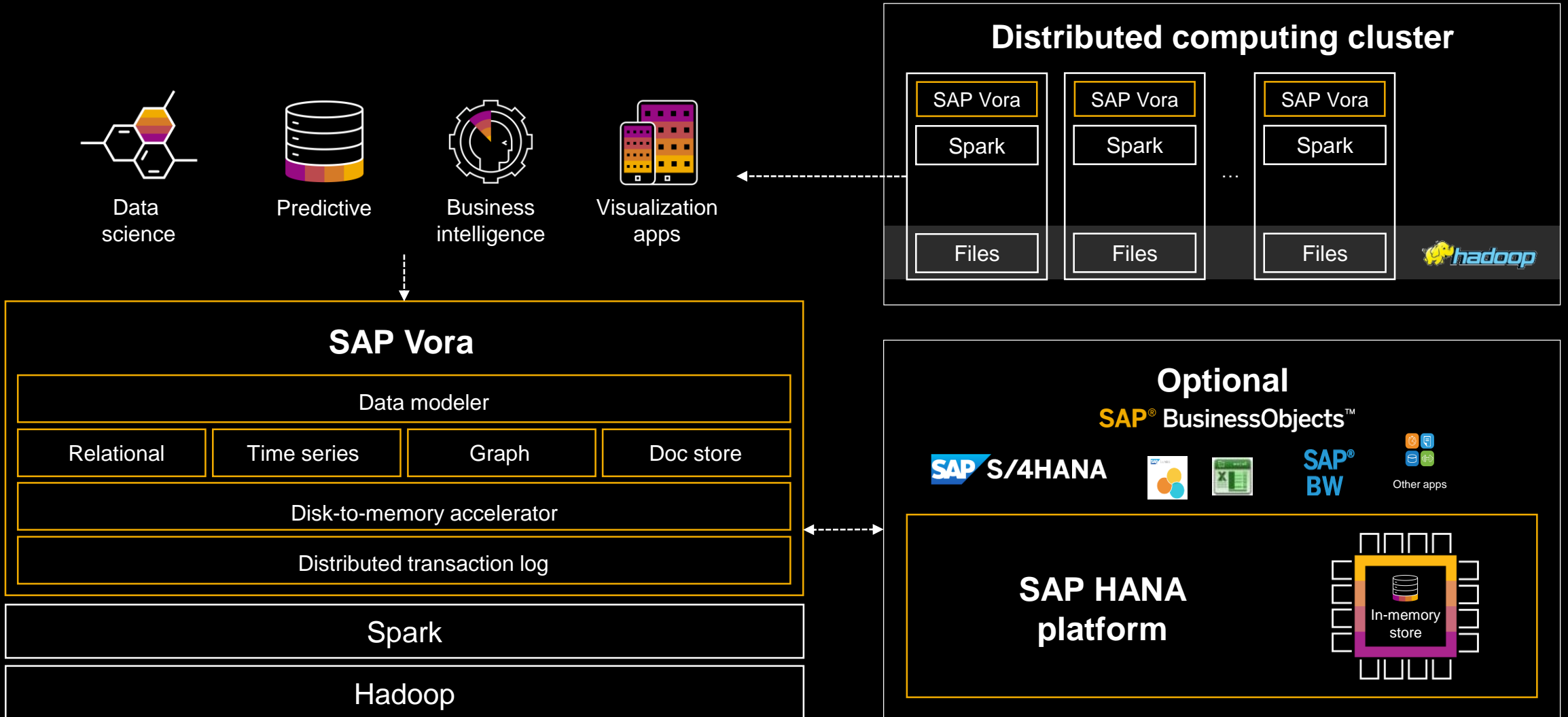
High availability and disaster recovery

ONE open platform   **OLTP + OLAP**   **ONE copy of the data**

SAP Vora

Distributed computing solution to uncover actionable insights from Big Data

Distributed computing cluster

SAP Vora	SAP Vora	SAP Vora
Spark	Spark	...	Spark
Files	Files	Files

hadoop

Data science   Predictive   Business intelligence   Visualization apps

SAP Vora

Data modeler			
Relational	Time series	Graph	Doc store

Disk-to-memory accelerator

Distributed transaction log

Spark

Hadoop

Optional

SAP® BusinessObjects™

SAP S/4HANA   **SAP® BW**   Other apps

SAP HANA platform

In-memory store

Rapidly and efficiently acquire and consolidate all data
Integrate and ingest massive amounts of diverse and arbitrary data

