

The background of the top half of the cover is an abstract, futuristic data visualization. It consists of numerous glowing, semi-transparent rectangular blocks in shades of orange, yellow, and blue, arranged in a curved, overlapping pattern that suggests a 3D bar chart or data flow. The lighting is soft and bokeh-like, creating a sense of depth and movement.

Data Management and Analytics with SAP® BTP

- › Explore SAP BTP's data management and analytics portfolio
- › Learn about key products, including SAP HANA Cloud, SAP Business Data Cloud, and SAP Analytics Cloud
- › Get to know SAP's artificial intelligence capabilities with SAP AI Launchpad and SAP AI Core

Gehlot • Gericke • Gupta • Isacc
Kalwachwala • Markham • Munshi • Sam

Contents

Preface	15
1 Introduction	25
1.1 Build, Run, and Optimize a Business	25
1.2 Challenges for Executives	30
1.3 Characteristics of a Successful, Data-Focused Organization	31
1.4 Data Strategy and Information Management Maturity Models	34
1.5 Data Integration	35
1.5.1 General Components of Data Integration	36
1.5.2 Data Integration and Evolution	41
1.6 Data Tiers in SAP	45
1.6.1 Transactional System	46
1.6.2 Operational Data Store	46
1.6.3 Enterprise Data Warehouse	46
1.6.4 Enterprise Analytics	47
1.6.5 Data Quality	48
1.6.6 Data Archiving	50
1.6.7 Security	52
1.7 SAP Business Technology Platform	52
1.8 SAP's Data and Analytics Portfolio	56
1.8.1 SAP Master Data Governance	57
1.8.2 SAP S/4HANA Migration Cockpit	57
1.8.3 SAP Business Data Cloud	59
1.8.4 SAP HANA Cloud	61
1.8.5 SAP BW/4HANA	62
1.8.6 SAP Analytics Cloud	63
1.9 Enterprise Value and Data Fabric	63
1.9.1 Approaches to Unifying Data	64
1.9.2 SAP's Component Map	67
1.9.3 Non-SAP Component Map	69
1.10 Bringing Them Together with SAP	71
1.10.1 Run the Business	71
1.10.2 Evolve the Business	72
1.11 Summary	74

- 2 Data Strategy** 77
- 2.1 Overview 78
- 2.2 Defining a Data Strategy 79
 - 2.2.1 Developing a Data Strategy 80
 - 2.2.2 Importance of a Data Strategy 91
 - 2.2.3 Characteristics of an Effective Data Strategy 96
 - 2.2.4 Applying the Clean Core Philosophy Within Data Management 104
- 2.3 Data and Information Management Maturity Models 104
 - 2.3.1 Methodology 105
 - 2.3.2 Catalog of Models 112
 - 2.3.3 Next Steps 118
- 2.4 SAP’s Advisory Methodologies 119
 - 2.4.1 SAP Application Extension Methodology 120
 - 2.4.2 SAP Integration Solution Advisory Methodology 121
 - 2.4.3 SAP Data and Analytics Advisory Methodology 123
- 2.5 Summary 124

- 3 Overview of SAP Business Technology Platform** 127
- 3.1 Fundamentals 128
 - 3.1.1 Global Account 130
 - 3.1.2 Directories 130
 - 3.1.3 Subaccounts 131
 - 3.1.4 Regions and Providers 131
 - 3.1.5 Environments 133
 - 3.1.6 Entitlements and Quotas 135
 - 3.1.7 Services 138
- 3.2 Data and Analytics 139
 - 3.2.1 Data Products 139
 - 3.2.2 SAP Business Data Cloud Cockpit 140
 - 3.2.3 SAP Datasphere 140
 - 3.2.4 SAP Analytics Cloud 140
 - 3.2.5 SAP Databricks 141
- 3.3 Application Development 141
 - 3.3.1 SAP Build Apps 143
 - 3.3.2 SAP Build Code 143
 - 3.3.3 SAP Build Process Automation 145

3.3.4	SAP Build Work Zone	145
3.3.5	Joule Studio	146
3.4	Integration	148
3.4.1	SAP Cloud Integration	149
3.4.2	Migration Assessment	150
3.4.3	Integration Technology Guidelines	151
3.4.4	Extend Connectivity	153
3.4.5	SAP API Management	153
3.4.6	Trading Partner Management	155
3.4.7	SAP Integration Suite, Advanced Event Mesh	156
3.4.8	Access Data in Classic SAP Business Suite with OData Provisioning	157
3.5	Automation	158
3.5.1	SAP Build Process Automation	158
3.5.2	SAP Automation Pilot	159
3.6	Artificial Intelligence	160
3.6.1	SAP Business AI	160
3.6.2	SAP AI Core	161
3.6.3	SAP AI Launchpad	161
3.6.4	SAP AI Services	161
3.7	Use Cases	162
3.7.1	Goals	162
3.7.2	Solution	163
3.7.3	Overall Architecture and Integration	170
3.8	Summary	171
4	Data Quality	173
4.1	Introduction to Data Quality	174
4.1.1	Data Quality Terminology	175
4.1.2	Roles and Responsibilities in Data Quality	179
4.1.3	Categories of Data	187
4.1.4	Type of Data Quality Defects	190
4.1.5	Causes of Data Quality Defects	194
4.1.6	Data Quality Approach	198
4.1.7	Benefits Realized from Profiling and Cleansing	203
4.2	Data Profiling Approach	203
4.2.1	Introduction to Data Profiling	204
4.2.2	Data Profiling Fundamentals	208

- 4.2.3 Identify Data Risks and Issues with a Data Quality Assessment 209
- 4.2.4 When to Perform a Data Quality Assessment 211
- 4.2.5 Data Quality Assessment Process 212
- 4.2.6 Data Profiling Best Practices 215
- 4.2.7 Other Considerations 216
- 4.3 Data Cleansing Approach 219**
 - 4.3.1 Introduction to Data Cleansing 221
 - 4.3.2 Data Cleansing Fundamentals 222
 - 4.3.3 Correcting Data Quality Defects with Data Cleansing 234
 - 4.3.4 Determining How Much Cleansing Is Enough 236
 - 4.3.5 Data Cleansing Methodology 238
 - 4.3.6 Data Cleansing Best Practices 248
 - 4.3.7 Other Considerations 249
- 4.4 User Training and Enablement 250**
 - 4.4.1 SAP Learning Journeys 251
 - 4.4.2 Training Courses 252
 - 4.4.3 Other Learning Resources 253
- 4.5 Summary 254**
- 5 Master Data Management 255**
 - 5.1 Overview 256**
 - 5.2 Key Features 262**
 - 5.2.1 Consolidation 263
 - 5.2.2 Mass Processing 268
 - 5.2.3 Central Governance 268
 - 5.2.4 Data Quality Management 272
 - 5.2.5 Process Analytics 276
 - 5.2.6 Federated Master Data Governance 279
 - 5.3 Architecture and Deployment Models 282**
 - 5.3.1 Choosing a Deployment Model 283
 - 5.3.2 SAP Master Data Governance on SAP S/4HANA Cloud,
Private Edition 286
 - 5.3.3 SAP Master Data Governance, Cloud Edition 290
 - 5.3.4 SAP S/4HANA Cloud Public Edition, Master Data Governance 290
 - 5.4 Master Data Integration 292**
 - 5.5 Enabling a Business Data Fabric 293**
 - 5.6 User Training and Enablement 294**
 - 5.7 Summary 295**

6	Data Storage with SAP HANA Cloud	297
6.1	Key Features	298
6.1.1	Data Storage Capabilities	298
6.1.2	Data Lake Capabilities	302
6.1.3	Real-Time Data Access	303
6.1.4	Native Multi-Model	303
6.1.5	Built-In Machine Learning, Predictive Analytics, and Search	310
6.1.6	Security	312
6.2	Use Cases	315
6.2.1	Integration with Business Intelligence Tools	315
6.2.2	Predictive Analytics	316
6.2.3	Transition to Cloud	316
6.3	User Training and Enablement	317
6.3.1	Training Courses	317
6.3.2	SAP Learning Journeys	318
6.3.3	Learning Resources	318
6.4	Summary	320
7	Data Fabric with SAP Business Data Cloud	321
7.1	Architecture	322
7.1.1	Source Systems	323
7.1.2	Foundation Services	323
7.1.3	Data Products	323
7.1.4	Delta Share Protocol	326
7.1.5	Onboarding	327
7.2	SAP Business Data Cloud Cockpit	329
7.3	SAP Datasphere	331
7.3.1	Architecture	331
7.3.2	Space and Connection Management	334
7.3.3	Modeling: Data Builder	338
7.3.4	Modeling: Business Builder	341
7.3.5	Data Integration Monitor	342
7.3.6	Catalog and Marketplace	343
7.4	SAP Business Warehouse and SAP BW/4HANA	345
7.5	SAP Databricks	347
7.6	SAP Analytics Cloud	350
7.7	Business Use Cases	353
7.7.1	Solution Overview	353

- 7.7.2 Data Integration and Modeling 354
- 7.7.3 C-Level Dashboards 355
- 7.7.4 Machine Learning Applications 356
- 7.7.5 Governance and Scalability 357
- 7.7.6 Conclusion 358
- 7.8 Summary 359

- 8 Data-Driven Decision-Making with SAP Analytics Cloud 361**
- 8.1 Introduction 362
 - 8.1.1 Data and Analytics Strategy 363
 - 8.1.2 Business Cases for SAP Analytics Cloud 363
 - 8.1.3 SAP Analytics Cloud Value Proposition 364
- 8.2 SAP Analytics Cloud Capabilities 366
 - 8.2.1 Analytical Capabilities 366
 - 8.2.2 Planning Capabilities 376
 - 8.2.3 AI Capabilities 383
 - 8.2.4 Integration with Microsoft Office 390
- 8.3 Decision-Making with SAP Analytics Cloud 391
 - 8.3.1 Operational 391
 - 8.3.2 Tactical 392
 - 8.3.3 Strategic 394
 - 8.3.4 On-the-Go 395
- 8.4 Sharing, Collaborating, and Exporting 396
 - 8.4.1 Sharing and Publishing Content 396
 - 8.4.2 Commenting and Discussing 396
 - 8.4.3 Exporting Content 397
- 8.5 Business Content and Accelerators 398
 - 8.5.1 Packages, Content, and Templates 399
 - 8.5.2 Getting Business Content 402
 - 8.5.3 Business Content Usage 403
- 8.6 User Training and Enablement 403
 - 8.6.1 Training Courses 403
 - 8.6.2 SAP Learning Journeys 404
 - 8.6.3 Learning Resources 404
- 8.7 Best Practices 405
- 8.8 Summary 407

9	Artificial Intelligence for Data and Analytics	409
9.1	AI Concepts	410
9.1.1	AI Models	411
9.1.2	AI Workflow and Interface	419
9.1.3	Additional AI Considerations	425
9.2	AI in SAP Solutions	431
9.2.1	Joule	432
9.2.2	Embedded AI	433
9.2.3	AI Foundation	435
9.3	Integration with SAP Products	440
9.4	SAP Joule Assistance During Implementation	442
9.5	User Training and Enablement	445
9.5.1	SAP Discovery Center	445
9.5.2	SAP Tutorials	448
9.6	Summary	449
10	Business and Integration Scenarios	451
10.1	Sustainable Enterprise	452
10.1.1	Significance of Addressing the Challenge	452
10.1.2	Overview of Potential Solutions	452
10.1.3	Components Contributing to the Solution	452
10.2	Ethical Supply Chain	454
10.2.1	Significance of Addressing the Challenge	454
10.2.2	Overview of Potential Solutions	454
10.2.3	Components Contributing to the Solution	454
10.3	Managing Private and Secure Data	455
10.3.1	Significance of Addressing the Challenge	455
10.3.2	Overview of Potential Solutions	456
10.3.3	Components Contributing to the Solution	456
10.4	Fraud Detection	457
10.4.1	Significance of Addressing the Challenge	458
10.4.2	Overview of Potential Solutions	458
10.4.3	Components Contributing to the Solution	458
10.5	Predictive Maintenance	460
10.5.1	Significance of Addressing the Challenge	460
10.5.2	Overview of Potential Solutions	461
10.5.3	Components Contributing to the Solution	461

10.6 Data Quality Firewall	462
10.6.1 Significance of Addressing the Challenge	463
10.6.2 Overview of Potential Solutions	463
10.6.3 Components Contributing to the Solution	464
10.7 Guided End-User Priorities and Application Navigation	465
10.7.1 Significance of Addressing the Challenge	465
10.7.2 Overview of Potential Solutions	466
10.7.3 Components Contributing to the Solution	467
10.8 Custom Operating Guides and User Documentation	468
10.8.1 Significance of Addressing the Challenge	468
10.8.2 Overview of Potential Solutions	469
10.8.3 Components Contributing to the Solution	469
10.9 Summary	471
The Authors	473
Index	477

Chapter 7

Data Fabric with SAP Business Data Cloud

SAP launched SAP Business Data Cloud in February 2025. This comprehensive software as a service (SaaS) solution features an object store for managing vast amounts of data at a lower cost. To tackle the crucial issue of business context, SAP introduced self-sustaining data products powered by artificial intelligence (AI). Furthermore, recognizing the evolving machine learning (ML) landscape, SAP partnered with Databricks to create SAP Databricks, an original equipment manufacturer (OEM) service included within SAP Business Data Cloud, bringing ML capabilities directly to organizations.

