SAP DATASPHERE

DIE CLOUD-DATENMANAGEMENTLÖSUNG VON SAP

ING. MAG. MARIO ROSENFELDER

CONSULTNETWORK GMBH

Rosenheim | Kufsteiner Straße 103 | 83026 Rosenheim
Klagenfurt am Wörthersee | Bahnhofstraße 49 | 9020 Klagenfurt

WWW.CONSULTNETWORK.COM

INHALTSVERZEICHNIS

1 EINLEITUNG 3

2 ARCHITEKTUR 4

2.1 Source Layer 4

2.2 ETL-Tools 5

2.3 SAP Datasphere 5

2.4 Target Layer 6

3 SAP BUSINESS DATA CLOUD 7

4 FUNKTIONALITÄTEN 8

4.1 Datenintegration 8

4.2 Governance 8

4.3 Semantische Modellierung 9

4.4 Catalog 9

4.5 SAP HANA Features 9

4.6 Datasphere als Quellsystem 9

4.7 SAP-Partner 10

5 SCHLUSSWORT 11

ABBILDUNGSVERZEICHNIS

Abbildung 1: SAP Datasphere Architekturschaubild 4
Abbildung 2: SAP Datasphere 5
Abbildung 3: SAP Datasphere als Herzstück der SAP Business Data Cloud 7

1 EINLEITUNG

Die SAP Datasphere ist die integrierten Datenmanagement-Plattform der SAP und ermöglicht es Unternehmen, Daten aus verschiedenen Quellen zu extrahieren, zu verlinken, zu speichern, zu transformieren und für die Verwendung in der SAP Analytics Cloud und anderen BI Tools zu modellieren. Dabei positioniert sich die SAP Datasphere als ein Cloud-Datawarehouse mit umfassenden ETL-Tools und einer sehr guten Connectivity hin zu den SAP eigenen Systemen aber auch hin zu Systemen anderer Hersteller!

In diesem Whitepaper erläutern wir Ihnen anhand des von uns entwickelten Architekturschaubilds (Kapitel 2), wie die SAP Datasphere technisch aufgebaut ist. Im Kapitel 3 gehen wir auf die SAP Business Data Cloud ein. Diese neue Suite vereint die zuvor separaten Systeme SAP Datasphere und SAP Analytics Cloud mit dem neuen SAP Databricks und dem Migrationstool SAP BW[4] Private Cloud Edition unter einem Dach. Im Kapitel 4 beschreiben wir die wichtigsten Funktionen, welche die SAP Datasphere anbietet. Wir gehen auf die nativen Funktionalitäten der SAP Datasphere ein und schauen uns zudem an, in welchen Bereichen Zusatzleistungen von SAP-Partner genutzt werden können.

Um weitere wissenswerte Informationen zur SAP Datasphere, aber auch zur SAP Analytics Cloud, zur SAP Business Data Cloud oder SAP HANA zu erhalten, lohnt es sich, einen Blick auf unsere Website www.controlling-strategy.com sowie auf unseren YouTube-Channel controlling-strategy | consultnetwork zu werfen.

Bei Fragen stehe ich Ihnen jederzeit gerne zur Verfügung!

Ihr Mario Rosenfelder

CONSULTNETWORK GMBH

Rosenheim | Kufsteiner Straße 103 | 83026 Rosenheim
Klagenfurt am Wörthersee | Bahnhofstraße 49 | 9020 Klagenfurt

DE +49 (0)8031 58180 11
AT +43 (0)463 219095

WWW.CONTROLLING-STRATEGY.COM

2 ARCHITEKTUR

Im ersten Kapitel möchten wir Ihnen mithilfe eines von uns entwickelten Architekturschaubildes den Aufbau der SAP Datasphere sowie der vor- und nachgelagerten Komponenten erläutern.

Unser Architekturschaubild gliedert sich in 4 Ebenen: den Source Layer, die ETL-Tools, die SAP Datasphere selbst und den Target Layer, welche nachfolgend beschrieben werden.

Wenn Sie zudem sehen möchten, wie die unterschiedlichen Komponenten im Betrieb funktionieren und wie Sie ein erstes kleines Projekt in der SAP Datasphere realisieren können, dann sehen Sie sich gerne unsere kostenlose Webinaraufzeichnung „SAP Datasphere": Einführung" dazu an.