Data
virtualization

ELT/
replications

Hadoop/
open source
integration

Remote data
sync

Data quality

Data quality

MDM

Data
integration

Data discovery
and
architecture

Information
lifecycle
management

Leverage data in real time for innovation, faster deployment, and lowered total cost of ownership (TCO)

- Quickly acquire, stream, replicate, and transform any kind of data at any velocity

- Connect relational databases, classic EDW, and Hadoop sources quickly

- Integrate any data source – flat file, relational, Big Data, or event

- Employ simple and powerful data acquisition, streaming, and transformation capabilities

- Use efficient tools for data movement, cleansing, and placement throughout the architecture

Manage massive volumes of data with elastic storage and persistence

Columnar OLTP+OLAP

Multicore and parallelization

Advanced compression

Multitenancy

Multitier storage

Data modeling

Admin and security

High availability and disaster recovery

amazon web services

MAPR

cloudera

Hortonworks

Leverage ALL data for innovation for quicker ROI and reduced latency and TCO

- Organize and store petabytes of data for ready analysis

- Use native multitemperature management and virtualization; reduce data movement by accessing and processing data from any source at its location

- Optimize and integrate data warehousing with logical data modeling

- Use full virtualization and schema extension capabilities; federate data from any data source

Achieve real results to get the insights you need, alone or in combination
Unified in-memory computing services and distributed processing