SAP Business Data Cloud represents a strategic evolution of SAP's data strategy. This SaaS offering on SAP Business Technology Platform (SAP BTP) combines the existing SAP Datasphere and SAP Analytics Cloud solutions and adds new technological stacks for low-cost storage and ML with SAP Databricks. AI-powered intelligent applications and data products provide advanced cross-analytics for actionable insights. SAP Business Data Cloud is comprised of the following:

- Foundation services
- SAP Analytics Cloud
- SAP Datasphere
- SAP Databricks
- SAP Business Warehouse (SAP BW) (optional)

For SAP BW organizations, SAP Business Data Cloud enables a cloud transition without disruption with extended support up to 2030. SAP BW data can be converted to structured data products for analytics.

The strategic goals of SAP Business Data Cloud are as follows:

- Provide curated access to SAP S/4HANA and line of business (LoB) data without having to build complex extraction processes.
- Provide a stable migration path for existing SAP BW organizations.
- Provide a seamless conversion for all existing SAP Datasphere implementations to SAP Business Data Cloud implementations.
- Provide the SAP Business Data Cloud offering as a SaaS, which can be easily subscribed to.

- Create a strong library of data products using a common modeling paradigm. Incentivize partners to create custom data products and establish a data product marketplace.
- Support ML use cases with the introduction of SAP Databricks.
- Enable quick, simplified, and seamless access to high-performing analytics in a unified offering.

The goal of this chapter is to help you get to know SAP Business Data Cloud and its various components. New concepts and terms will be explained that will aid in your move to SAP Business Data Cloud. We'll start with its architecture and then drill down into its components. We'll explain the concepts of a data fabric, data products and their various flavors, foundation services, intelligent applications, and zero-copy Delta Share. Next, we'll talk about the main components of SAP Business Data Cloud, which are the SAP Business Data Cloud cockpit in Section 7.2, SAP Datasphere in Section 7.3, SAP BW in Section 7.4, SAP Databricks in Section 7.5, and finally SAP Analytics Cloud in Section 7.6. Business use cases are discussed in Section 7.7.

7.1 Architecture

Figure 7.1 shows the high-level architecture of SAP Business Data Cloud and aims to explain all the inter-connected components that make up this offering. Each component brings unique capabilities and contributes toward a comprehensive cloud solution that combines data storage, data modeling, analytics, and ML in one place.

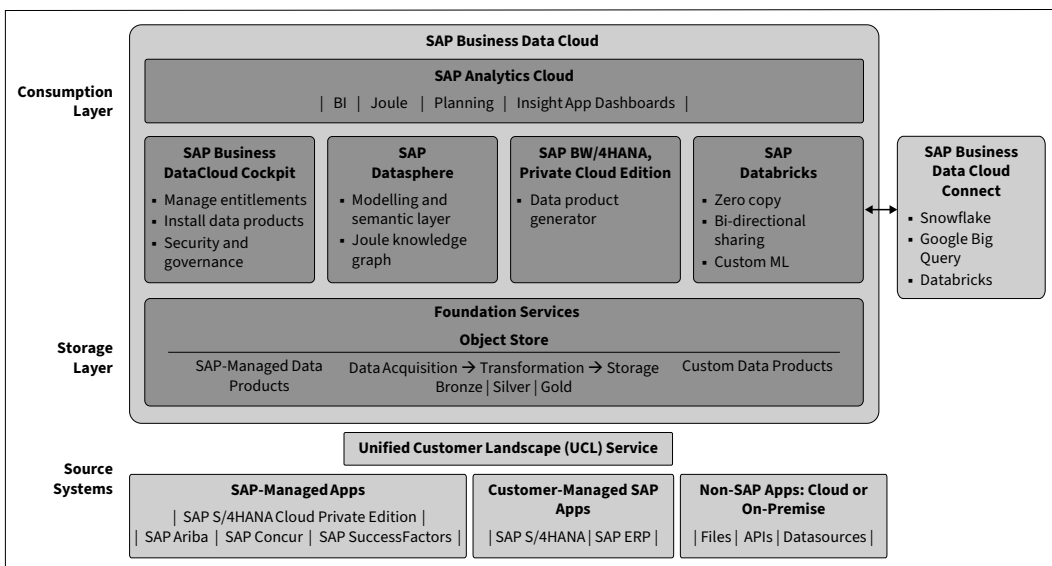


Figure 7.1: SAP Business Data Cloud Architecture

First, we'll talk about the source systems in Section 7.1.1, which can be made a part of the SAP Business Data Cloud formation. Next, we'll talk about the newly introduced stack of

foundation services in Section 7.1.2. Data products are the cornerstone of the SAP Business Data Cloud strategy and are elaborated on in Section 7.1.3. Data sharing efficiency introduced by the Delta Share protocol is explained in Section 7.1.4. Finally, onboarding an organization on SAP Business Data Cloud is elaborated on in Section 7.1.5.

7.1.1 Source Systems

The bottom-most part of the architecture diagram displays the various source systems. Data from these source systems can be brought into the SAP Business Data Cloud landscape via prebuilt data products and a connection to the source system in the landscape formation (this is shown later in this chapter in Figure 7.4). *Unified Customer Landscape* (UCL) is a collection of central foundation services that are used to create a *formation*—a logical grouping—of an organization’s source SAP systems (SAP S/4HANA, LoB) within SAP Business Data Cloud. The systems trust each other when in a formation, making data exchange seamless. These services also play a role in sharing the metadata of data products with consumers such as the SAP Business Data Cloud cockpit, SAP Datasphere catalog, and SAP Business Accelerator Hub. Connection to various third-party systems can also happen at the SAP Datasphere level through the native adapter capability within SAP Datasphere.

7.1.2 Foundation Services

At the base of SAP Business Data Cloud lies its foundation services, an internal, customer-invisible component that is crucial for delivering SAP Business Data Cloud’s innovative data products and establishing core cloud data platform capabilities. The foundation services handle the end-to-end creation of these data products internally and replicates datasets from formation source systems into the object store, making them ready for consumption. Key functions of these services include management operations (streamlining onboarding and configuration), data integration (seamlessly transforming data from various SAP solutions), data governance and storage (securely and efficiently storing vast data volumes), data lineage and monitoring (continuously overseeing data pipelines for accuracy), and usage metering and cross-charging (supporting SAP Business Data Cloud’s commercial objectives).

The technical backbone of the foundation services consists of SAP HANA Cloud, data lake files (these act as an object store governance layer using native hyperscaler object storage such as Amazon Web Services [AWS] S3), a managed Spark service for computation, and SAP HANA Cloud serving as a database sidecar. This robust foundation underpins SAP’s comprehensive data strategy, enabling seamless data sharing and integration.

7.1.3 Data Products

Data products aren’t new or unique to SAP Business Data Cloud. They have been used in the industry for a long time. However, they indicate a new data integration unit that gives

access to a specific set of business data with minimal or extensive add-on capabilities built in. Operational data is treated as a product with a consistent, well-defined structure and interfaces, supported by lifecycle management, metadata, and documentation. This enables easy, reliable, and secure consumption of the data. The data product owner is responsible for managing any changes to the data preparation process.

With data products, you can simply use the provided data and tailor it to your specific needs. Data products have owners, typically domain owners or the team managing the operational data, who are responsible for defining the data products to meet consumer needs. The data product provider handles all the extraction, harmonization, and integration of various tables. They follow a standardized process that includes needs analysis, design, and delivery phases, ensuring that the data is accurate and in the correct format. The data product includes a self-description provided by the producer using *Open Resource Discovery*. Data products are well described, easily consumable business datasets that can be easily discovered.

Data products in the SAP Business Data Cloud context can exist at the source system and foundation services level and must be pure data products only. This means the data products have been created specific to an SAP application such as SAP S/4HANA, and all they do is bring forth the data from SAP S/4HANA with the business context for consumption and analytics without any organizational investment in building an extract, transform, and load (ETL) mechanism. Each data product source may use different technologies to build the data product. For example, SAP S/4HANA data products use ABAP core data services (CDS) views to combine several tables into a harmonized view. The dataset is designed for intensive reads and is accessed in a read-only format.

Examples of SAP S/4HANA data products include country/region, SAP HANA currency, cost center, journal entry header, ledger, and financial planning entry item. For instance, the financial planning entry item data product combines entity types such as the financial planning entry item and planning category.

Other SAP LoB solutions, such as SAP SuccessFactors, SAP Ariba, SAP Concur, and SAP Fieldglass, are also envisioned to have data products following the SAP One Domain Model. This harmonization allows easy combination of factual data from these LoBs with SAP S/4HANA data using common entities such as customer or material.

Data products are organized into *data packages*. A data package combines two or more data products. A data package might include, for example, both country/region and SAP HANA currency data products as cross-application items. Another data package could encompass all the data products related to accounting and financial close, such as journal entry header and financial planning entry item. Figure 7.5 later in the chapter shows the SAP Business Data Cloud cockpit and a list of the available data products and intelligent applications.

Tip

A list of data products and descriptions is available on the SAP Business Accelerator Hub at <https://hub.sap.com/dataproducts>.

An *intelligent application* is a high-level data product that integrates interconnected data products, base data models, SAP Datasphere analytical models, and analytical dashboards to provide actionable insights. Intelligent applications are adaptive, AI-driven applications that use your business data to learn, reason, and perform actions. These fully managed applications are installed via the SAP Business Data Cloud cockpit, automatically deploying and configuring data products, data models, analytics, and application assets. The prebuilt, SAP-managed data products included in these application packages enable rapid development of new insights and accelerate AI innovation. Figure 7.2 shows the full packaging pyramid of an intelligent app. At the top are the predefined AI-infused analytics in the form of SAP Analytics Cloud visualizations or SAP Cloud Application Programming Model applications. SAP Datasphere analytical or base models form the center core. Base data products pulling data from the source systems are the base layer.

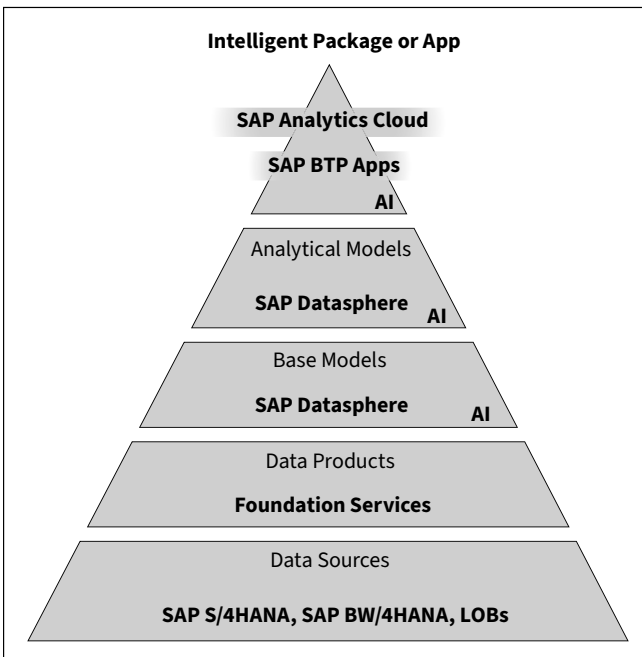


Figure 7.2: Intelligent Application

The key components of an intelligent application include the following:

- **Data products**

The foundation of an intelligent application is business data sourced from various systems. During the installation of an intelligent app, data is replicated from source systems

to SAP Business Data Cloud and the data lake, where one or more data products are created.

- **Data models**

Built on top of data products, data model artifacts enable model extensions and user authorization definitions. These models are deployed in a read-only SAP-managed secure space.

- **Dashboards**

The analytical dashboard forms the top layer of an intelligent app, offering features such as forecasting, AI-powered decision support, insight-to-action capabilities, and planning tools. This capability is provided by SAP Analytics Cloud.

Installation of an intelligent application leads to an automated deployment of the underlying data products. This helps organizations get quick no-effort access to data and discover actionable insights quickly. Managing the complexity of the underlying models is taken out of the equation and managed by SAP. Updates and enhancements can be received over time.

With the data product strategy, SAP simplifies the availability of LoB application data by taking care of the extraction, curation and storage of it in the distributed data lake of the foundation services. Organizations don't have to worry about understanding the source systems, knowing where to extract data from, or building and maintaining extraction routines. To gain access to LoB data products, organizations must purchase the corresponding SAP LoB application. For example, to gain access to the SAP SuccessFactors data products, the organization must have a subscription to SAP SuccessFactors and include it in the SAP Business Data Cloud formation.

7.1.4 Delta Share Protocol

The base storage layer in SAP Business Data Cloud is an object store of data lake files. Data is stored in the format of Delta Lake files, not tables. This provides for sharing a large amount of data in an efficient way. This storage also extends the number of tools that can access data. Delta Lake is an open-source file format that provides significantly faster metadata operations for large tabular datasets. It offers relational-like features such as caching mechanisms and specific SQL-on-files commands.

Tip: Delta Share

You can learn more details about Delta Share from the Delta Lake website at <https://delta.io/>.

Delta Share is an open-source framework for secure, real-time data sharing across organizations and platforms. Data stored in the Delta Lake format can be accessed via the Delta Share protocol. It enables efficient sharing of extensive datasets through a streamlined

REST interface. It uses modern cloud storage systems, such as AWS S3, Azure Data Lake Storage, Google Cloud Storage, or Cloudflare R2, to reliably transfer data without requiring data replication. This capability plays a key role in sharing data between the foundation services and SAP Databricks for ML scenarios.

Shared data is primarily accessed in a read-only format, supporting various query types, including batch, streaming, and change data feed for tracking table modifications, though direct writes to shared tables aren't permitted, thus ensuring data integrity. When data is shared in the Delta Lake format, the protocol supports advanced features such as time travel queries, allowing users to query specific table versions, metadata, and the latest snapshots, along with limit and predicate pushdown for optimized data retrieval.