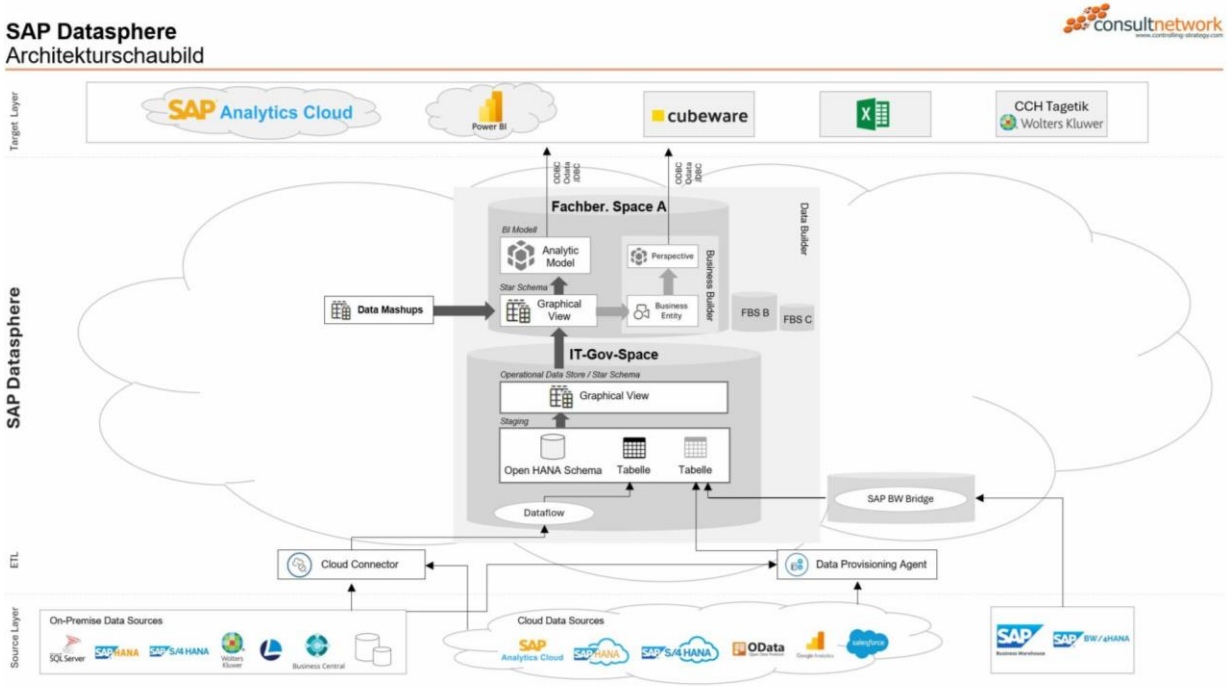


Abbildung 1: SAP Datasphere Architekturschaubild

2.1 SOURCE LAYER

Auf der untersten Ebene finden wir den Source Layer. Im Wesentlichen die unterschiedlichen Datenquellen, auf welche wir mit der SAP Datasphere zugreifen können. Es ist möglich, Daten aus SAP-Systemen sowie Daten aus nicht SAP-Systemen in die SAP Datasphere mithilfe von ETL-Tools hineinzuladen bzw. zu verlinken.

Im Architekturschaubild sehen Sie links On-Premise-Datenquellen wie den Microsoft SQL-Server, SAP HANA, SAP S/4HANA, CCH Tagetik, Lucanet oder auch Business Central. In der Mitte haben wir die Cloud-Datenquellen wie die SAP Analytics Cloud, SAP S/4 HANA in der Cloud, OData oder auch salesforce. Sowohl On-Premise als auch Cloud Datenquellen können mithilfe von ETL-Tools — auf welche wir im nächsten Kapitel näher eingehen — an die SAP Datasphere angebunden werden. An dieser Stelle möchten wir festhalten, dass die hier genannten Datenquellen nur eine beispielhafte Aufzählung darstellen und keinesfalls vollständig sind!

Zusätzlich kann auch ein SAP Business Warehouse (SAP BW, SAP BW/4HANA) über eine sogenannte SAP BW Bridge angebunden werden, was im Schaubild ganz rechts unten zu sehen ist.

