



Data Mesh & Data Vault @ Deichmann

27.11.2025



Das sind wir!



Anh Le
Data Engineer
IT Data & Analytics

- DataOps Advocate
- data4good Volunteer
- Hobby: Musicals & Theater



Sabine Siebenmorgen
Data Engineer
IT Data & Analytics

- Agile Mindset Driver
- Databricks Crafter
- DV2 Modelling Enthusiast
- Hobbies: photography, board games



Sebastian Ennigkeit
Developer
IT Data & Analytics

- DV2 Practitioner
- Agile Enthusiast
- Hobbies: Music, Travel, Basketball

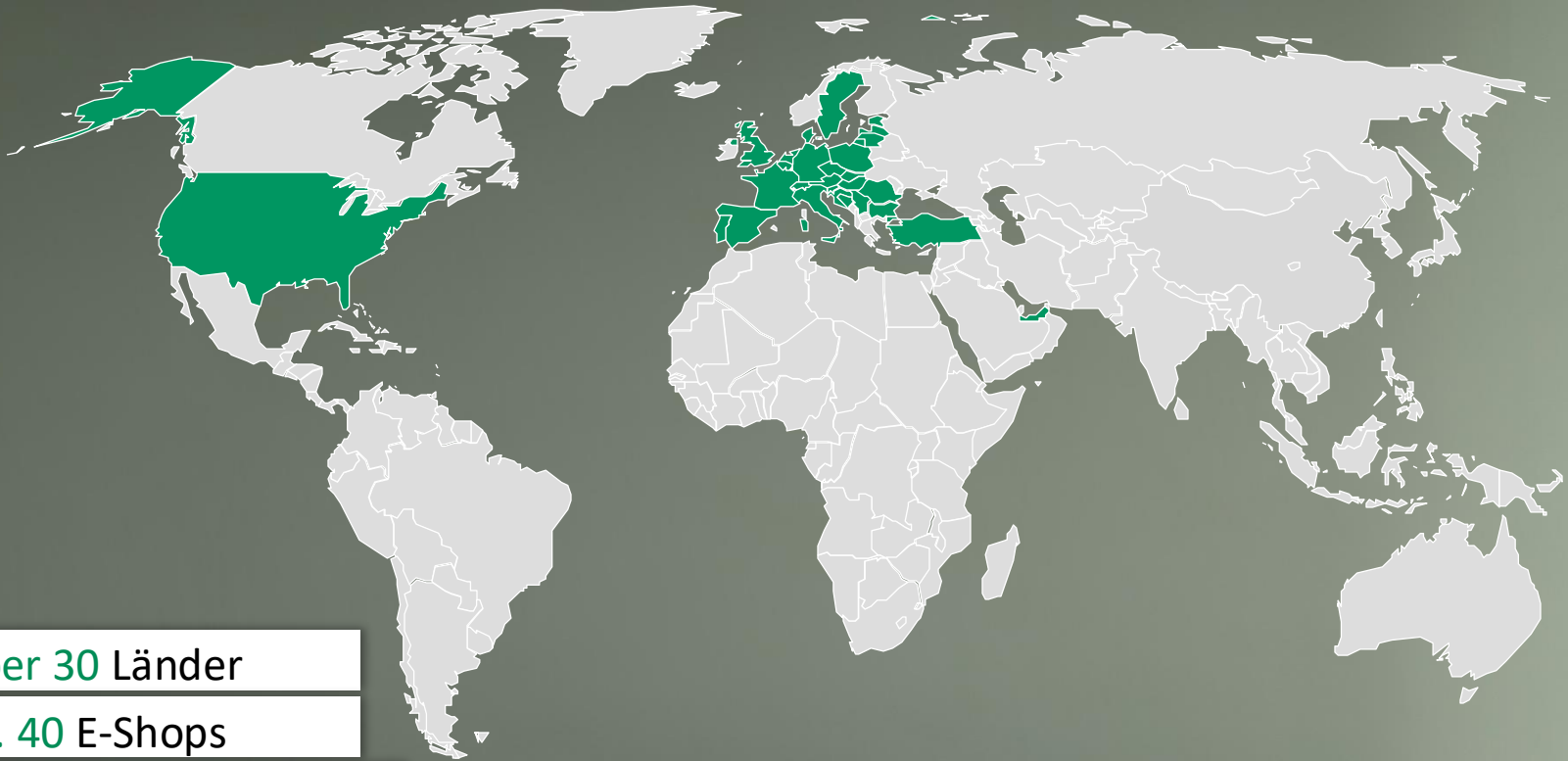


Kevin Haferkamp
Head of Data Platform
& Engineering CoE
IT Data & Analytics

- DV2 Practitioner
- Data Modelling Afficionado
- Functional Programming
- Hobbies:



Nr. 1 in Europa



über 30 Länder

ca. 40 E-Shops

ca. 4.700 Filialen

ca. 49.900 Mitarbeitende

über 180.000.000 Paar Schuhe

ca. 8.700.000.000 Euro Umsatz

D DEICHMANN
GROUP

DEICHMANN



DOSENBACH

snipes®

vanHaren

solebox.