The protocol offers strong data management and tracking features, empowering data providers to control access permissions, grant specific recipient access, and track usage. Integration with platforms such as the SAP Databricks Unity Catalog further enhances centralized management and auditing. Designed for scalability, Delta Sharing efficiently handles terabyte-scale datasets by using cloud storage. Its flexible client support means users can access data via connectors such as Python (for Pandas or Spark DataFrames) or Apache Spark (supporting SQL, Python, Java, Scala, or R), or even build custom clients using the REST API. Finally, real-time synchronization is a key advantage; changes to shared Delta Lake tables are recorded in metadata, ensuring recipients have immediate access to the latest updates, which is crucial for streaming analytics and timely decision-making.

SAP Business Data Cloud Connect is the new connectivity layer available in SAP Business Data Cloud that provides bidirectional, zero-copy data and metadata sharing between SAP Business Data Cloud and external platforms such as Snowflake, Databricks, and Google Big-Query. Specifically for Snowflake, this is presented as SAP Business Data Cloud Connect for Snowflake, and it allows an existing Snowflake instance to consume semantically rich SAP data products in real time, as well as share Snowflake data back into SAP Business Data Cloud without building ETL-style replication.

7.1.5 Onboarding

To get onboarded onto SAP Business Data Cloud, a number of steps happen behind the scenes. Figure 7.3 shows these steps, from provisioning the SAP Business Data Cloud tenant to consuming intelligent applications.

Pricing

For pricing the capacity units for SAP Business Data Cloud, you can use the Estimator tool at <https://bdc-pricing-estimator-sac-sacus10.cfapps.us10.hana.ondemand.com/>.

At a minimum, 640 capacity units is required for an SAP Business Data Cloud core subscription. This presents the smallest possible use case with 25 SAP Analytics Cloud BI users.

Note

SAP Business Data Cloud is offered through subscription only. It isn't available through SAP BTP Enterprise Agreement (BTPEA) or SAP Cloud Platform Enterprise Agreement (CPEA).

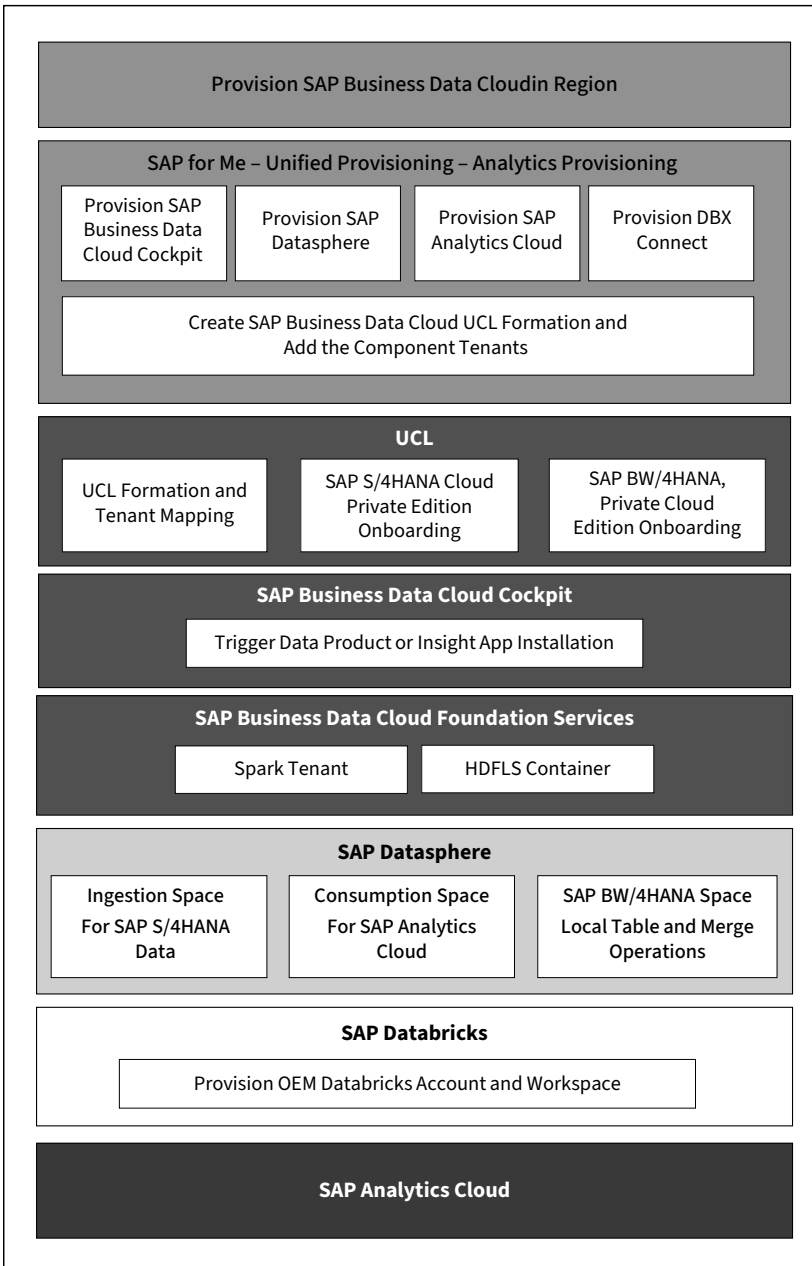


Figure 7.3: SAP Business Data Cloud Organization Onboarding

Following is a high-level overview of the necessary steps to onboard onto SAP Business Data Cloud:

1. Deployment of SAP Business Data Cloud in a region
2. SAP Business Data Cloud onboarding
3. SAP S/4HANA Cloud Private Edition client onboarding
4. SAP BW/4HANA, private cloud edition onboarding
5. Provision SAP Datasphere
6. Provision SAP Analytics Cloud
7. Provision SAP Databricks
8. Creation of the SAP Business Data Cloud unified organization landscape formation
9. Installation of intelligent application or data package activation
10. Consumption of the intelligent app

7.2 SAP Business Data Cloud Cockpit

The SAP Business Data Cloud cockpit is the central point of access for SAP Business Data Cloud for displaying, configuring, and finding any artifacts that interact with the SAP Business Data Cloud landscape.

The first step when using SAP Business Data Cloud is to set up the formation in the landscape directory. This involves making connections to the various source systems using SAP for Me and using the SAP BTP unified organization landscape services in the background. An already set up landscape is shown in Figure 7.4.

System Landscape	
Basic Trial BDC (4)	
System Name	System Type
trial-bdc-datasphere.eu10	SAP Datasphere (cf-eu10)
HE4_400_BDC_BasicTrial_2	SAP S/4HANA Cloud Private Edition
trial-bdc-analytics.eu10	SAP Analytics Cloud (cf-eu10)
databricks-ee9691c9	SAP Databricks

Figure 7.4: SAP Business Data Cloud System Landscape (Formation)

The main components as you can see in Figure 7.4 are SAP Datasphere, SAP S/4HANA Cloud Private Edition, SAP Analytics Cloud, and SAP Databricks. Not all of these systems are mandatory to be included in the landscape, but these four components give the maximum SAP Business Data Cloud capabilities to an organization. SAP BW is optional and can be included if you have it in your landscape.

Intelligent applications deploy analytical stories and models that can be used only if you have SAP Datasphere and SAP Analytics Cloud in your formation. Similarly, SAP Databricks isn't mandatory to have if advanced ML capabilities aren't immediately needed.

SAP Datasphere, SAP Analytics Cloud, and the SAP Business Data Cloud cockpit must be co-located on the same data center, but SAP BW/4HANA can remain in a different data center on a private cloud. Attempts should be made to co-locate SAP BW/4HANA, private cloud edition and SAP S/4HANA Cloud Private Edition.

Using the SAP Business Data Cloud cockpit, you can see a list of or search for available data packages or intelligent applications for usage, as shown in Figure 7.5.

Intelligent Applications and Data Packages					
Available (11)		Installed (8)			
Available applications and packages (11)					
Name	Type	Category	Version	System Type	Minimum System Version
Working Capital Insights	Intelligent Application	Finance	1.3.0	SAP S/4 HANA Cloud Private Edition	2022.0.0
SAP Accounting and Financial Close Data Pro...	Data Package	Finance	1.2.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Cross Applications Data Products (SAP Si...	Data Package	Other	1.1.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Database and Data Management Data Pro...	Data Package	Other	1.1.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP EHS Environment, Health and Safety and ...	Data Package	Other	1.0.0	SAP S/4 HANA Cloud Private Edition	2023.0.0
SAP Financial Operations Data Products (SAP ...	Data Package	Finance	1.2.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Group Reporting Data Products (SAP S/4...	Data Package	Finance	1.2.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Manufacturing Data Products (SAP S/4HA...	Data Package	Other	1.2.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Sourcing and Procurement Data Products...	Data Package	Other	1.2.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Supply Chain Data Products (SAP S/4HA...	Data Package	Other	1.1.0	SAP S/4 HANA Cloud Private Edition	2021.0.0
SAP Sales Data Products (SAP S/4HANA Clou...	Data Package	Unknown	1.1.0	SAP S/4 HANA Cloud Private Edition	2023.0.0

Figure 7.5: Available Intelligent Applications and Data Packages

The intelligent applications collection contains adaptive, AI-driven applications that use your business data to learn, reason, and perform actions. These fully managed applications are installed via the SAP Business Data Cloud cockpit, automatically deploying and configuring data products, data models, analytics, and application assets. The prebuilt, SAP-managed data products included in these application packages enable rapid development of new insights and accelerate AI innovation. Intelligent applications can't be consumed by non-SAP products.

Data packages are collections of related data products that can be installed in the SAP Business Data Cloud landscape. Installing a single data package also installs all the data

products contained within it. Data packages focus on data provisioning, that is, providing access to curated, semantically rich data. They don't include pre-created analytical stories or dashboards.

7.3 SAP Datasphere

Launched on March 8, 2023, SAP Datasphere is SAP's cloud data warehousing solution, succeeding SAP Data Warehouse Cloud. Built on SAP BTP and powered by SAP HANA Cloud, SAP Datasphere provides a unified experience for business-critical data from both SAP and non-SAP systems. It integrates data, offers cataloging and semantic modeling, and handles warehousing, federation, and virtualization. This comprehensive approach empowers organizations to make agile, data-driven decisions. More than just a data warehouse, SAP Datasphere offers a broader vision that includes data integration, governance, and real-time access, connecting diverse data landscapes seamlessly.

SAP introduced the *data fabric architecture* with SAP Datasphere, and this continues with SAP Business Data Cloud. The key concept with the data fabric was to provide seamless, governed, and unified access to distributed SAP data while maintaining business context and minimizing data duplication. SAP refers to this as the business data fabric, and it's central to the vision of SAP Business Data Cloud.

SAP Datasphere is a key component and is at the heart of SAP Business Data Cloud. Hence, we'll discuss the capabilities of SAP Datasphere and its workings in more detail. We'll talk about the SAP Datasphere layered architecture in Section 7.3.1. Spaces and connectivity options are discussed in Section 7.3.2. SAP Datasphere provides a rich set of semantic modeling features, which are covered in Section 7.3.3 and Section 7.3.4. Monitoring SAP Datasphere for performance and capacity is discussed in Section 7.3.5. Finally, cataloging capabilities and how the marketplace can be used to extend functionality is covered in Section 7.3.6.

7.3.1 Architecture

At the heart of SAP Datasphere is its cloud-native foundation: SAP BTP. SAP BTP delivers the essential infrastructure and platform services—including security, identity management, and integration frameworks—that empower SAP Datasphere's capabilities. This ensures the solution is scalable, secure, and enterprise-ready.

Driving the processing and storage layer is SAP HANA Cloud, a high-performance in-memory database that enables real-time analytics and complex data transformations, and uses its in-memory processing for high performance. Key capabilities such as columnar storage, parallel processing, and built-in ML libraries—Predictive Analysis Library (PAL) and Automated Predictive Library (APL)—make SAP HANA Cloud a cornerstone of SAP Datasphere.

The platform is designed to support a modular, cloud-native environment that integrates diverse data sources and tools. Figure 7.6 shows a high-level overview of the SAP Datasphere architectural components. We've attempted to align this with the SAP Datasphere UI shown in Figure 7.7 so you can make connections with the components shown on the homepage when logged in to SAP Datasphere. Certain components may require multiple navigational steps to get to.