Spatial

Graph

Machine learning

Search

Text analytics
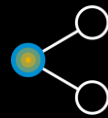
Streaming analytics

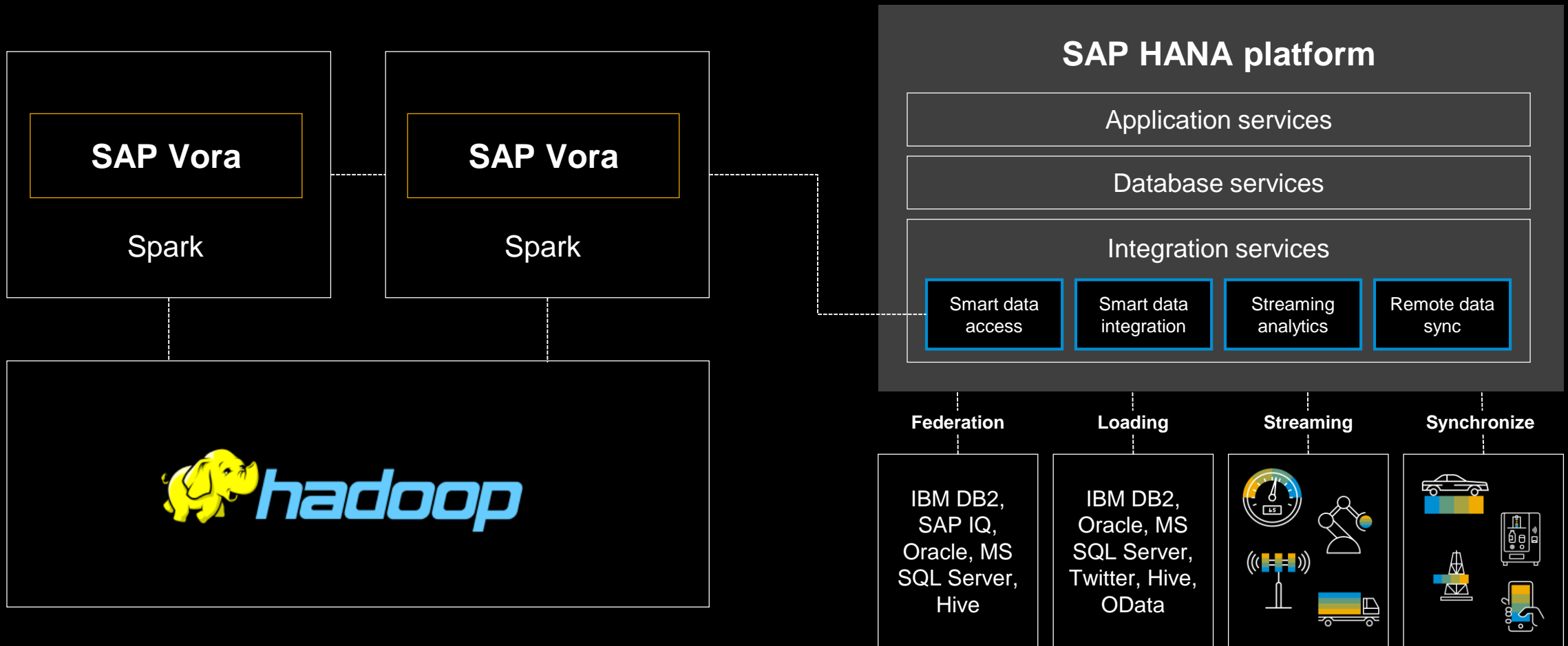Series data

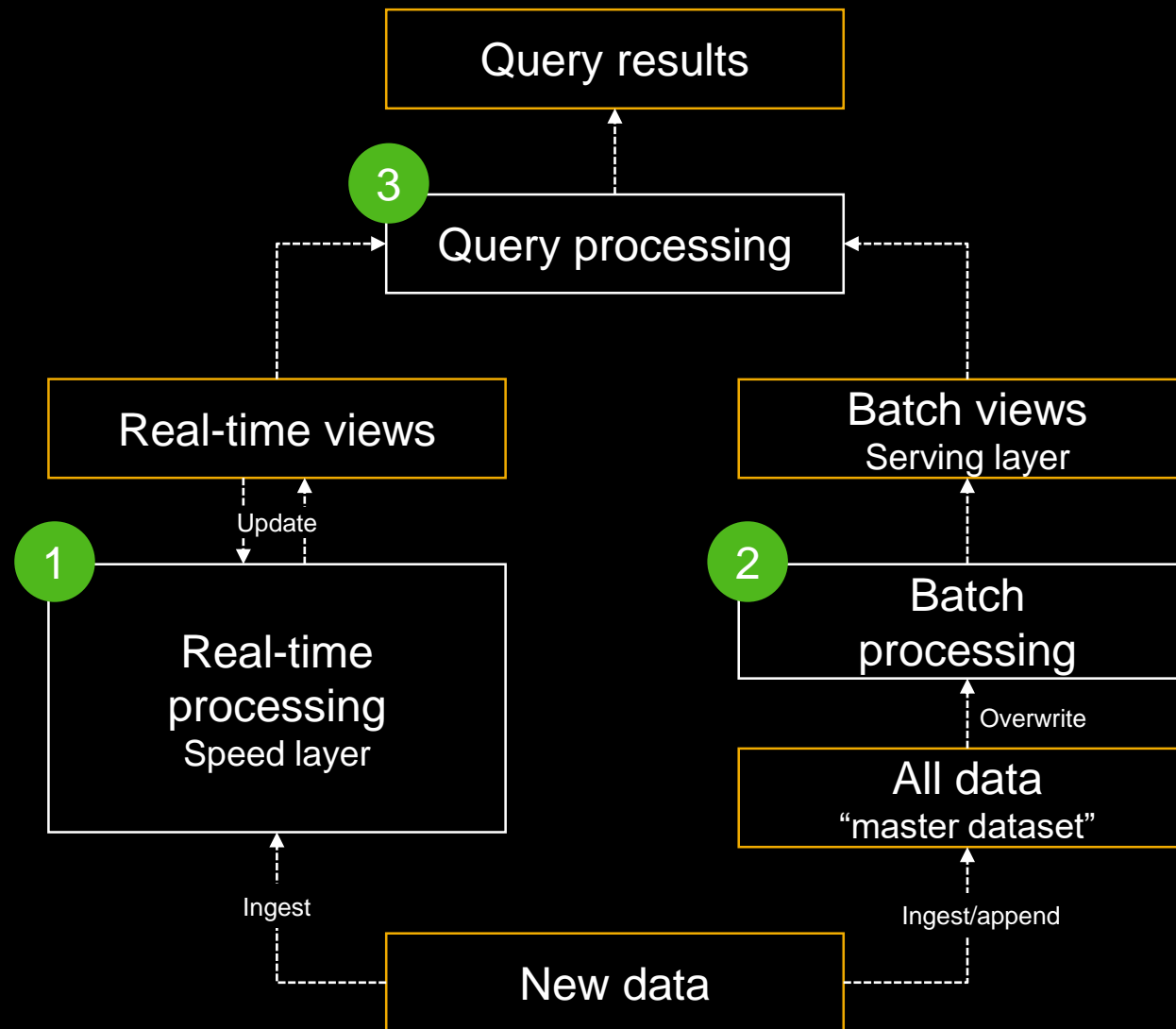Business functions

OLAP

Time series

Graph

Doc store

Discover insights to drive innovation and make better and faster business decisions