2.2 ETL-TOOLS

Auf der Ebene der ETL-Tools haben wir den Cloud Connector und den Data Provisioning Agent, beide werden im Netzwerk installiert. Die beiden ETL-Tools ermöglichen die Anbindung von Daten aus den Vorsystemen an die SAP Datasphere. Mit den Werkzeugen können Remote-Tabellen in die Datasphere verknüpft oder Daten durch einen Dataflow direkt in die Datasphere hineingeschrieben werden.

2.3 SAP DATASPHERE

Die SAP Datasphere selbst ist in zwei Spaces gegliedert: Den IT-Governance-Space, über welchen die IT- bzw. BI-Abteilung die Hoheit hat, und den Fachbereichs-Space. Im Frontend greift man über den Data Builder auf die Spaces zu.

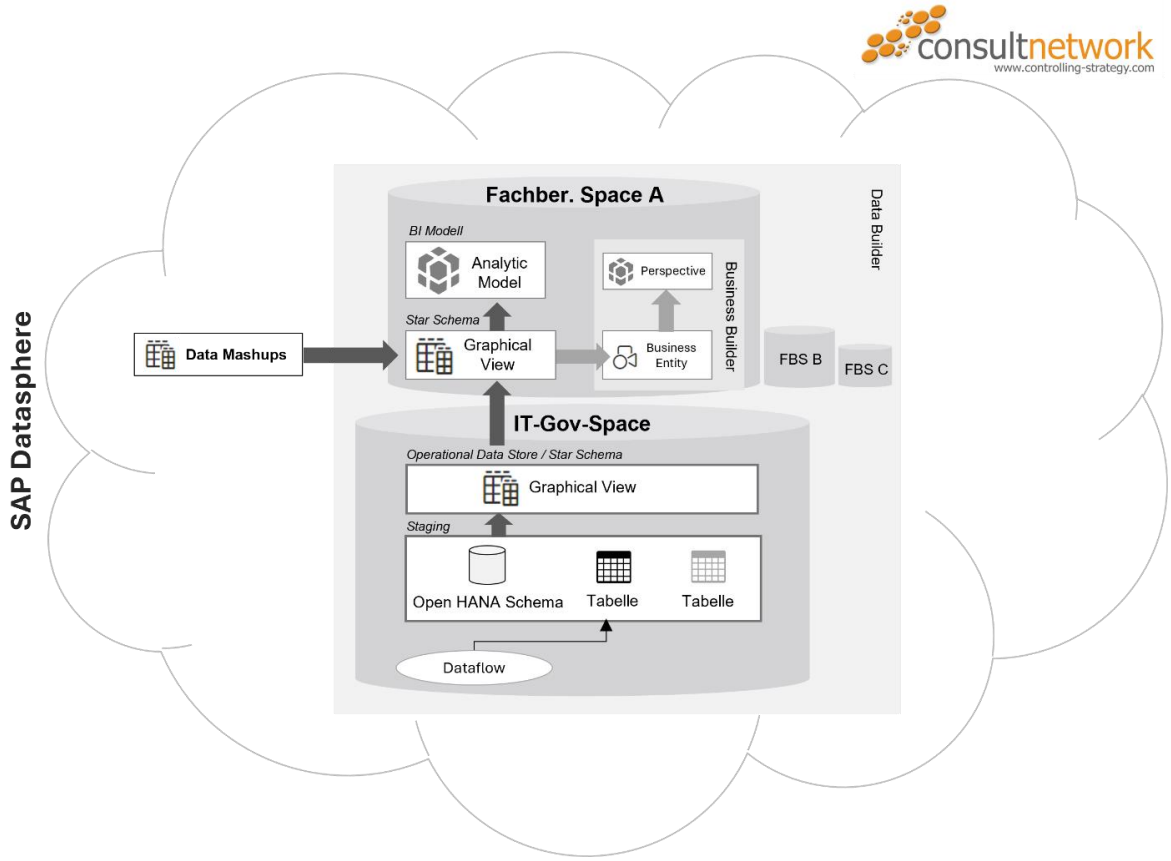
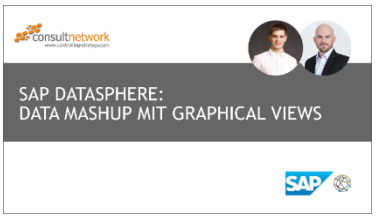


Abbildung 2: SAP Datasphere

Die Spaces folgen dem typischen Datawarehouse-Aufbau, bestehend aus Staging, Operational Data Store, Star Schema und BI-Modellen.