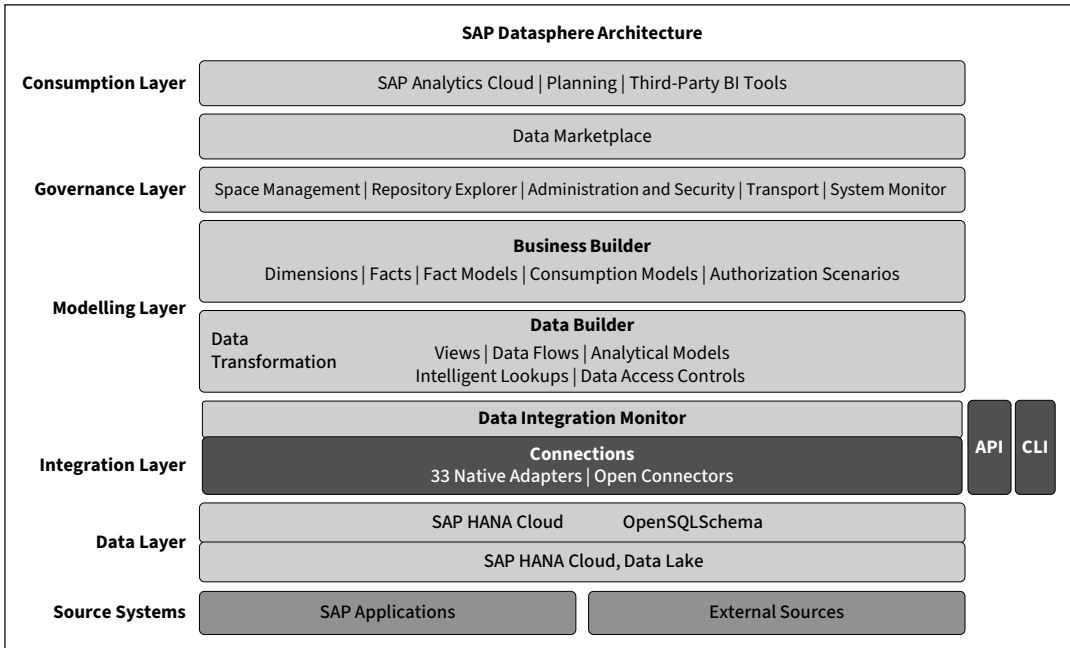


Figure 7.6: SAP Datasphere Architecture

Let's look at the core components of SAP Datasphere in a little more detail:

■ Data layer

- SAP HANA Cloud: The core database service for storing and processing structured and semi-structured data underneath SAP Datasphere.
- Open SQL schema: A schema created within SAP HANA Cloud for enabling data storage for third-party client access to BI or ETL tools.
- SAP HANA Cloud, data lake files: An object store governance layer that integrates with native cloud storage (e.g., AWS S3, Azure Data Lake, Google Cloud Storage), provides an additional low-cost storage layer, and can only be enabled for one space within SAP Datasphere.

■ Integration layer

- Connections: Built-in connectors for SAP (e.g., SAP S/4HANA, SAP ERP) and non-SAP sources (e.g., Microsoft Azure, Google BigQuery, Oracle). These support federation, replication, and real-time data access. There are 33 native adapters, and the rest can

be supported via open connectors. This supports the integration of SAP applications and multiple external sources.

- Data Integration Monitor: Provides visibility into data flows, replication schedules, and persisted tables.
- Replication flows: Enable batch and delta data loading from source systems with a single flow, reducing ETL complexity.

■ Modeling layer

- Data Builder: A graphical, low-code/no-code interface for technical users to acquire, prepare, and model data. It supports SQL modeling, semantic layers, and data previews.
- Business Builder: A semantic layer for business users to define business entities, key performance indicators (KPIs), and consumption models without technical expertise.
- Repository Explorer: Allows users to browse and filter existing models (e.g., dimensions, facts) with associated metadata.

■ Consumption layer

- SAP Analytics Cloud: Tightly integrated for reporting, dashboards, and planning. SAP Analytics Cloud uses SAP Datasphere’s data models for real-time analytics.
- Third-party tools: Supports integration with Tableau, Power BI, and other visualization tools via OData APIs.
- Data marketplace: A platform to discover, purchase, and integrate third-party data products into SAP Datasphere spaces.

■ Governance and security

- Spaces: Virtual, secure workspaces for different departments or use cases, with allocated storage, user roles, and monitoring tools.
- Data catalog: Centralized metadata repository for data lineage, data quality, and data discovery. Organizes data and metadata, enabling confident decision-making with centralized governance.
- Security features: Includes encryption, row-level security, data masking, and space-dependent permissions (scoped roles).

SAP Datasphere operates as a public data warehouse as a service (DWaaS) on SAP BTP, offering fast release cycles and automatic updates. It supports hybrid deployments, allowing access to on-premise data via federation or replication, and to multicloud environments for flexibility. We’ll also talk more in depth about the capabilities of SAP Datasphere that we touched on in Section 7.1.

You can provision SAP Datasphere as a part of your SAP Business Data Cloud subscription. When you log in to SAP Datasphere, the homepage appears as shown in Figure 7.7 with all of its capabilities listed in the left menu.

The screenshot shows the SAP Datasphere Home page. At the top, there is a notification banner: "We made updates since you were last here. Click here to find out more." Below this, the user is greeted with "Hello, Asim Munshi" and the date "Jun 27, 2025".

The main content area features a large banner with the text "Unleash the power of business data" and a "Welcome to SAP Datasphere" section. The welcome message states: "SAP Datasphere enables a business data fabric architecture that uniquely harmonizes mission-critical data across the organization, unleashing business experts to make the most impactful decisions. It combines previously discrete capabilities into a unified service for data integration, cataloging, semantic modeling, data warehousing, and virtualizing workloads across SAP and non-SAP data." A "Learn More" button is present.

Below the welcome message is a "Spaces" section titled "Storage / Status - Top 5". It contains a table with the following data:

Name	Priority	Storage	Status
Finance	8	Used 9.24 GB of 12 GB	Working
HR	5	Used 1.03 GB of 4 GB	Working

On the right side, there are two sections: "SAP Datasphere Blog" with recent posts and "Data Integration Tasks" showing recently started tasks, including "DS_TO_TCMP_STAGE" and "Data_Flow_1_AWS".

Figure 7.7: SAP Datasphere Homepage

7.3.2 Space and Connection Management

Deployment of data products results in the creation of SAP-managed spaces on SAP Datasphere. Think of a *space* in SAP Datasphere as a secure logical area created by an administrator, which holds a collection of SAP Datasphere artifacts. For example, you may have a space catering to finance and a separate space catering to human resources. The concept is similar to the spaces concept discussed in Chapter 3. Space members can acquire, prepare, and model data in the space. The administrator allocates disk and in-memory storage to the space, sets its priority, and limits how much memory and how many threads its statements can consume. A space can be thought of as a small data mart for a specific business segment. Technically, a space is created as a schema within the underlying SAP HANA Cloud layer. SAP Datasphere allows for the controlled sharing of artifact data between spaces.

When an intelligent application is installed, special spaces called *ingestion spaces* and *consumption spaces* are created in SAP Datasphere to facilitate the ingestion, preparation, and exposure of data. These spaces are specifically designed to contain replication flows, tables, views, and analytic models required for the data product. Ingestion spaces use the SAP HANA Cloud, data lake files as the underlying object store, which is fully managed by SAP and transparent to organizations and system integrators. This ensures scalability and cost-effectiveness for large data volumes. In the context of SAP BW integrations, a dedicated SAP BW space is created for each InfoProvider subscribed to. Figure 7.8 shows the

SAP S/4HANA Ingestion space and the intelligent application **Working Capital** insight consumption space. These spaces are automatically created on the installation of the intelligent application, and objects in these spaces can't be modified.

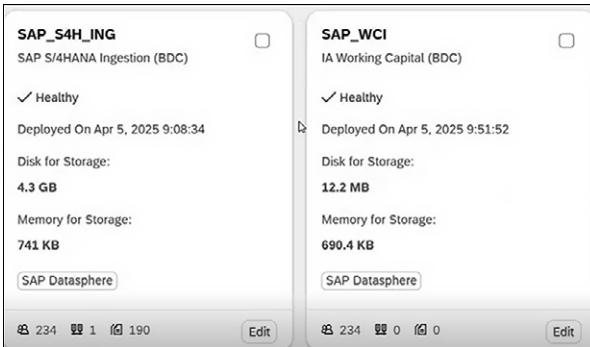


Figure 7.8: Ingestion and Consumption Intelligent Application Spaces

SAP Datasphere creates a data fabric architecture by protecting your investments in pre-existing products and by providing rich connectivity to varied systems. This allows you to connect to multiple SAP and non-SAP systems with the option of federating data in a single place to combine it, creating a true data fabric. For SAP systems, if data products have been released, they serve as the preferred way of accessing data; however, where data products aren't available, the SAP Datasphere native connectivity can be used.

Connections created are confined to the space they are created in and accessible only to the artifacts created in that space. Figure 7.9 shows the non-SAP native connectors available within SAP Datasphere. Note also all the **Features** you can filter the connections by: **Data Flows**, **Model Import**, **Remote Tables**, and **Replication Flows**.

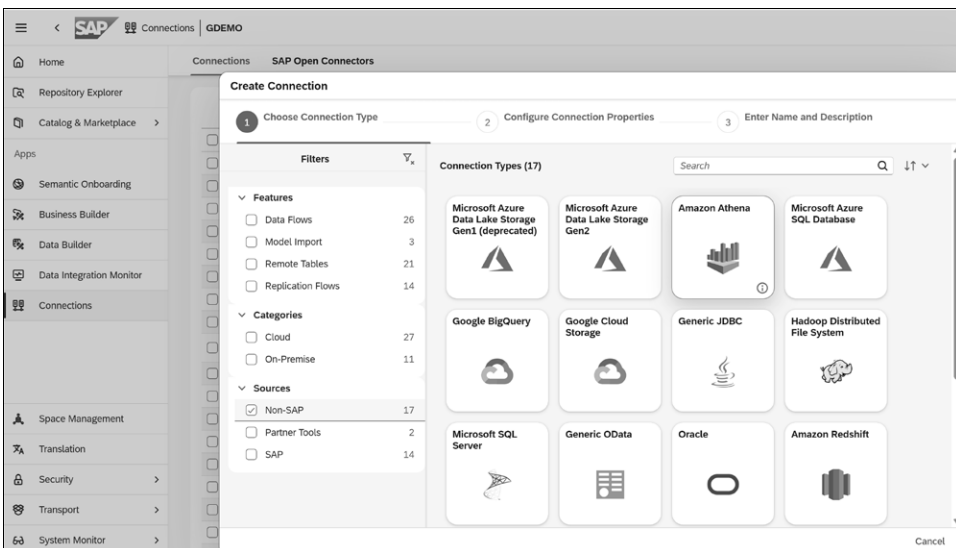


Figure 7.9: SAP Datasphere Native Connections

Let's first talk about the connectivity options, and then we'll cover these features. SAP Datasphere offers robust support for connecting to a wide array of data sources, both in the cloud and on-premise. It handles various integration types, including data federation (virtual access), replication (batch or delta loading), and near real-time access. This comprehensive connectivity relies on four key technologies: SAP HANA smart data integration, SAP HANA smart data access, the data provisioning agent, and the cloud connector. Let's understand the core technologies a little bit more:

- **SAP HANA smart data integration**

This is SAP Datasphere's main engine for data integration. SAP HANA smart data integration excels at real-time replication, batch processing, and transforming data. It uses adapters, hosted by the provisioning agent, to connect to a large number of SAP and non-SAP sources, making it simple to ingest data into SAP Datasphere for analytics. For example, you could use SAP HANA smart data integration to replicate customer data from Salesforce, including delta updates to capture changes.

- **SAP HANA smart data access**

SAP HANA smart data access focuses on data federation, which means you can access data from remote sources without actually moving or replicating it. It creates virtual tables in SAP Datasphere that map directly to your source tables. This is particularly useful for real-time reporting where you need the most up-to-date information, such as querying live inventory data from an on-premise SAP ERP or SAP S/4HANA system. SAP HANA smart data access is great for scenarios where data movement isn't necessary, saving on storage costs and reducing latency.

- **SAP HANA smart data integration, provisioning agent**

The provisioning agent is a lightweight component that runs on your on-premise network. It acts as a secure bridge between SAP Datasphere and your on-premise data sources. The provisioning agent hosts the necessary SAP HANA smart data integration adapters (like those for ABAP, Online Database Connectivity [ODBC], or Oracle) and establishes secure connections. It's crucial for both SAP HANA smart data integration-based replication and SAP HANA smart data access-based federation, ensuring secure data access from your on-premise systems. For instance, the provisioning agent would connect an on-premise Oracle database to SAP Datasphere for batch replication of financial data.

- **Cloud connector**

This component creates a secure tunnel between your on-premise systems and SAP BTP, which includes SAP Datasphere. It's primarily used for secure application-level integration, handling authentication and encryption. The cloud connector is vital for connecting specific SAP systems, such as importing metadata from SAP BW/4HANA or accessing ABAP CDS views from SAP S/4HANA. It can also support some replication flows alongside the provisioning agent.

These technologies work in concert to provide a flexible and secure hybrid integration model:

- SAP HANA smart data integration and the provisioning agent go hand-in-hand for robust data replication and transformation, enabling SAP Datasphere to pull and persist data from various sources.
- SAP HANA smart data access uses the provisioning agent for federated access to on-premise data, allowing real-time querying without data movement.
- Cloud connector provides secure, application-level connectivity, especially for SAP systems, streamlining metadata imports and supporting specific data replication scenarios.

This unified approach allows organizations to combine federation for real-time insights with replication for deeper, persisted analytics, all while ensuring secure on-premise connectivity.

Imagine a manufacturing company using SAP Datasphere to combine its on-premise SAP S/4HANA data with cloud-based Google BigQuery data. The cloud connector and provisioning agent establish secure links to SAP S/4HANA for real-time production data via SAP HANA smart data access, while SAP HANA smart data integration replicates customer data from BigQuery into SAP Datasphere. This setup provides unified analytics, which can then be visualized in tools such as SAP Analytics Cloud.

All connections created in a space are accessible to members of that space and within the space. Each connection can support features such as the following:

- **Remote tables**

SAP Datasphere has virtual tables that provide direct, real-time access to data stored in remote source systems without physically replicating or moving the data into SAP Datasphere's storage layer. They are a key feature of SAP Datasphere's federation-first approach, using SAP HANA smart data access to enable seamless querying of external data sources while maintaining performance and minimizing storage costs. Queries on remote tables are executed directly on the source system, with results retrieved in real time.

Remote tables are managed within spaces and security features such as encryption and row-level security ensure compliance and data protection.