- In-memory first architecture for operational systems and Big Data landscape

- Integrated multimodal advanced analytics to converge multiple complex analytic processing without data movement or duplication; native multimodal SQL (predictive, search, text, spatial, graph, series, streaming)

- Accelerated OLAP analysis and scoring with built-in algorithms; ability to adapt the models frequently

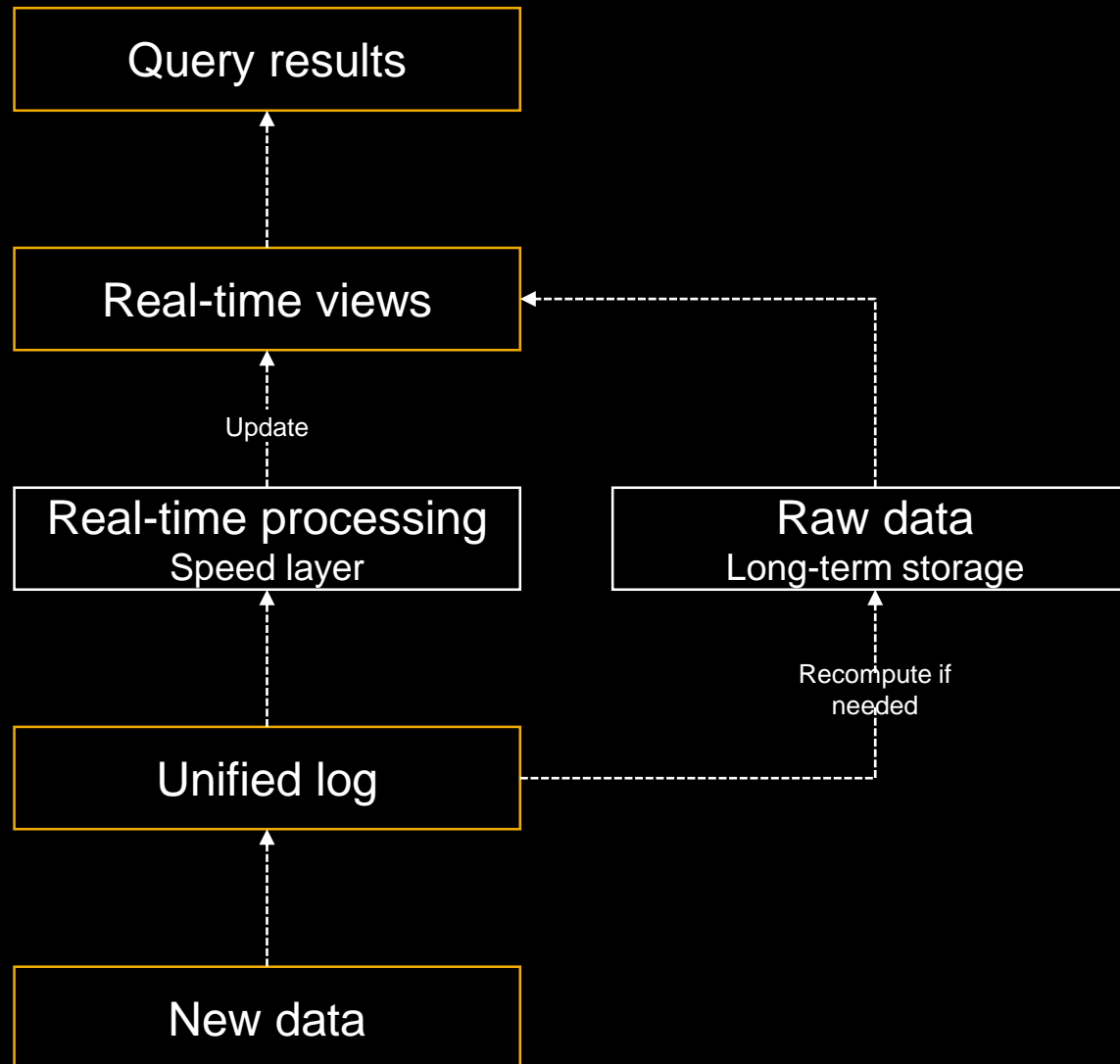- Multiplatform architecture to deploy in the cloud, on premise, or hybrid