Im IT-Governance-Space werden die Basisdaten zunächst in der Staging-Schicht zusammengezogen und dann mithilfe von Graphical Views im Operational Data Store (SPOT) aufbereitet. Das Star Schema und die BI-Modelle können sowohl im IT-Governance-Space als auch im Fachbereichs-Space konfiguriert werden. Im Fachbereichs-Space können zudem Daten ergänzt und ein Data Mashup durchgeführt werden.

Sollten Sie mehr über Data Mashups in der SAP Datasphere erfahren wollen, schauen Sie sich gern unsere kostenlose Webinaraufzeichnung zu diesem Thema an.



Zusätzlich gibt es im Fachbereich-Space auch noch den Business Builder. Mit dem Business Builder kann der Fachbereich Business Entites (Star Schema) erstellen, also Dimensionen, Hierarchien, Fakten und Kennzahlen definieren, die dann in das Consumption Model, welches das BI Model darstellt, weitergegeben werden. Auf diese Consumption Models kann nicht als Ganzes zugegriffen werden. Man verwendet die sogenannten Perspectives als Abfragen auf das Model, auf die dann mit externen Tools zugegriffen werden kann. Im Business Builder wird rein mit Business Kontext gearbeitet und nicht mit „kalten" Daten, wie es im Data Builder der Fall ist. Damit ist der Business Builder für weniger technikaffine User geeignet.
Nachdem die BI Modelle im Data Builder „Analytic Models", im Business Builder die „Perspektives" erstellt wurden, können diese in unterschiedlichen Reporting-Tools dargestellt werden.

2.4 TARGET LAYER

Auf der obersten Ebene, dem Target Layer, finden wir die unterschiedlichen Reporting-Tools. Die Tools zapfen via ODBC, OData oder JDBC die BI-Modelle aus der SAP Datasphere an.

Dabei ist die SAP Analytics Cloud (SAC) sehr eng mit der SAP Datasphere integriert, da die Modelle, die normalerweise in der SAC erstellt werden, jetzt aus der SAP Datasphere kommen und optimal für die SAC aufbereitet werden. Aber auch die Anbindung anderer Drittherstellertools wie bspw. Microsoft Power BI, Cubeware oder CCH Tagetik von Wolters Kluwer funktioniert problemlos.

3 SAP BUSINESS DATA CLOUD

Anfang 2025 wurde der Release der neuen SAP Business Data Cloud bekanntgegeben. Diese neue Suite vereint die zuvor separaten Systeme SAP Datasphere und SAP Analytics Cloud mit dem neuen SAP Databricks und dem Migrationstool SAP BW[4] Private Cloud Edition unter einem Dach.