OCHSNER
SHOES

supremo

OCHSNER
SPORT

urbanstyles.



DEICHMANN: Ein Familienunternehmen



Heinrich Deichmann († 1940)

- Gründete das Unternehmen 1913 im Ruhrgebiet
- Startete mit einem Schuhmacherladen im Arbeiterbezirk Essen-Borbeck



Dr. Heinz-Horst Deichmann († 2014)

- 1956 übernahm der Sohn des Gründers die Leitung der Firma mit Filialen in Düsseldorf und Oberhausen
- Baute das Unternehmen zu Europas größtem Schuheinzelhändler auf



Heinrich Deichmann

- Der Enkel des Gründers leitet das Unternehmen seit 1999
- Unter seiner Leitung werden die internationale Expansion, die Digitalisierung und der Omnichannel-Commerce stetig voran getrieben
- Darüber hinaus setzte er im Einkauf konsequent die Vertikalisierung der Beschaffung um



Samuel Deichmann

- Seit 2020 ist auch der Sohn von Heinrich Deichmann und Urenkel des Gründers Mitglied der Geschäftsleitung





01.

Data Mesh @ Deichmann

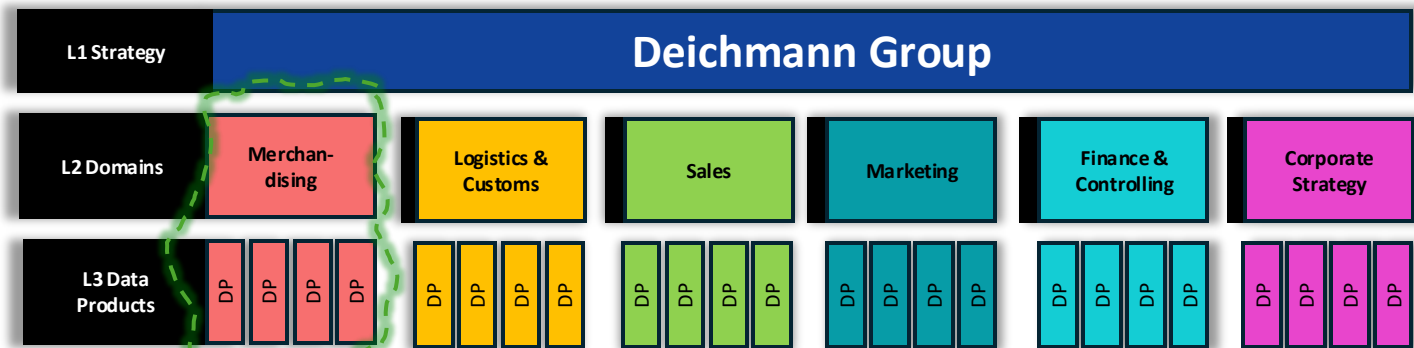


Deichmanns “Road to Data Mesh”