- **Model import**

This is a feature in SAP Datasphere designed to accelerate data modeling by reusing existing models from other SAP systems—especially SAP BW/4HANA and SAP S/4HANA. It allows users to import metadata and semantic definitions (not actual data) into SAP Datasphere, providing a fast, consistent, and governed way to build new data models without starting from scratch. No data is imported. Complex ABAP logic or SAP BW models may require manual adjustment.

- **Data flows**

Users can easily design data flows in the Data Builder using a low-code interface, defining mappings, transformations, and schedules. This allows for the harmonization of structured and semi-structured data into SAP Datasphere's internal storage (SAP HANA

Cloud or data lake). Data flows are particularly beneficial for scenarios requiring persisted data for advanced analytics, complementing the real-time access offered by remote tables. They also offer features such as automated delta updates and integration with the Data Integration Monitor for pipeline monitoring.

■ Replication flows

Replication flows in SAP Datasphere are automated, cloud-based pipelines designed to efficiently move and synchronize data from various sources to targets within the platform. They enable both initial full data loads and continuous delta updates, ensuring data is up-to-date for analytics and modeling. Supporting a wide range of sources, including SAP systems (e.g., SAP S/4HANA CDS views, Operational Data Provisioning [ODP] providers) and non-SAP systems (e.g., Salesforce, Oracle), as well as cloud storage (e.g., AWS S3), replication flows are defined in the Data Builder with user-friendly source-to-target mappings and scheduling options. Powered by SAP HANA smart data integration and replication management, they use the data provisioning agent and the cloud connector for secure on-premise connectivity. Key benefits include simplified data integration, efficient delta synchronization, and robust monitoring via the Data Integration Monitor. Unlike data flows, which focus on SAP Datasphere transformations, replication flows prioritize data replication and persistence, making them ideal for scenarios requiring consistent, high-quality data for advanced analytics.

Tip: Features by Connection

You can look up the features each connection supports at SAP Help: <http://s-prs.co/v612200>.

Now let's look at the Data Builder within SAP Datasphere and dig a little bit deeper.

7.3.3 Modeling: Data Builder

Let's start off our discussion of the model by taking a look at the Data Builder, as shown in Figure 7.10).

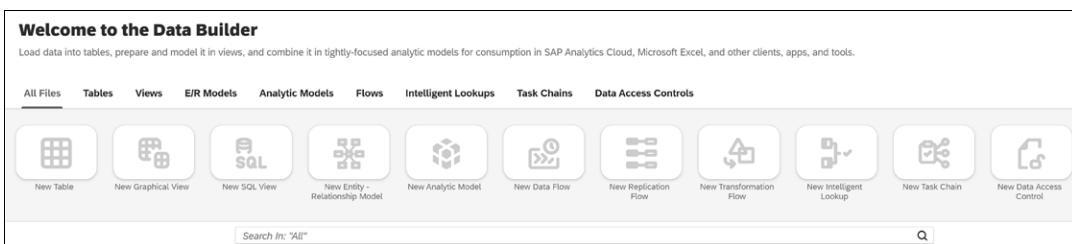


Figure 7.10: SAP Datasphere Data Builder

The Data Builder contains the following components:

- **Views (SQL and graphical)**
Supports creating both SQL views for advanced querying and graphical views for a visual, low-code approach, enabling flexible data modeling and previews for validation.
- **Entity-relationship models**
Allows users to design entity-relationship models to define relationships between data entities, facilitating structured data representation for analytics.
- **Analytics models**
Provides tools to build optimized models for consumption in SAP Analytics Cloud and other tools, supporting facts, dimensions, and KPIs for business insights.
- **Replication flows**
Enables configuration of automated pipelines for initial and delta data replication from SAP and non-SAP sources using SAP HANA smart data integration. Not all targets are supported for replication and require primary keys on sources and targets.
- **Transformation flows**
Supports creation of transformation flows to prepare and enrich data within SAP Datasphere, enhancing data quality for downstream use.
- **Data flows**
Facilitates automated ingestion, transformation, and loading of data into SAP Datasphere, supporting batch and delta updates with a user-friendly interface.
- **Intelligent lookups**
Offers data enrichment by linking datasets with common keys, improving the accuracy and depth of analytical models.
- **Task chains**
Allows sequencing of multiple tasks (e.g., data flows, transformations) for automated execution, streamlining complex workflows. It's possible to create branches that execute in parallel.
- **Data access controls**
Provides role-based security and permissions management within spaces, ensuring data governance and compliance through features such as row-level security and masking.

A key component of SAP Datasphere is SAP HANA Cloud. Access to the database is available via an Open SQL schema in SAP Datasphere that is automatically created when you establish a database user for your SAP Datasphere space. This database user is a technical user that allows external SQL clients (such as third-party reporting tools, ETL tools, or custom applications) to connect directly to the underlying SAP HANA Cloud database of your SAP Datasphere instance using Java Database Connectivity (JDBC) and Open Database Connectivity (ODBC). Figure 7.11 shows an Open SQL schema created in the **Finance** space.

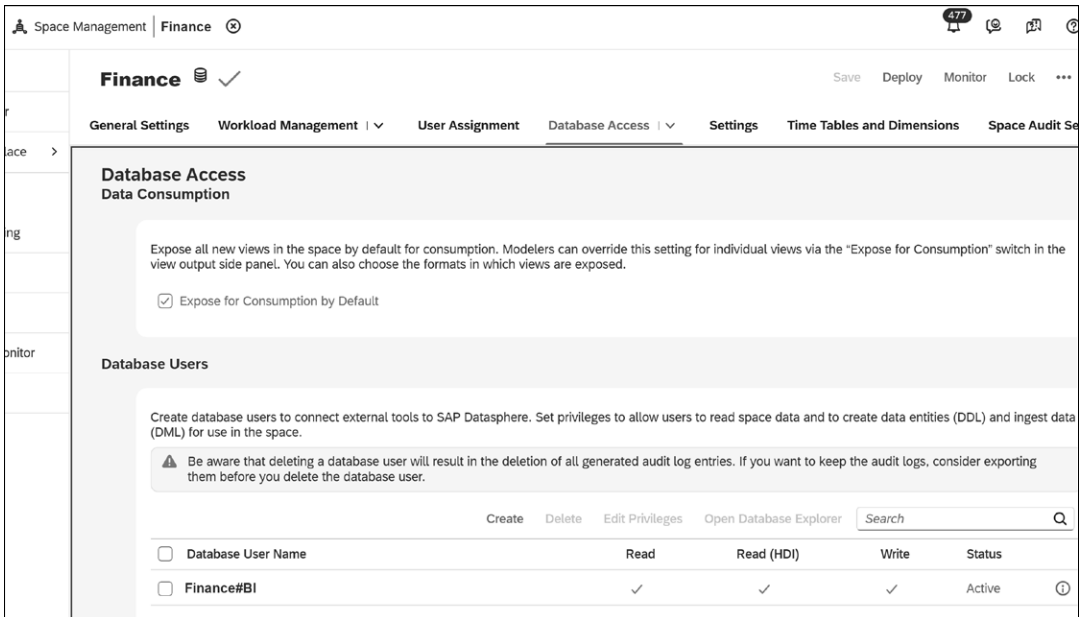


Figure 7.11: Open SQL Schema

The Open SQL schema acts as a dedicated area where you can expose data for external consumption or ingest data from external sources using standard SQL. This is crucial for scenarios where you need to interact with your SAP Datasphere data outside of the SAP Datasphere UI.

By default, the database user associated with the Open SQL schema is granted read privileges to data in the SAP Datasphere space schema (where your deployed SAP Datasphere models reside), provided those models are explicitly exposed for consumption. Crucially, the database user also has write access (SQL, data definition language [DDL], data manipulation language [DML]) to its own Open SQL schema. This means external tools can create tables and views, as well as insert, update, and delete data directly within this schema.

Data can be loaded into Open SQL schema tables by external tools. These tables can then be used as sources within SAP Datasphere's data flows or views. This provides a flexible way to bring data into SAP Datasphere, especially for scenarios with specific enterprise security protocols or when direct SAP Datasphere connectors aren't suitable.

Data processed within SAP Datasphere (e.g., from data flows) can be written to tables in the Open SQL schema, making that data directly accessible to external consuming applications.

While it provides direct SQL access, the Open SQL schema is controlled through the database user's privileges. Space administrators can manage these privileges, enabling or disabling read/write access to the space schema and the Open SQL schema. Unlike SAP Datasphere local tables, which are managed entirely within the SAP Datasphere environment

and not directly accessible via external SQL clients, Open SQL schema tables offer this direct database access, serving as a bridge for specific integration patterns.

7.3.4 Modeling: Business Builder

SAP Datasphere's Business Builder is a key component designed to empower business users to create semantic data models in business terms, without help from IT. It sits on top of the data layer. This is similar to the universe concept in SAP BusinessObjects BI, which obfuscated the technical details and presented a business view to the users for creating reports. Figure 7.12 shows the Business Builder capabilities.

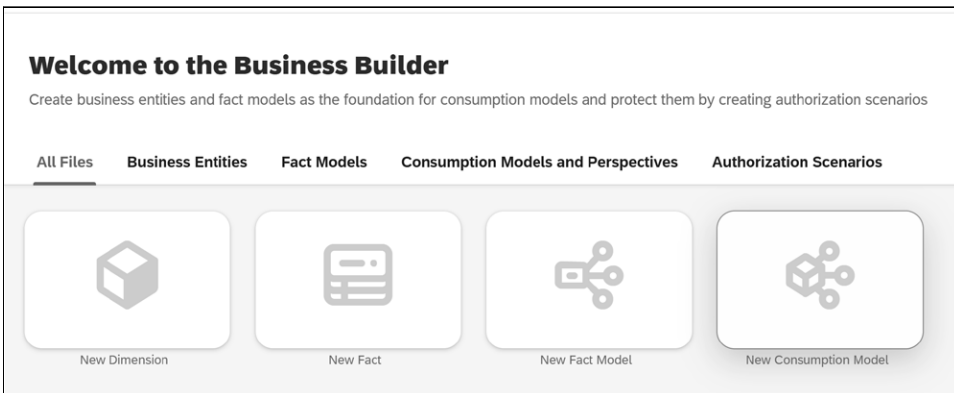


Figure 7.12: SAP Datasphere Business Builder

Let's look at each of these capabilities in more detail:

■ Business entities

These are the foundational building blocks in the Business Builder. A business entity represents a real-world business concept, such as customer, product, or sales order item. They are built on views or tables from the Data Builder (the technical layer) and are categorized as either dimensions or facts:

- Dimensions: Typically contain master data (e.g., customer master, product hierarchy) and usually have a key defined.
- Facts: Typically contain transactional data (e.g., sales transactions, production orders) and must have at least one measurable value (a measure) defined. Business entities define attributes (descriptive characteristics) and measures (quantifiable values) and establish associations between each other, much like an entity-relationship model, but presented in business language.

■ Fact models

These are optional, reusable intermediate models that help streamline the creation of consumption models. A fact model allows you to predefine a collection of related measures, attributes, and filters that are frequently used together in a specific business context. For instance, if you have several consumption models that all require the same set

of calculated sales metrics, you can define these once in a fact model and then reuse it across those consumption models, preventing redundancy and ensuring consistency.

■ Consumption models and perspectives

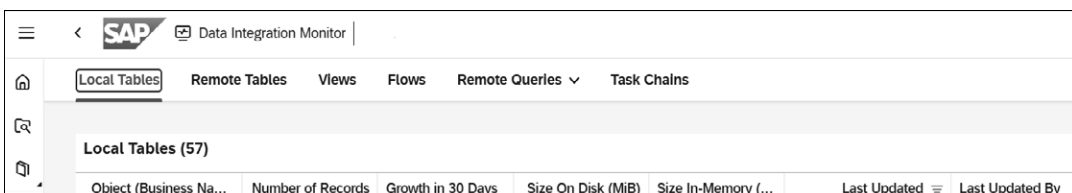
- Consumption models: These are the final, comprehensive data models designed for analytical consumption. They are built on top of business entities or fact models. Here, business users can combine multiple facts and dimensions to create a complete star schema–like structure, add further calculations, use filters, and define hierarchies, tailoring the data to specific reporting and analytical requirements.
- Perspectives: A perspective is a subset or tailored view of a consumption model. You can create multiple perspectives from a single consumption model, each focusing on a specific business question or audience. For example, a sales performance consumption model might have different perspectives for a regional sales manager (showing only their region’s data) versus a global sales executive. Perspectives are the actual objects that are exposed and consumed by frontend analytical tools such as SAP Analytics Cloud.

■ Authorization scenarios

These enable fine-grained data access control for the business models created in the Business Builder. Authorization scenarios use data access controls defined in the Data Builder (the technical layer). They allow you to define rules that filter the data visible to specific users or user groups based on their roles or characteristics (e.g., a manager only sees data for their team). By assigning an authorization scenario to a business entity or consumption model, you ensure that business users only access the data they are permitted to see, maintaining data security and compliance within the business context.

7.3.5 Data Integration Monitor

SAP Datasphere’s Data Integration Monitor (see Figure 7.13) is a core operational tool within SAP Datasphere designed to give administrators and integrators a comprehensive, real-time overview of all data integration and persistence activities occurring within their assigned spaces. Acting as a centralized dashboard, it ensures that the movement, transformation, and replication of data across SAP and non-SAP systems are transparent, traceable, and manageable.