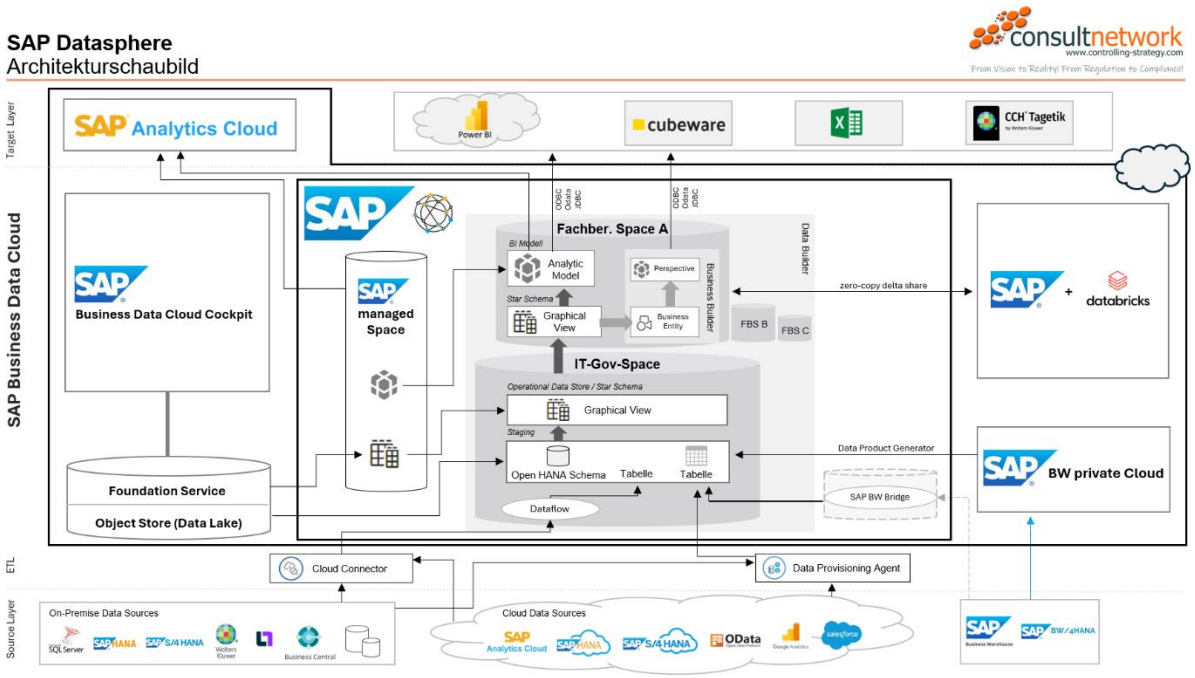


Abbildung 3: SAP Datasphere als Herzstück der SAP Business Data Cloud

In diesem Kapitel stellen wir dar, welche Rolle die SAP Datasphere im Kontext der SAP Business Data Cloud spielt. Wenn Sie detaillierte Informationen zur SAP Business Data Cloud wünschen, besuchen Sie gerne unsere Webseite.

Die SAP Business Data Cloud befindet sich, wie es der Name schon verrät, ausschließlich in der Cloud. Dies war der nächste natürliche Schritt bei der Cloud-First-Strategie von SAP, um den Kunden den Weg von den alten On-Premise Systemen zur Cloud zu erleichtern. SAP Datasphere spielt dabei eine ganz besondere Rolle. Als zentrale Instanz der Integration, Modellierung und Bereitstellung der Daten stellt die SAP Datasphere das Herzstück der SAP Business Data Cloud dar.

Eng verbunden ist die SAP Datasphere mit der **SAP BW[4] Private Cloud Edition**, die das alte On-Premise SAP BW oder SAP BW4/HANA nach einem Lift in die Cloud darstellt. In der Folge kann die Zeit genutzt werden, um die Investments im alten SAP BW in das neue SAP Datasphere System zu shiften und die wichtigsten Modellierungsbausteine auch in Zukunft verwenden zu können. Damit wird eine sanfte Migration des alten SAP BW in die Cloud ermöglicht. Die Komponente SAP BW Bridge ist damit obsolet, wird aber weiterhin unterstützt, sollte man damit bereits eine Migration begonnen haben.

Das neue **SAP Databricks** ist als eine Komponente in SAP Business Data Cloud aus der Partnerschaft zwischen SAP und Databricks entstanden und ebenfalls eng mit der SAP Datasphere verbunden. Die SAP Datasphere wirkt als zentraler Hub, aus dem Daten via Zero Copy Delta Share bereitgestellt werden, um in der Pro Code Umgebung von SAP Databricks Maschine Learning Algorithmen umsetzen zu können und die Ergebnisse wieder in den zentralen Datenspeicher zurückzuspielen. Damit ergeben sich völlig neue Möglichkeiten, auf Fragen der Entscheidungsträger Antworten zu finden.

Nicht zuletzt wäre auch die **SAP Analytics Cloud** zu erwähnen, die vorher schon eng mit der SAP Datasphere vernetzt war. Es ist schon seit einer Weile möglich, SAP Datasphere Modelle für Stories zu

verwenden, ohne separat ein Modell in der SAP Analytics Cloud definieren zu müssen. Mit der allgemeinen Verfügbarkeit der SAP Business Data Cloud wird nun aber auch das Feature des Seamless Planning allgemein verfügbar. Das ermöglicht die Datenhaltung von Planungsmodellen in der SAP Analytics Cloud in SAP Datasphere. Damit lassen sich Ergebnisse der Unternehmensplanung leichter und schneller in das zentrale Datenmodell integrieren. Vergleiche mit Actuals oder mit Simulationen werden damit leichter als je zuvor.