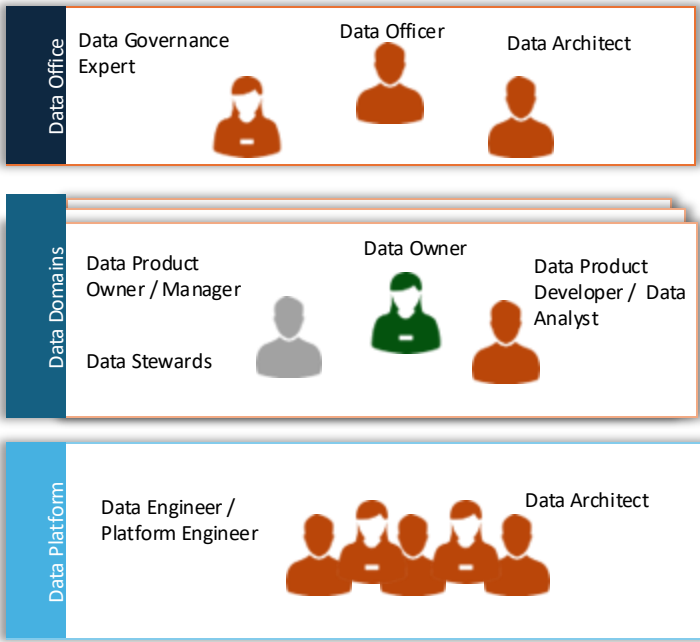


Data Domains* @ Deichmann

Auszug von Domänen in der Deichmann Gruppe, mit ihren jeweiligen Datenprodukten



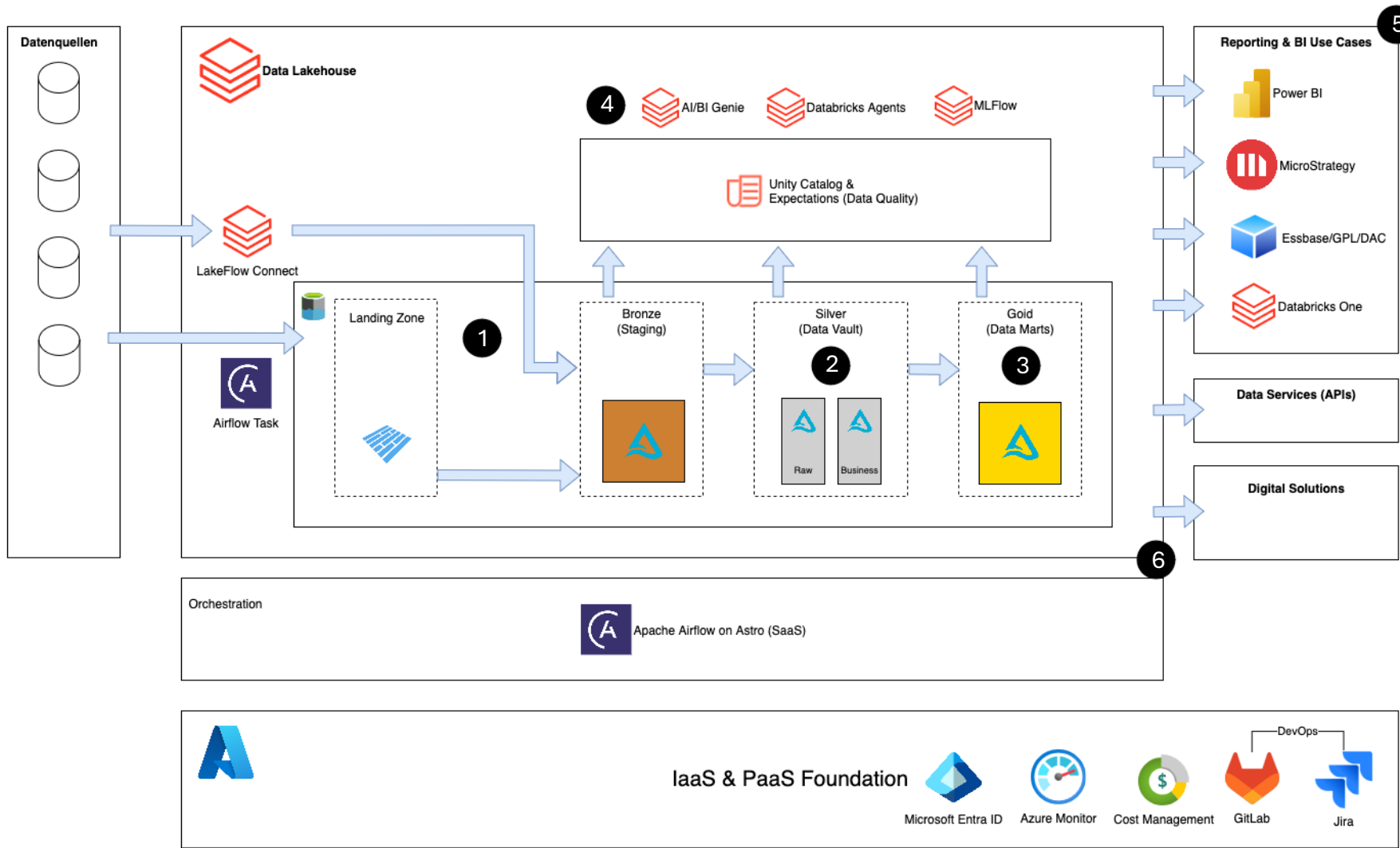
Exemplarische Rollen in einer Domäne



*unvollständig und stark vereinfacht



Deichmanns Data Analytics Platform (DAP)