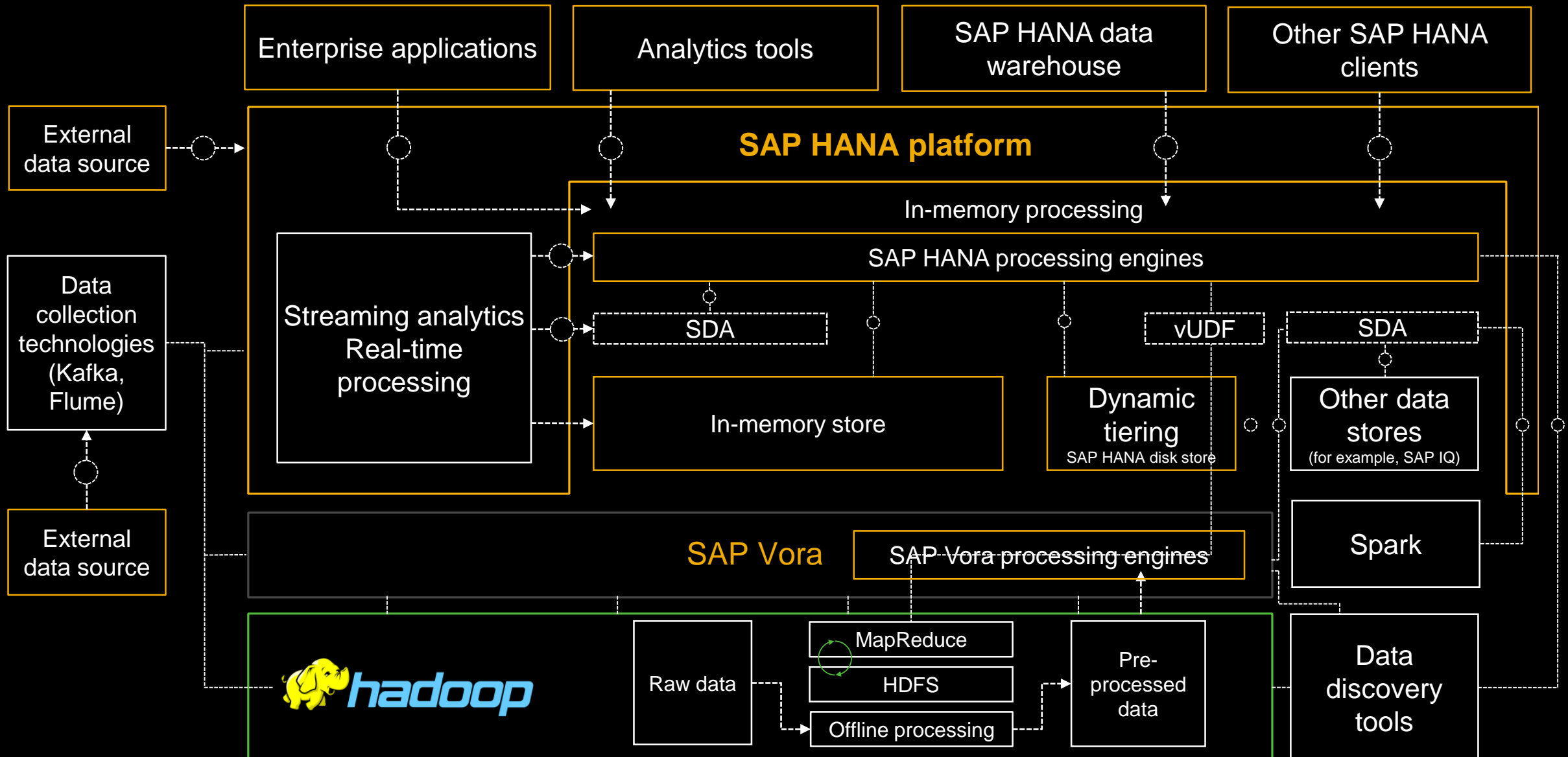# SAP HANA and SAP Vora: Connecting data from many sources

SAP Vora

Spark

SAP Vora

Spark

hadoop

SAP HANA platform

Application services

Database services

Integration services

| Smart data access | Smart data integration | Streaming analytics | Remote data sync |