Object (Business Na...	Number of Records	Growth in 30 Days	Size On Disk (MiB)	Size In-Memory (...)	Last Updated	Last Updated By
Local Tables (57)						

Figure 7.13: SAP Datasphere Data Integration Monitor

The components of the Data Integration Monitor in SAP Datasphere include the following:

- **Remote table monitor**
 - Tracks replication status and data transfer history for remote tables.
 - Allows scheduling of periodic replication or manual snapshot loads.
 - Supports copying remote tables into local persistent storage.
- **View persistency monitor**
 - Monitors views that have been set up for *persisted storage*.
 - Displays update frequency, last load date, and data volume.
 - Helps manage performance trade-offs between virtual access and materialization.
- **Flows monitor (data flows and replication flows)**
 - Tracks the execution history and current state of data flows (used for transformation) and replication flows (used for movement of data).
 - Shows execution details such as start/end time, row count, and error messages.
 - Enables reruns, status checks, and dependency reviews.
- **Remote query monitor**
 - Offers diagnostic insights into how SAP Datasphere queries external data sources in real time via SAP HANA smart data access or SAP HANA smart data integration.
 - Useful for performance tuning and identifying slow or failing queries.
- **Task chain monitor**
 - Visualizes the status and sequence of *task chains*—grouped data tasks executed in a logical or time-based order.
 - Helps in managing dependencies and tracking the progress of complex workflows.

7.3.6 Catalog and Marketplace

The SAP Datasphere catalog (aka data catalog) is an integrated metadata management and data discovery tool. It allows users—especially data analysts, data stewards, and business users—to easily search, explore, and understand the available data assets across their SAP Datasphere spaces.

Using the data catalog, SAP Datasphere users can do the following:

- Create and standardize business terms across the organization.
- Link definitions to specific data assets, promoting semantic consistency.
- Create KPIs along with information on how to calculate them and link them to assets and documentation. This goes a long way in removing ambiguity on how a certain metric is to be viewed across the organization. Clarity with consistent jargon is a major boost in productivity for any organization.
- Enrich metadata by tagging, annotating, and classifying datasets with custom labels.

- Ensure integrated security and governance so that users only see the metadata they are permitted to view.
- Search for assets, terms, or KPIs from a single interface using the search feature. Users can search for datasets, views, tables, and models using business or technical terms. Figure 7.14 shows the **Search** feature and a listing of the data providers.
- Easily view metadata such as data types, descriptions, lineage, and creation/modification history.

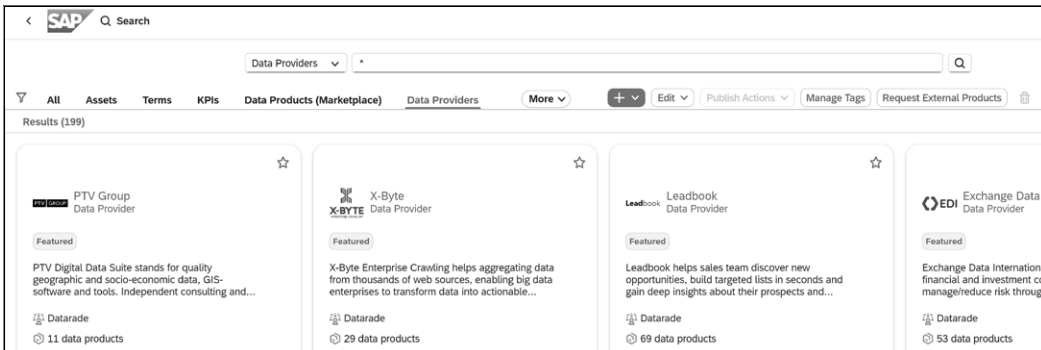


Figure 7.14: Catalog and Marketplace Search

The SAP Datasphere data marketplace is a secure platform for exchanging data assets—both internally (within an enterprise) and externally (with partners or third parties). It allows data providers to publish and monetize data and data consumers to subscribe to and integrate trusted datasets into their own data landscapes. Using the data marketplace enables the users to do the following:

- Share curated datasets securely across teams and business units or with external parties.
- Use contract-based governance for external sharing scenarios.
- Subscribe to datasets and immediately use them in modeling and analytics. These datasets appear as local assets available for further transformation or direct use.
- Find listed data providers and lists of the data products each offers. Data providers are able to define dataset metadata and descriptions, as well as update SLA information and schedules.
- Search for data products of interest and subscribe to them.

Marketplace data is integrated directly into spaces, allowing for seamless data modeling and analytics within the subscriber’s environment.

Together, the catalog and marketplace in SAP Datasphere provide a powerful ecosystem for discovering, governing, and sharing data. The catalog enhances internal visibility, understanding, and trust in data. The marketplace facilitates external and internal data collaboration, supporting use cases such as supply chain collaboration, industry benchmarking, or data monetization.

These features are essential for enabling data democratization while maintaining strong governance and security controls—key pillars of the modern business data fabric approach that SAP Datasphere embodies.

7.4 SAP Business Warehouse and SAP BW/4HANA

SAP Business Data Cloud provides a lifeline to existing organizations on SAP BW. With SAP Business Data Cloud, organizations have the option of moving their on-premise SAP BW installation to a private cloud edition with extended support. This safeguards their existing investments and operations and gives them an extended time period for moving to the cloud with a transition to SAP Business Data Cloud. The minimum version required for moving SAP BW to SAP Business Data Cloud is SAP BW 7.5 powered by SAP HANA.

Here are the key strategies for moving SAP BW to SAP Business Data Cloud:

- **Lift to private cloud and shift with the SAP BW bridge**

This strategy involves lifting the existing SAP BW 7.5 (SAP NetWeaver or SAP BW/4HANA) system to a private cloud edition without immediate architectural changes. This allows organizations to retain their current setups while extending mainstream maintenance (e.g., SAP BW 7.5 to 2030, SAP BW/4HANA to 2040). The SAP BW bridge within SAP Datasphere is used to transfer metadata and data (via remote or shell conversion). It's a low-risk starting point but may delay full modernization benefits.

- **Phased migration with Data Product Generator and SAP BW bridge**

The Data Product Generator utility service helps convert existing SAP BW InfoProviders to data products that can be consumed within SAP Business Data Cloud. After lifting to the cloud, this approach shifts metadata into SAP Datasphere (a core SAP Business Data Cloud component). It enables zero-copy data management via Delta Sharing, preserving semantics and allowing gradual integration with tools such as SAP Databricks for AI/ML.

The SAP BW bridge in SAP Datasphere is a SAP BW-compatible staging and modeling environment inside SAP Datasphere that allows existing SAP BW and SAP BW/4HANA data flows, ABAP routines, and ODP extractors to be moved to and run in SAP Datasphere with minimal rework. This phased strategy supports incremental innovation but requires careful planning to avoid data silos.

- **Full transformation to SAP Datasphere**

For organizations seeking a complete overhaul, this strategy reengineers SAP BW data models and processes into SAP Datasphere's cloud-native architecture. It uses advanced analytics, real-time capabilities, and prebuilt intelligent applications, but involves reworking customizations and transformations, posing higher up-front effort and potential disruption.

- **Selective migration**

Focusing on critical data and objects (e.g., key reports, KPIs), this method migrates only essential elements to SAP Business Data Cloud, optimizing costs and time. It's suitable

for businesses with dormant data or complex landscapes, though it risks overlooking long-term scalability.

■ **Hybrid approach**

Combining on-premise SAP BW with SAP Business Data Cloud services, this strategy uses the SAP BW bridge for legacy retention while adopting SAP Datasphere for new analytics use cases. It offers flexibility but may complicate governance and integration over time.

Tip: Get a Free Modernization Assessment

A free modernization assessment is a good starting point for organizations that already use SAP. Get the assessment at www.sap.com/products/data-cloud/sap-migration-assessment.html.

SAP Business Data Cloud is the envisioned end state for existing SAP BW and SAP BW/4HANA organizations with innovative enhancements beyond what SAP BW can offer. After migrating to SAP Business Data Cloud, organizations can do the following:

- Use SAP Databricks for AI/ML models, and use SAP BW or SAP BW/4HANA data for predictive analytics, time-series forecasting, and generative AI applications.
- Deploy intelligent applications for prebuilt, cross-functional analytics (e.g., profitability, sustainability) tailored to specific LoBs.
- Use the object store for cost-effective storage of large datasets and Delta Share for zero-copy data sharing with lakehouse architectures.
- Gradually decommission legacy SAP BW systems as new use cases are fully transitioned to SAP Business Data Cloud.

To successfully migrate SAP BW to SAP Business Data Cloud, organizations should follow these practical steps:

1. **Conduct a readiness assessment.**

Use SAP's free SAP BW Cloud Modernization Assessment or partner tools to evaluate the current SAP BW environment and prioritize migration scope.

2. **Engage SAP Partners.**

Collaborate with certified partners for customized road maps, license consulting, and implementation support.

3. **Plan a phased migration.**

Adopt the lift, shift, and innovate approach—lift to SAP BW 7.5 Private Cloud Edition, shift the metadata to SAP Business Data Cloud, and innovate by using data products—to minimize disruption and align with business priorities.

4. **Use the SAP BW bridge.**

Use this tool to transfer SAP BW data and logic to the cloud, ensuring compatibility with SAP Datasphere and SAP Analytics Cloud.

5. **Validate data products.**

Use the Data Product Generator to create data products and consume them in SAP Datasphere, ensuring semantic integrity.

6. **Train teams.**

Provide training on SAP Business Data Cloud components (SAP Datasphere, SAP Analytics Cloud, SAP Databricks) to empower IT and business users.

7. **Monitor and optimize.**

Use the data catalog and knowledge graph to track data usage and governance, refining processes as needed.

7.5 SAP Databricks

SAP Databricks is an SAP-managed, OEM version of the Databricks Data Intelligence Platform, seamlessly embedded within SAP Business Data Cloud. It combines advanced AI/ML, data science, and data engineering capabilities with semantically rich, business-ready data from SAP applications. Organizations can access SAP data products via zero-copy Delta Sharing and integrate semi-structured and unstructured data from any source into a unified, harmonized data model. These curated data products simplify data access, eliminating complex data movement while ensuring high-quality, reliable data. The SAP Databricks and SAP Business Data Cloud pair deliver a trusted, governed, and harmonized data foundation, streamlining operations, lowering total cost of ownership, and empowering intelligent applications for unparalleled business insights and AI innovation.

SAP Databricks uses SAP-managed SAP HANA Cloud, data lake files as cloud storage. The key capabilities of SAP Databricks are as follows:

■ **Zero-copy data harmonization**

Blend data without complex ETL processes or data duplication. Connect to external object stores containing non-SAP data and integrate semi-structured and unstructured data from any source with SAP application data products into a single, harmonized data model.

■ **Trusted data products**

Use SAP-managed, curated data products to accelerate AI/ML and data engineering. Create custom derived data products from enriched SAP and non-SAP data.

■ **Pro-code development**

Use pro-code tools and notebooks to build custom AI/ML models, perform data engineering tasks, or develop Spark pipelines to integrate SAP and third-party data.

■ **Scalable, on-demand analysis**

Use SAP Databricks SQL for large-scale, on-demand data analysis with SQL, enabling faster, data-driven decisions.

- **Secure and governed AI/ML**

Develop secure, governed custom AI/ML solutions using the SAP Databricks data science workbench and Mosaic AI.

SAP Databricks includes almost all the features and capabilities of the full Databricks Data Intelligence Platform: Delta Share, unity catalog, data science workbench, Mosaic AI, notebook, and Databricks SQL.

SAP Business Data Cloud natively implements the medallion architecture to organize, refine, and deliver trusted, high-quality data across the enterprise. This multilayered architecture—popularized in data lakehouse paradigms—forms the backbone for scalable data pipelines, enabling AI, analytics, planning, and data productization. This architecture typically consists of three main layers, as follows:

- **Bronze layer**

This is the initial landing zone for all raw data, regardless of its source (SAP S/4HANA, SAP SuccessFactors, non-SAP systems, etc.). The data is ingested as is, meaning its structure largely mirrors that of the source system. The bronze layer serves as a historical record, providing data lineage and auditability. It allows for reprocessing data if needed without having to reread from the source. Data in this layer is stored in cost-effective, scalable data lake files within the SAP HANA Cloud, data lake files system.

- **Silver layer**

Data from the bronze layer undergoes initial processing in the silver layer. This involves matching, merging, conforming, and cleansing the data to remove inaccuracies, inconsistencies, and duplicates. The silver layer aims to provide an enterprise view of key business entities, concepts, and transactions (e.g., master customers, nonduplicated transactions). This brings data from disparate sources into a more unified and structured format.

The philosophy here is to apply just-enough transformations and cleansing rules to make the data ready for general exploration and self-service analytics. More complex business rules and project-specific transformations are typically reserved for the gold layer.

The foundation services discussed in Section 7.1.2 play a key role in this layer, replicating bundled data, harmonizing it with other sources, and transforming it into data products. This integration, transformation, and cleansing is managed by SAP.

- **Gold layer**

Optional for SAP Business Data Cloud, the gold layer contains highly curated, aggregated, and modeled data that's optimized for specific business use cases, such as analytics, reporting, and AI/ML applications. Data in this layer is optimized for intensive reads and consumption in a read-only fashion, ensuring efficiency and reliability for dashboards, analytical models, and AI applications. It can be enhanced with business semantics and provide a single unified layer for analytics.

SAP Datasphere (discussed in Section 7.3) plays a key role in the medallion architecture. It defines data models on top of these data products, supporting prebuilt intelligent applications and custom models. It also manages analytical roles and data access control. SAP Analytics Cloud (discussed in Chapter 8) provides the visualization and planning tools, enabling the creation of dashboards and reports on top of the gold layer data, including out-of-the-box intelligent applications.

SAP Databricks doesn't support advanced pipeline orchestration for complex workflows, but organizations can integrate external data using the following approaches:

- **Scheduled notebook jobs**

Organizations can develop Spark pipelines within SAP Databricks notebooks and schedule them using calendar-based triggers for automated execution. Notebooks also support the creation of bronze, silver, and gold jobs.