In der SAP Business Data Cloud werden zudem die **Intelligent Applications** angeboten. Das sind von SAP bereitgestellte und gewartete Stories in der SAP Analytics Cloud, die mit Datenmodell, Integrationen zu Quellsystemen und Berechtigungskonzepten geliefert und nach Best Practise Ansätzen aufgebaut sind. Die Modelle werden dabei in einem separaten von SAP verwalteten Space innerhalb der SAP Datasphere gespeichert und können kopiert und anschließend an eigene Anforderungen angepasst werden.

4 FUNKTIONALITÄTEN

In diesem Kapitel beschreiben wir Ihnen die nativen Funktionen, die die SAP Datasphere anbietet. In einigen Punkten können auch Zusatzleistungen von SAP-Partner in Anspruch genommen werden, die eng mit SAP Datasphere integrierte Lösungen anbieten. Diese werden ebenfalls in einem Unterkapitel behandelt.

4.1 DATENINTEGRATION

Eine wesentliche Stärke der SAP Datasphere besteht darin, viele verschiedene Datenquellen ohne komplexe Konfiguration anbinden zu können. Das gilt sowohl für SAP-Systeme als auch für Nicht-SAP Systeme. Dabei ist es nicht nur möglich, sondern sogar Best Practise, auf die Quellsysteme live via Remote-Tabellen zuzugreifen. Die Daten liegen damit stets aktuell in der SAP Datasphere vor. Dies ermöglicht Ihnen zeitnahe Analysen von neuen Daten im Quellsystem.

In großen Systemlandschaften kommt es jedoch vor, dass Tabellen zu groß sind, um eine gute Performance der Datenmodelle bei Live-Zugriff zu ermöglichen. In diesen Fällen bietet Ihnen die SAP Datasphere mit den Dataflows ein vollumfängliches Werkzeug für den ETL-Prozess. Zusätzlich besteht die Möglichkeit, für SAP-Systeme den Replication Flow zu verwenden. Dieser eignet sich besonders, wenn mit einem Massendatenload eine größere Anzahl von Tabellen aus dem Quellsystem übernommen werden soll, wobei festzuhalten ist, dass hier die Transformationsfunktionen eingeschränkt sind.

Zusätzlich gibt es die Möglichkeit, von Drittanbietern Daten aus dem Data Marketplace in das eigene System zu integrieren. Dabei kann es sich um öffentlich zugängliche Statistiken handeln oder auch um Datenpakete von Partnern, die nur privat zur Verfügung gestellt werden. Innerhalb eines Konzerns können so auch einzelne Firmen untereinander Daten austauschen. Werden öffentliche Daten in das eigene System integriert, kann es manchmal schwierig sein, diese Daten mit den eigenen Daten zu harmonisieren. In der SAP Datasphere hilft dabei der Intelligent Lookup. Damit können Spalten, die das Gleiche aussagen aber unterschiedlich definiert sind, aufeinander abgestimmt werden.

4.2 GOVERNANCE

Historisch ist die IT-Abteilung für den Schutz der Daten im Unternehmen verantwortlich gewesen. Wenn Fachbereiche Anpassungen an Datenmodellen gewünscht haben, kam es nicht selten zu Verzögerungen durch Bottlenecks bei der IT. Aufgrund der zum Teil recht hohen Wartezeiten hat sich bei den Fachbereichen mehr und mehr das Bedürfnis gebildet, selbstständig in den Aufbau der Berichte und auch in die Modellierung der Daten einzugreifen. Damit wird die IT, die nach wie vor für den Datenschutz Verantwortung trägt, nicht selten in Verlegenheit gebracht.

Mit der SAP Datasphere kann dieser Konflikt aufgelöst werden. Das Konzept der Spaces hilft Ihnen dabei, die Aktivitäten vom Fachbereich von den Aktivitäten der IT abzugrenzen und gleichzeitig dafür Sorge zu tragen, dass die Daten aus dem zentralen Datawarehouse optimal genutzt werden können.

Dabei kann der Single Point of Truth (SPOT) gepflegt und Rowlevel Security eingestellt werden, sodass jeder User nur kontrollierte und freigegebene Daten einsehen kann, die für ihn relevant sind.