Federation

Loading

Streaming

Synchronize

IBM DB2, SAP IQ, Oracle, MS SQL Server, Hive

IBM DB2, Oracle, MS SQL Server, Twitter, Hive, OData

Enabling Big Data processing patterns: Lambda architecture

Enabling Big Data processing patterns: Kappa architecture

Query results

Real-time views

Update

Real-time processing
Speed layer

Raw data
Long-term storage

Recompute if needed

Unified log

New data

Universal Big Data foundation for SAP Leonardo

Enterprise applications

Analytics tools

SAP HANA data warehouse

Other SAP HANA clients

External data source

SAP HANA platform

In-memory processing

SAP HANA processing engines

Data collection technologies (Kafka, Flume)

Streaming analytics Real-time processing

SDA

vUDF

SDA

In-memory store

Dynamic tiering
SAP HANA disk store

Other data stores
(for example, SAP IQ)

External data source

SAP Vora

SAP Vora processing engines

Spark

Raw data

MapReduce

HDFS

Offline processing

Pre-processed data

hadoop

Data discovery tools

Data hub for the Internet of Things (IoT)

IoT apps (SAP Cloud Platform)

Data processing and orchestration

↑

Big Data lakes

↑

Ingestion and stream

↑

IoT gateway and services

↑
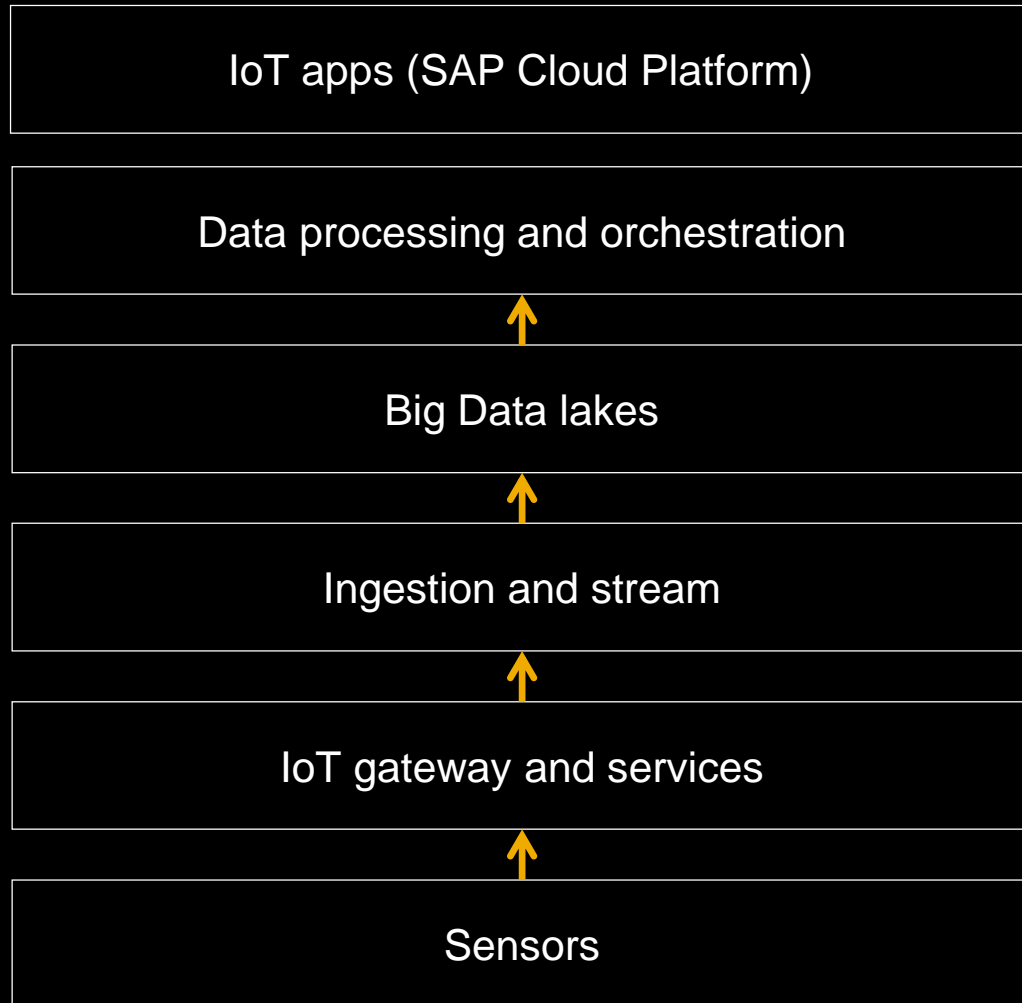
Sensors

Powerful end-to-end IoT scenario
Go from ingestion to consumption with a modern architecture and ingestion pipeline