- **Data replication via SAP Datasphere**

Organizations can replicate data into SAP Datasphere using replication flows, convert it into data products, and share these with SAP Databricks. This capability is now generally available and uses Delta Sharing for zero-copy, governed data delivery.

- **External object store integration**

Organizations can mount external object stores directly into SAP Databricks for seamless access to large-scale data without manual data movement or duplication. However, frequent read/write operations may increase cloud storage and egress costs depending on the storage provider and access patterns.

- **Delta Sharing**

If supported by the source system, organizations can share external data with SAP Databricks via Delta Sharing. This data isn't persisted in SAP Databricks, as it's shared without physical copying.

SAP addresses the following situations for new organizations with SAP Databricks. These two situations are classified as *greenfield*:

- Totally new to SAP Databricks: Go to SAP Business Data Cloud
- Have existing Databricks, but not processing any SAP data: Go to SAP Business Data Cloud

However, if an existing Databricks organization is already using the product to process SAP data, they are classified as a *brownfield* organization. Here it's important to assess the extent of this processing. For limited processing, the recommended path is still to go to SAP Business Data Cloud. For extensive integrations with Databricks, a connector is planned that will enable zero-copy delta sharing with SAP Business Data Cloud from the existing Databricks.

The SAP Databricks services, including SAP-managed SAP HANA Cloud, data lake files, are planned to be made available as capacity services of SAP Business Data Cloud. This means

that subscribed SAP Business Data Cloud core capacity units (CUs) will be available to allocate to and be consumed by the SAP Databricks services.

Figure 7.15 shows the steps involved in integrating data from SAP Business Data Cloud with SAP Databricks. This data can be combined with external data, and ML algorithms can be run on this data to generate insights. These insights can be shared back with SAP Business Data Cloud using the SAP Datasphere component and can be used in visualizations served by SAP Analytics Cloud.

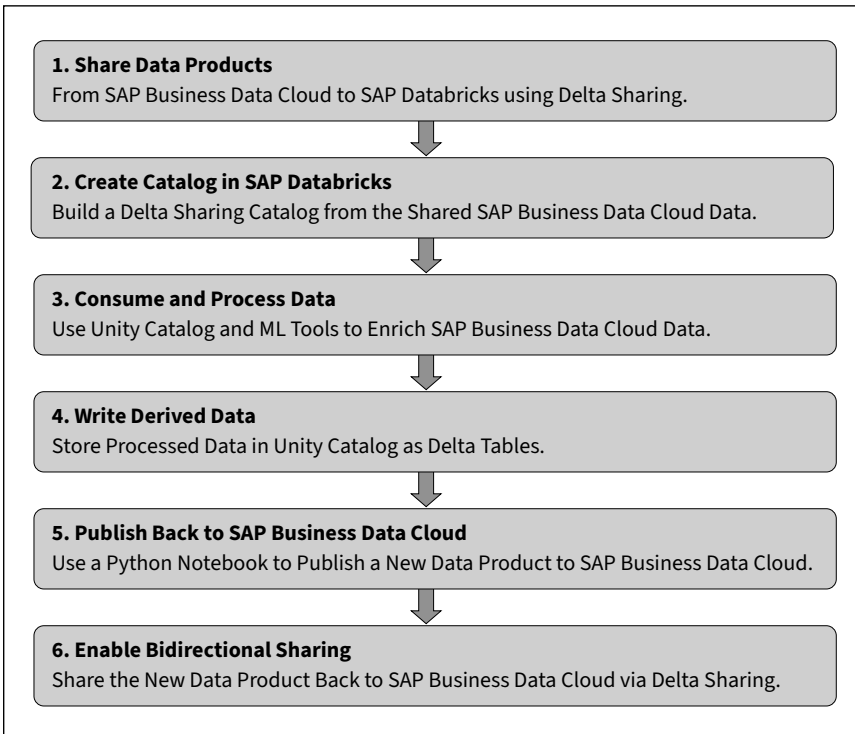


Figure 7.15: SAP Databricks and SAP Business Data Cloud Data Integration Flow

7.6 SAP Analytics Cloud

With the tight integration in SAP Business Data Cloud, SAP's flagship visualization product—SAP Analytics Cloud—is at the forefront. SAP Analytics Cloud serves as the primary consumption layer for analytics, using data managed within SAP Business Data Cloud's ecosystem, particularly through SAP Datasphere and data products.

SAP Analytics Cloud plays a critical role in SAP Business Data Cloud as the frontend analytics and visualization tool, enabling business users to consume and analyze data prepared by SAP Business Data Cloud's components, such as SAP Datasphere and data products. Its integration within SAP Business Data Cloud makes it easier to deliver actionable insights,

real-time reporting, and AI-driven analytics. Key aspects of SAP Analytics Cloud's role in SAP Business Data Cloud include the following:

■ **Consumption layer for data products and intelligent applications**

- Data products: When SAP Business Data Cloud data products are installed, semantically rich datasets sourced from SAP (e.g., SAP S/4HANA) and non-SAP systems can be ingested into SAP Datasphere spaces. SAP Analytics Cloud consumes these data products directly from the SAP Datasphere consumption spaces, enabling users to create dashboards, reports, and visualizations without needing to handle raw data or complex ETL processes.
- Intelligent applications: These high-level data products bundle data products, data models, analytical models, and SAP Analytics Cloud visualizations into prebuilt, AI-driven applications. When an intelligent application (e.g., Working Capital Insight) is installed via the SAP Business Data Cloud cockpit, SAP Analytics Cloud automatically deploys the associated dashboards, providing immediate access to predefined analytics, such as forecasting or KPI tracking.

■ **Real-time analytics and reporting**

- SAP Analytics Cloud uses SAP Datasphere's real-time data access capabilities (e.g., via remote tables or replication flows) to provide up-to-date insights. For example, SAP Analytics Cloud can query live data from SAP S/4HANA or Google BigQuery through SAP Datasphere's federation capabilities, enabling real-time reporting without data replication. A live connection is maintained between SAP Analytics Cloud and SAP Datasphere, which means that SAP Business Data Cloud data isn't stored in SAP Analytics Cloud.
- Features such as Smart Insights and Smart Discovery in SAP Analytics Cloud use AI to analyze data products, identifying trends, correlations, and anomalies that are critical for SAP Business Data Cloud's goal of delivering actionable insights.

■ **Planning and predictive analytics**

- SAP Analytics Cloud's planning capabilities allow users to create financial and operational plans using data from data products. For instance, a data product such as a financial planning entry item can feed into SAP Analytics Cloud's planning tools for budgeting or forecasting.
- Predictive analytics features, powered by built-in ML models, enable scenario analysis and what-if simulations, using SAP Business Data Cloud's integrated data to predict outcomes such as sales trends or supply chain risks.

■ **Unified user experience**

- SAP Analytics Cloud provides a user-friendly, browser-based interface accessible via the SAP Business Data Cloud cockpit, ensuring business users can interact with data without technical expertise. Its integration with SAP Datasphere's Business Builder allows users to consume semantically rich models (e.g., KPIs, business entities) created by domain experts.

- SAP Analytics Cloud’s augmented analytics features, such as natural language processing (NLP) via Joule integration, enable users to ask questions in plain language (e.g., “Show me sales by region”) and receive visualized responses, enhancing accessibility within SAP Business Data Cloud.

■ Integration with SAP Databricks and third-party tools

- Within SAP Business Data Cloud, SAP Analytics Cloud complements SAP Databricks, which handles advanced AI/ML workloads. For example, predictive models developed in SAP Databricks (e.g., demand forecasting) can be shared back with SAP Datasphere and visualized in SAP Analytics Cloud dashboards, creating a seamless analytics pipeline.
- SAP Analytics Cloud supports integration with third-party tools (e.g., Tableau, Power BI) via OData APIs, allowing SAP Business Data Cloud’s data to be consumed in external visualization platforms, though SAP Analytics Cloud remains the primary tool for SAP-centric analytics.

■ AI integration

- The Just Ask feature within SAP Analytics Cloud allows the user to use the power of generative AI to create stories quickly by prompting with very specific analytic questions based on an analytical model.

Prebuilt content analytics is delivered in the form of SAP Analytics Cloud dashboards. Figure 7.16 shows the visualizations of the People Intelligence in SAP Business Data Cloud intelligent application.

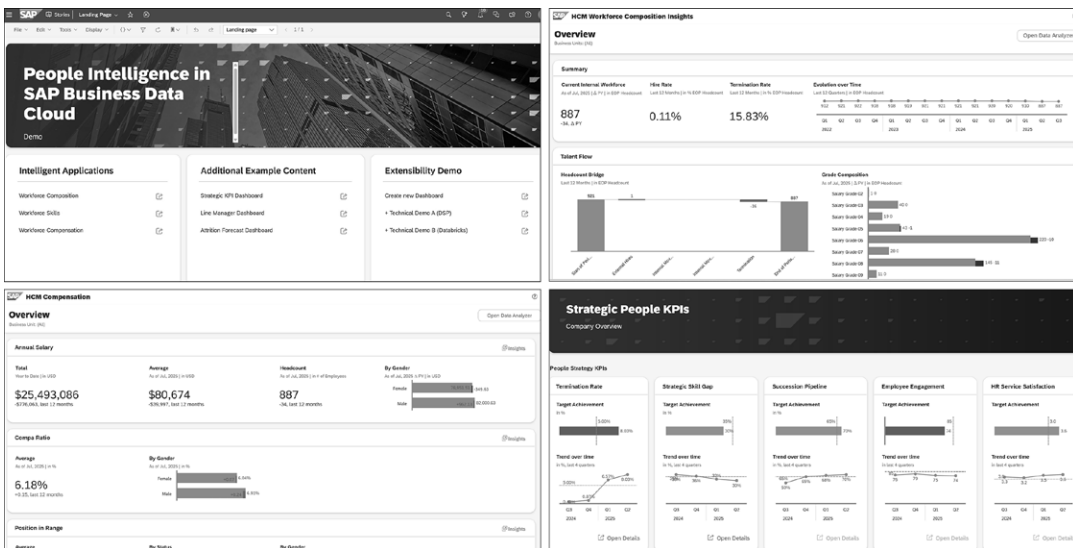


Figure 7.16: People Intelligence in SAP Business Data Cloud Application

7.7 Business Use Cases

Consider the following real-life business situation showing how SAP Business Data Cloud can be used. A multinational retailer with a large SAP S/4HANA implementation seeks to enhance its analytics capabilities to drive data-driven decision-making. The chief technology officer (CTO) empowered by the Board requirements, aims to create agile analytics by integrating financial and sales data from SAP S/4HANA with employee data from SAP SuccessFactors. Certain subsidiaries have created a data lake in Snowflake. Data from Snowflake will have to be brought in for a complete picture. Other third-party data (competitor metrics, campaign spends, and social sentiment) may also be required. The objectives are to do the following (the C-level executives in focus are shown in parentheses):

1. Calculate KPIs such as revenue per employee (chief executive officer [CEO]- or chief financial officer [CFO]) and profit per employee (CEO or CFO), segmented by profit center, cost center, and company code.
2. Develop two C-level dashboards:
 - An internal company dashboard focusing on internal financial metrics with planning and ML-driven forecasts. Predictive analytics catering to seasonal scenarios may be incorporated.
 - A competitive dashboard comparing internal profitability metrics with third-party data.
3. Apply ML to do the following:
 - Analyze spending patterns and campaign spends to predict campaign performance (CFO and chief marketing officer [CMO]).
 - Provide forward-looking financial guidance to capital markets (CFO).
4. Use five years of historical data trends for ML model development.

For this use case, Section 7.7.1 provides an overall view of SAP Business Data Cloud as a solution choice. The role of each component is also summarized here. Integrating data from the various sources and modeling is covered in Section 7.2 using SAP Business Data Cloud components and capabilities.

7.7.1 Solution Overview

SAP Business Data Cloud addresses the retailer's business case by providing a unified platform for data integration, analytics, and AI-driven insights, enabling agile analytics, actionable dashboards, and predictive capabilities. The solution uses the following products:

- **SAP Datasphere**

For data integration, modeling, and governance of SAP S/4HANA, SAP SuccessFactors, and third-party data.

- **SAP Analytics Cloud**
For creating C-level dashboards with real-time visualizations, planning, and predictive analytics.
- **SAP Databricks**
For advanced ML model development using historical data.
- **SAP Business Data Cloud cockpit**
For streamlined deployment of data products and intelligent applications.
- **SAP Business Data Cloud Connect**
For connecting to Snowflake using the Delta Share technology.
- **SAP HANA Cloud, data lake files and Delta Share**
For cost-efficient storage and zero-copy data sharing.

It's necessary to understand all four objectives mentioned earlier (calculate KPIs, develop dashboards, apply ML, use historical data trends for model development) and come up with a comprehensive plan that addresses your organization's needs. For this use case, we assume that the system has been provisioned and a formation has been created with SAP S/4HANA, SAP Datasphere, and SAP Analytics Cloud.