Mit den Fachbereichen können dann bestimmte Datenpakete geteilt werden. Innerhalb des eigenen Spaces hat dann der Fachbereich die Freiheit, eigene Modelle zu erstellen oder auch selbstständig Daten zu importieren. Dieses Modell ist zudem flexibel auf den jeweiligen Use Case anpassbar. Wenn der Fachbereich Interesse an fertigen BI-Modellen hat, können diese ebenfalls im IT-Space erstellt und dann dem Fachbereich zur Verfügung gestellt werden.

4.3 SEMANTISCHE MODELLIERUNG

Wenn Sie eine Tabellenspalte mit einer Zahl sehen, ist es nicht immer sofort ersichtlich, was die Spalte aussagt. Handelt es sich um eine Kundennummer? Ist es ein Umsatz? Vielleicht ist auch ein Datum gemeint. Um Eindeutigkeit für die Daten zu schaffen, benötigt es eine saubere Semantik bei der Modellierung. Dies ist eine weitere Kernkompetenz der SAP Datasphere.

Sie haben aber nicht nur volle Kontrolle über die Semantik neuer Modelle, denn eine große Stärke der SAP Datasphere ist die Übernahme der Semantik aus dem Quellsystem. Damit müssen Sie nicht nochmals Modelle gänzlich neu erstellen, die es bereits in der Datenquelle gegeben hat. Speziell mit dem Semantic Onboarding können Sie Objekte aus angebundenen SAP-Systemen (z.B. CDS-Views) importieren und wiederverwenden.

Dazu gehört auch die Möglichkeit, Membertexte in verschiedenen Sprachen und Hierarchien darzustellen. Die hier erstellten Logiken sind vielseitig wiederverwendbar.

4.4 CATALOG

In einem komplexen Datawarehouse kann es mitunter sehr schwierig sein, ein bestimmtes Objekt wiederzufinden. Das Catalog Feature verschafft dabei Abhilfe. Mit detaillierter Such- und Filterfunktionalität lassen sich gezielt Objekte finden und wiederfinden.

Zusätzlich können Sie ein Glossar anlegen. Hier können Begriffe, die im Geschäftskontext verwendet werden, erklärt und zu bestimmten Datenobjekten gemapped werden. Damit lässt sich sicherstellen, dass jeder Business User mit den Begrifflichkeiten vertraut ist und schnell zu den Daten Zugriff hat, die für den jeweiligen User relevant sind.

4.5 SAP HANA FEATURES

In jedem Datasphere System ist ein SAP HANA Cloud Tenant enthalten, auf die Definitionen der Remote Tabellen, Views und auch persistierte Daten gespeichert sind. Innerhalb eines Space kann ein Datenbank User eingestellt werden, der direkten Zugriff auf die Datenbank via Database Explorer erhält. Damit ist es auch möglich, mit der SAP Datasphere verschiedene SAP HANA Features zu verwenden. Darunter fällt die prozedurale Datentransformation, das Jobscheduling und unter anderem können auch SAP HANA Predictive Libraries oder SAP HANA Graph Engine verwendet werden. Die Tabellen oder Views, die als Ergebnis entstehen, können zudem ohne weitere Bearbeitungsschritte im Data Builder der SAP Datasphere verwendet werden.

4.6 DATASPHERE ALS QUELLSYSTEM

Nicht nur Daten können in die SAP Datasphere integriert werden, sondern auch die Integration von Daten aus der SAP Datasphere in andere Systeme ist möglich. Besonders mit der SAP Analytics Cloud ist die SAP Datasphere sehr eng verbunden. Vor allem die Live-Anbindung der BI Modelle in die SAC erfolgt nahtlos, ohne Erstellung eines zusätzlichen Modells innerhalb der SAC. Es lassen sich aber auch Daten in Planungsmodelle importieren. Um die Planung ebenso nahtlos zu gestalten, stehen in Zukunft weitere Entwicklungen in der SAP Datasphere auf der Roadmap.

Zusätzlich können OAuth Clients angelegt werden, die den Zugriff externer Systeme auf Datasphere Modelle via OData zulassen. Bei diesen OData Schnittstellen wird sehr auf Sicherheit geachtet, sodass kein unbefugter Zugriff auf die Daten ermöglicht wird.