- Harness high-volume data results (for example, sensor data)

- Manage single or multiple data lakes and environments with the needed data quality
  On-premise Hadoop, SAP Cloud Platform Big Data Services, and/or Amazon S3

- Ingest and process data for enterprise applications

- Automate, design, and run all data processes on the data hub

- Define and execute the IoT ingestion process

Utilities rise to the smart meter challenge
Smart meters lead utilities to reconsider how they use data

Current situation

- 8 TB a month and over 1 PB in a 10-year period
- TBs of data each month
- Must retain data for 10 years
- Must accurately forecast energy usage
- Current DW technology is approaching end of life
- Massive amounts of data stored in proprietary solution is hard to manage and with a high total cost of ownership

SAP Vora

Historical sensor data

Parallelized queries

SAP HANA

Dynamic tiering

Query data within and across tiers
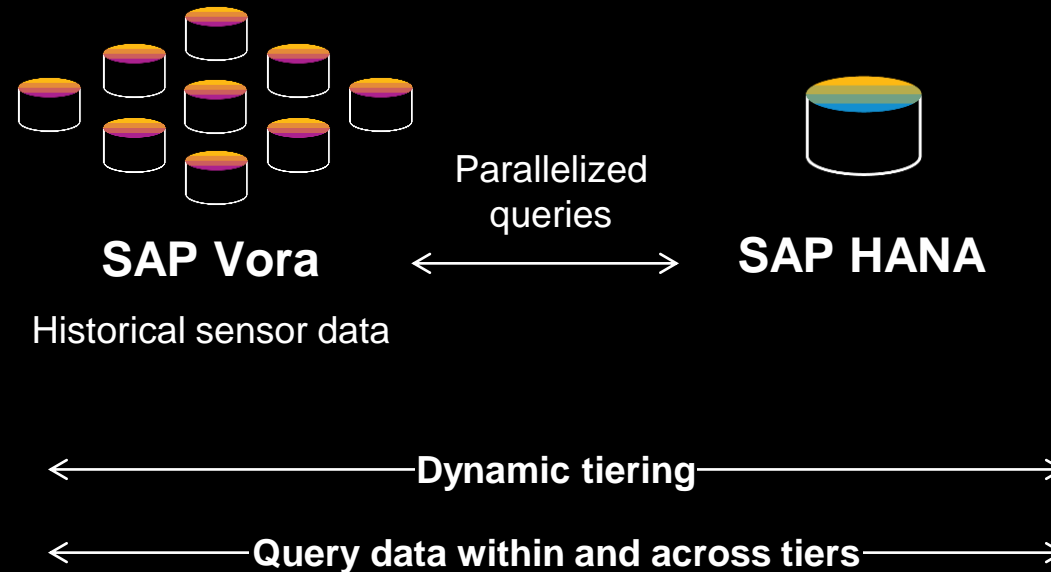
Benefits of integrating Big Data

- Meet regulatory requirements
- Forecast energy usage
- Detect fraud and predict maintenance requirements
- Respond intelligently to variations in supply and demand with smart grids

The results of simplified Big Data ownership with SAP HANA, dynamic tiering, SAP Vora, and Hadoop

Solution

- HOT+WARM+COLD data management strategy leverages SAP HANA data compression and data tiering

- Combined SAP HANA, dynamic tiering, and Hadoop into a single landscape

- Automatic access to stored data in SAP HANA, dynamic tiering, and Hadoop and SAP Vora with query execution

- Automated data movement between storage tiers using the data lifecycle manager

- Foundation for advanced predictive analytics and future business capabilities
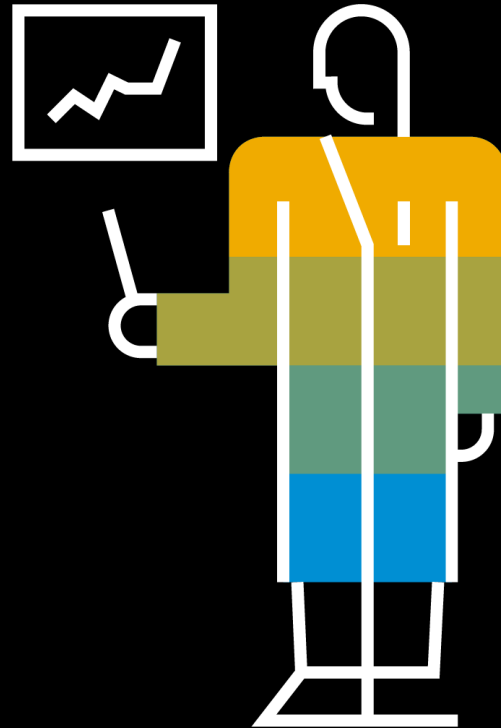
Results