7.7.2 Data Integration and Modeling

Our requirements are to integrate financial and sales data from SAP S/4HANA, employee data from SAP SuccessFactors, and third-party data (competitor metrics, campaign spends, social sentiment) to calculate KPIs (revenue per employee, profit per employee) segmented by profit center, cost center, and company code. Let's see how SAP Business Data Cloud helps us meet these requirements:

- **Data integration**
 - SAP S/4HANA and SAP SuccessFactors: Deploy prebuilt data products (e.g., journal entry header, cost center, company code for SAP S/4HANA; employee data for SAP SuccessFactors) via the SAP Business Data Cloud cockpit. These create ingestion spaces in SAP Datasphere, storing data in SAP HANA Cloud for high-performance queries or SAP HANA Cloud, data lake files) for cost-efficient storage of large datasets. This data will land in the bronze layer (SAP HANA Cloud, data lake files/object store) in its raw, immutable form, serving as a single, trusted source of truth with full data lineage. SAP Business Data Cloud's zero-copy data integration through Delta Lake and Delta Sharing will ensure efficient access without costly data replication.
 - Snowflake data: Use the solution extension of SAP Business Data Cloud Connect to share and obtain data from Snowflake.
 - Third-Party data: Use SAP Datasphere's native adapters (e.g., for Google BigQuery, AWS S3) or open connectors to ingest competitor metrics, campaign spends, and social sentiment data. Replication flows enable batch or delta loading for persistent

storage, while remote tables (via SAP HANA smart data access) provide real-time federation for live data access.

■ **Modeling and KPI calculation**

- Data harmonization: Use SAP Datasphere’s Data Builder to create data flows, merging SAP S/4HANA sales revenue, SAP SuccessFactors employee headcount, Snowflake, and third-party data into unified datasets. For example, combine revenue data with employee counts to calculate KPIs. The gold layer of the data from which dashboards will fetch the metrics is created in SAP Datasphere.
- Semantic modeling: In SAP Datasphere’s Business Builder, define business entities and KPIs (e.g., Revenue per employee = Total sales revenue ÷ Employee count) with segmentation by profit center, cost center, and company code. These models preserve business context (e.g., company code hierarchies).

7.7.3 C-Level Dashboards

The requirement is to develop two SAP Analytics Cloud-based dashboards, as follows:

■ **Internal dashboard**

Focuses on internal financial metrics (revenue per employee, profit per employee) with planning and ML-driven forecasts, including seasonal scenarios.

■ **Competitive dashboard**

Compares internal profitability metrics with third-party data (competitor metrics, campaign spends, social sentiment).

Let’s see how SAP Business Data Cloud helps us meet this requirement:

■ **Internal company dashboard**

- Content: Visualizes KPIs (revenue per employee, profit per employee) segmented by profit center, cost center, and company code using SAP Analytics Cloud’s interactive charts (e.g., bar charts, heatmaps). Incorporates planning scenarios via SAP Analytics Cloud’s planning modeler for budgeting and forecasting (e.g., seasonal sales projections).
- ML-driven forecasts: Integrates predictive outputs from SAP Databricks (e.g., revenue forecasts based on spending patterns) into SAP Analytics Cloud dashboards, displayed as trend lines or predictive KPIs. SAP Analytics Cloud’s Smart Predict supports time-series forecasting for seasonal scenarios (e.g., holiday sales peaks).
- Features: Uses SAP Analytics Cloud’s Smart Insights to highlight trends (e.g., profit declines in specific cost centers) and NLP via Joule for ad hoc queries (e.g., “Show profit per employee by company code for Q1”).
- Deployment: Deploys via an intelligent application (e.g., Financial Analytics) in the SAP Business Data Cloud cockpit, which auto-configures SAP Analytics Cloud dashboards, data models, and underlying data products for rapid setup.

■ Competitive dashboard

- Content: Compares internal KPIs (e.g., profit per employee) with third-party data (e.g., competitor revenue per employee, campaign return on investment [ROI], social sentiment scores) using side-by-side visualizations (e.g., comparative bar charts, sentiment trend graphs).
- Data sources: Pulls internal KPIs from SAP Datasphere’s consumption spaces and third-party data from ingestion spaces, accessed via replication flows or real-time federation (e.g., live competitor data from Google BigQuery).
- Features: Uses SAP Analytics Cloud’s Smart Discovery to identify correlations (e.g., campaign spends vs. social sentiment) and NLP for dynamic queries (e.g., “Compare our profit per employee with competitors in Europe”).
- Deployment: Deploys via an intelligent application (e.g., Competitive Insights) in the SAP Business Data Cloud cockpit, ensuring seamless integration of internal and external data.

■ Technical integration

- SAP Analytics Cloud consumes analytical models and views from SAP Datasphere’s consumption spaces, built in Data Builder or Business Builder. For example, a view combining SAP S/4HANA profit data with competitor metrics feeds the competitive dashboard.
- Real-time access: SAP Analytics Cloud uses SAP Datasphere’s remote tables (via SAP HANA smart data access) for live third-party data or replicated data in SAP HANA Cloud/SAP HANA Cloud, data lake files for performance.
- Mobile access: SAP Analytics Cloud’s iOS and Android apps, with voice mode support, allow C-level executives to access dashboards on the go.

7.7.4 Machine Learning Applications

First, let’s look at our requirements. We’re looking to apply ML to do the following:

- Analyze spending patterns and campaign spends to predict campaign performance (CFO and CMO).
- Provide forward-looking financial guidance for capital markets (CFO) using five years of historical data.

Next, let’s see how SAP Business Data Cloud helps us meet these requirements:

■ Data preparation

- Store five years of historical data (SAP S/4HANA sales, SAP SuccessFactors employee metrics, campaign spends, social sentiment) in SAP HANA Cloud, data lake files, SAP Business Data Cloud’s cost-efficient object store.
- Share data with SAP Databricks via Delta Share, enabling zero-copy access to large datasets (terabytes) without replication, reducing storage costs.

■ Campaign performance prediction (CFO/CMO)

- Model development: Use SAP Databricks’ Spark-based processing to build ML models (e.g., regression, decision trees) analyzing spending patterns and campaign spends. For example, correlate campaign budgets with sales uplift and social sentiment to predict ROI.
- Output: Feed predictions (e.g., Predicted Campaign ROI by Region) to SAP Analytics Cloud’s internal dashboard, visualized as interactive charts or KPIs.
- Example: A model predicts that a \$1 M campaign spend in Q4 will yield a 15% sales increase based on historical trends and sentiment data.

■ Financial guidance for capital markets (CFO)

- Model development: Build time-series forecasting models (e.g., AutoRegressive Integrated Moving Average [ARIMA], neural networks) in SAP Databricks to predict revenue and profit trends over the next 12–24 months, incorporating seasonal patterns (e.g., holiday sales).
- Output: Integrate forecasts into SAP Analytics Cloud’s planning modeler for scenario analysis (e.g., “Impact of 5% cost reduction on profit per employee”). Visualizations in the internal dashboard provide CFOs with forward-looking guidance for investor communications.
- Example: A forecast predicts a 10% revenue growth in 2026, adjusted for seasonal fluctuations, supporting confident capital market guidance.

7.7.5 Governance and Scalability

First, let’s look at our requirement: Ensure secure, scalable, and compliant analytics for C-level stakeholders. Next, let’s see how SAP Business Data Cloud helps us meet this requirement:

■ Governance

- Use SAP Datasphere’s data catalog to manage metadata, ensuring data lineage and quality for KPIs and third-party data.
- Implement spaces in SAP Datasphere for role-based access control, restricting dashboard access to C-level executives (e.g., CEO, CFO, CMO) and relevant teams (e.g., finance, marketing).
- Apply encryption, row-level security, and data masking to comply with regulations such as General Data Protection Regulation (GDPR), especially for sensitive employee and financial data.

■ Scalability

- SAP HANA Cloud supports high-performance queries for real-time analytics, while SAP HANA Cloud, data lake files scales to handle terabytes of historical data cost-effectively.

- Store curated KPI datasets in SAP HANA Cloud for fast analytics or SAP HANA Cloud, data lake files for cost-effective storage of five years of historical data, accessible via Delta Share for zero-copy integration with SAP Analytics Cloud and SAP Databricks.
- Distributed computing in SAP Databricks handles large-scale data processing, ensuring efficient model training and deployment. The Delta Lake format supports advanced features such as time travel queries for analyzing historical trends.
- Delta Share enables zero-copy data sharing between SAP Datasphere, SAP Analytics Cloud, and SAP Databricks, reducing storage and processing overhead.
- SAP Business Data Cloud’s cloud-native architecture on SAP BTP ensures scalability for growing data volumes and user bases.

■ **Monitoring**

- Use SAP Datasphere’s Data Integration Monitor to track replication flows, data transfers, and query performance, ensuring reliable data pipelines for dashboards and ML models.

7.7.6 Conclusion

Overall, SAP Business Data Cloud provides a comprehensive solution for the multinational retailer’s business case by using its business data fabric to integrate SAP S/4HANA, SAP SuccessFactors, and third-party data. SAP Datasphere enables seamless data ingestion, modeling, and governance, while SAP Analytics Cloud delivers C-level dashboards with real-time KPIs, planning, and predictive analytics. SAP Databricks powers ML models for campaign performance and financial guidance, using five years of historical data. The SAP Business Data Cloud cockpit streamlines deployment, and SAP HANA Cloud, data lake files and Delta Share ensure cost efficiency and scalability. By addressing data integration, analytics, and AI requirements, SAP Business Data Cloud empowers the CTO to deliver agile, data-driven insights for the CEO, CFO, and CMO, aligning with the Board’s vision. The specific benefits realized are as follows:

■ **Agile analytics**

SAP Analytics Cloud’s real-time dashboards and NLP capabilities enable C-level executives to access insights quickly, supporting agile decision-making. The Just Ask feature allows executives to use generative AI prompting to get answers to analytical queries.

■ **Unified data fabric**

SAP Business Data Cloud integrates SAP S/4HANA, SAP SuccessFactors, and third-party data, eliminating silos and ensuring consistent KPIs.

■ **Predictive power**

ML in SAP Databricks models delivers accurate campaign performance predictions and financial forecasts, enhancing CMO and CFO decision-making.

■ **Cost efficiency**

SAP HANA Cloud, data lake files and Delta Share minimize storage and replication costs, optimizing large-scale data handling.

- **Scalability**

SAP Business Data Cloud's cloud-native architecture supports growing data volumes and user demands.

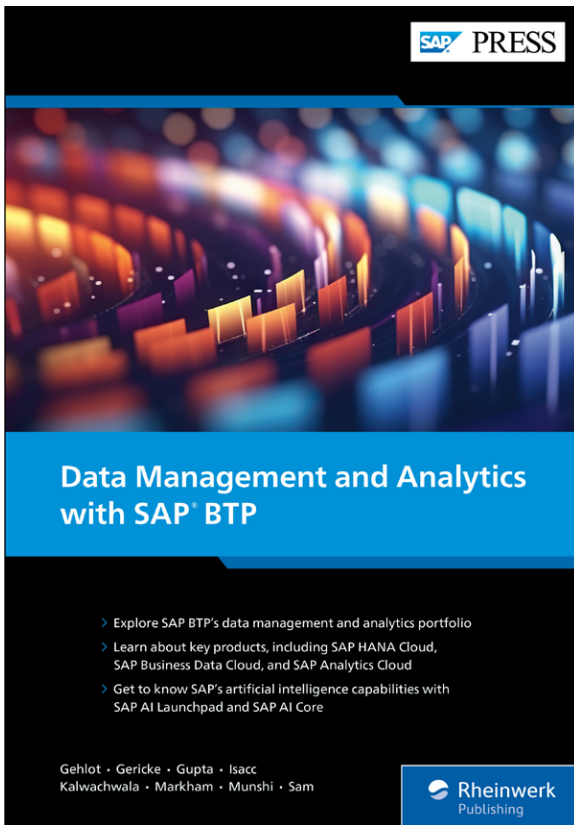
- **Governance and compliance**

The SAP Datasphere data catalog and security features ensure data integrity and regulatory compliance.

7.8 Summary

SAP Business Data Cloud has quickly evolved into SAP's flagship SaaS data fabric on SAP BTP by combining SAP Datasphere, SAP Analytics Cloud, SAP Databricks, and SAP BW (optional) for unified analytics and AI. The number of data products provide a quick path to value for organizations without having to worry about underlying structures or maintenance. By deploying intelligent applications, complex analytical needs can be satisfied in a very short time frame as SAP has taken away the complexity of figuring out underlying data structures and conducting complex ETL to extract data. The combination and close integration of SAP Datasphere and SAP Analytics Cloud creates enables you to combine SAP data with virtually any unstructured or structured source, harmonizing it and creating rich analytics for decision-making and planning. Efficient storage is achieved by the underlying foundation services resulting in data lake scales. The infusion of SAP Databricks provides the latest ML capabilities without redundant storage.

With SAP's commitment to the data product strategy, more and more LoB data will be available in a ready-to-consume manner. The entire SAP Business Data Cloud ecosystem is structured to address the challenges of complex data extraction, efficient storage, and rich analytics, as well as to handle AI/ML futuristic use cases.



Gehlot, Gericke, Gupta, Isacc,
Kalwachwala, Markham, Munshi, Sam

Data Management and Analytics with SAP BTP

- Explore SAP BTP's data management and analytics portfolio
- Learn about key products, including SAP HANA Cloud, SAP Business Data Cloud, and SAP Analytics Cloud
- Get to know SAP's artificial intelligence capabilities with SAP AI Launchpad and SAP AI Core

 www.sap-press.com/6122

We hope you have enjoyed this reading sample. You may recommend or pass it on to others, but only in its entirety, including all pages. This reading sample and all its parts are protected by copyright law. All usage and exploitation rights are reserved by the author and the publisher.

The Authors

Dhirendra Gehlot, Jeff Gericke, Shibajee Dutta Gupta, Antony Isacc, Homiar Kalwachwala, Rick Markham, Asim Munshi, and Chris Sam have decades of combined experience with SAP solutions for EIM, analytics, and data management.

ISBN 978-1-4932-2729-7 • 483 pages • 04/2026

E-book: \$84.99 • Print book: \$89.95 • Bundle: \$99.99

 **Rheinwerk**
Publishing