Zudem kann via JDBC und ODBC auf die Daten der SAP Datasphere zugegriffen werden. Da hier eine SAP HANA Cloud Datenbank vorliegt, können Sie dafür die normalen SAP HANA Treiber verwenden.

Die Replication Flows können neben dem Import zusätzlich auch für den Export von Daten verwendet werden. Da die Quelle und das Ziel flexibel einstellbar sind, lässt sich auch ein Use Case realisieren, in dem die SAP Datasphere nur als Durchfluss verwendet wird. In diesem Fall wären externe Systeme sowohl die Quelle als auch das Ziel des Replication Flows.

4.7 SAP-PARTNER

Wie erwähnt arbeitet SAP eng mit einigen Partnerfirmen zusammen, um die SAP Datasphere in bestimmten Aspekten noch leistungsfähiger zu machen und an die Bedürfnisse der Kunden anzupassen.

Ein Partner ist Collibra, der einen erweiterten Datenkatalog bereitstellt, mit dem Abhängigkeiten der Daten über Systeme hinweg, SAP und Nicht-SAP, zurückverfolgt werden können. Das macht Collibra zu einem wirksamen Tool beim Aufbau eines Business Data Fabrics, wo jeder Business User schnell zu den Daten gelangt, die er benötigt und zugleich effektive Governance über die Daten Assets ermöglicht.

Confluent bietet eine Data Streaming Plattform, die den Zugriff auf SAP-Daten in Echtzeit mit Apache Kafka Technologie ermöglicht. Damit ist es möglich die Zeit zwischen der Schaffung neuer Daten und der Nutzung für Analyse, Simulationen und letztendlich Entscheidungsfindung zu minimieren.

DataRobot stellt den Kunden eine Plattform zur Verfügung, auf der AI-getriebene Applikationen schnell und effizient erstellt und verwendet werden können. Dabei werden sowohl Pro Code als auch No Code Ansätze unterstützt.

Die Partnerschaft zwischen SAP und Databricks geht noch einen Schritt weiter. In der neuen SAP Business Data Cloud wurde, wie oben schon erwähnt, eine Komponente eingebaut, die die Verbindung zwischen SAP und Databricks untermauert. In SAP Databricks können die Modelle aus SAP Datasphere einfach geladen und in Maschine Learning Modellen verwendet werden, um wertvolle Erkenntnisse aus den SAP-Daten zu ziehen.

5 SCHLUSSWORT

Wir hoffen, dieses Whitepaper konnte dabei unterstützt, den technischen Aufbau der SAP Datasphere verständlicher zu machen. Mit der SAP Datasphere bekommen Sie eine integrierte Datenmanagement-Plattform mit umfassenden ETL-Tools, um Daten aus verschiedenen Quellen zu extrahieren, zu verlinken, zu speichern, zu transformieren, um diese für die Verwendung in verschiedenen BI-Tools zur Verfügung zu stellen. Hinsichtlich der Integrität der unterschiedlichen BI-Tools sticht vor allem die sehr gute Connectivity zu den SAP-eigenen Systemen hervor. Aber auch nicht SAP-Systeme können problemlos mit der SAP Datasphere verwendet werden.

Wenn Sie wissen möchten, ob die SAP Datasphere die für Ihr Unternehmen passende Lösung ist oder Sie Interesse an der SAP Analytics Cloud haben, dann buchen Sie über meinen Buchungskalender gerne eine unverbindliche, kostenlose Anforderungsbesprechung mit mir oder kontaktieren Sie mich direkt.

Mehr über die SAP Datasphere erfahren Sie auch auf unserer Webseite, auf unserem YouTube-Channel oder über meinen LinkedIn-Newsletter „Unternehmenssteuerung", über welchen ich monatlich einen neuen Artikel veröffentliche.

Besten Dank für Ihr Interesse!
Ihr Mario Rosenfelder

CONSULTNETWORK GMBH

Rosenheim | Kufsteiner Straße 103 | 83026 Rosenheim
Klagenfurt am Wörthersee | Bahnhofstraße 49 | 9020 Klagenfurt

DE +49 (0)8031 58180 11
AT +43 (0)463 219095

WWW.CONSULTNETWORK.COM