- Instant real-time analytics through SAP HANA

- 75% savings in storage costs compared to current solution

- Seamless integration with Hadoop technology, enabling data scientists to access and manage Hadoop data

- Ability to charge business based on data storage and performance requirements

Big Data strategy from SAP

Top reasons and capabilities

Comprehensive approach

Combines integration, workflow, pipelining, and data governance in one place

Embracing the entire data landscape

Diversity of data sources, data types, and endpoints; cloud or on-premise; SAP and non-SAP data

Unified management

Enables more users with a visual and unified solution to manage, monitor, and develop powerful data pipelines

Enterprise-wide discovery

Provides the ability to understand the entire landscape, the individual systems, and the data within

Data-driven architecture

Massively scalable, containerized, and distributed architecture build for serverless computing

Data sharing

Governance | Streams and pipelines | Orchestration

Enterprise data

Big Data

SAP HANA platform

SAP BusinessObjects solutions

SAP solutions for EIM

SAP Vora

SAP Cloud Platform Big Data Services

Hadoop integration

mahout

Spark

SQL

Spark Streaming
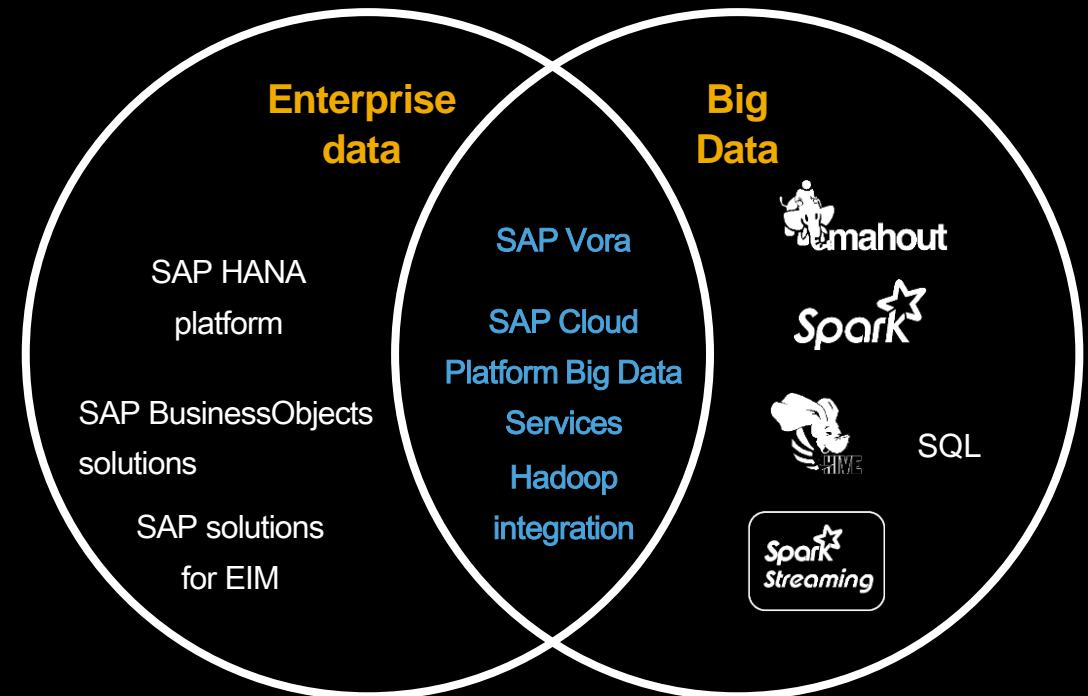
Join the Big Data product leadership program
Unearth business signals from Big Data landscapes to grow value

Early adopter care

Gain competitive advantage Get implementation support Provide product feedback

BigData@SAP.com

Be in the know

Get the latest news and access to special activities