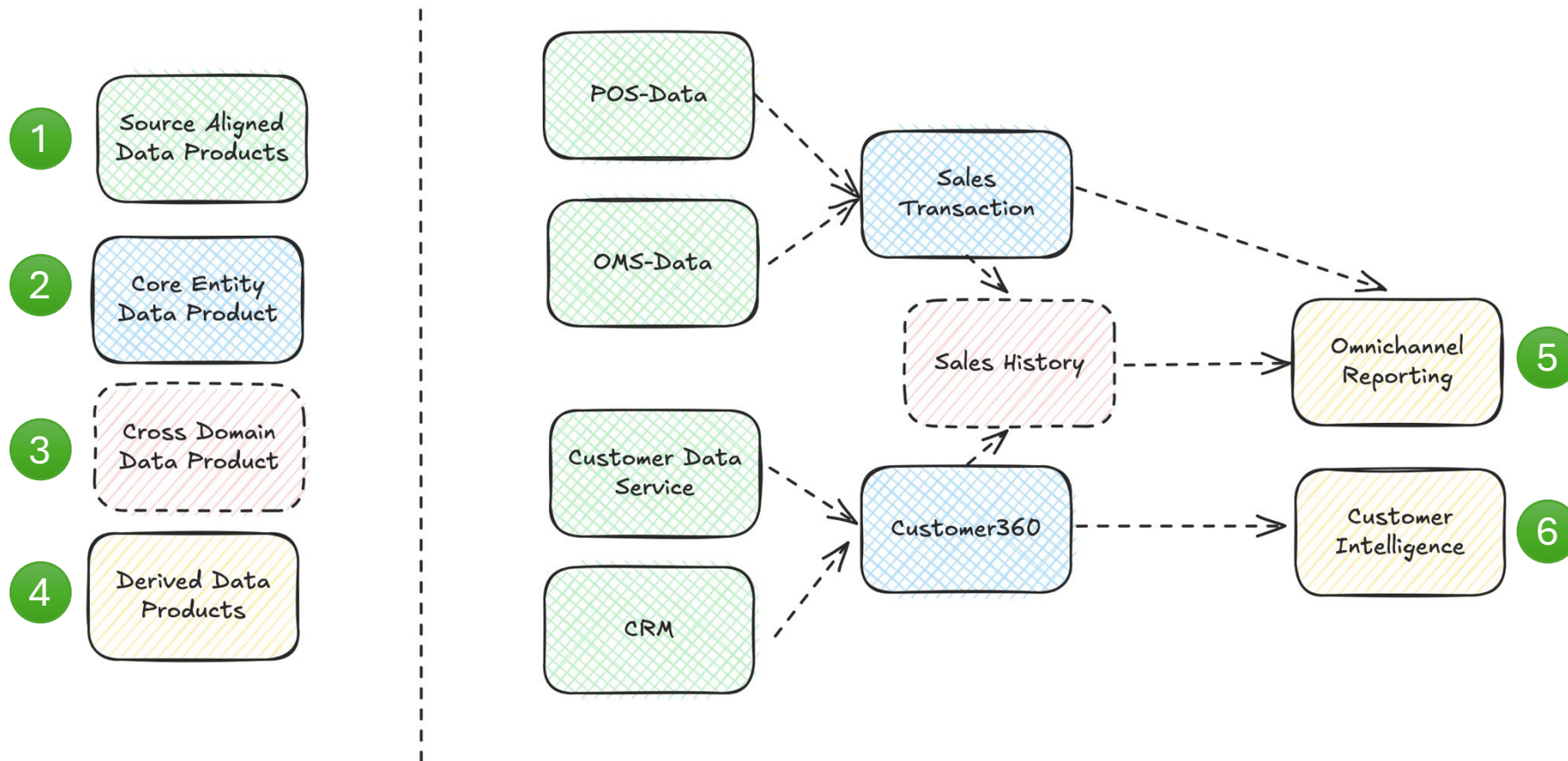


- 1 Data Engineers ingest data into the Bronze layer.
- 2 Data Modelers design and evolve the Data Vaults.
- 3 Data Analysts deliver insights from the Gold layer.
- 4 ML Engineers develop and deploy machine learning models.
- 5 BI Engineers transform data into dashboards and reports.
- 6 Platform Engineers build and maintain the infrastructure that powers the platform.



Data Products & Data Assets

Bei Deichmann verfolgen wir eine Datenproduktstrategie, in der wir die **Datenprodukte inkl. ihrer Data Assets** bündeln und eine **klare Ownership** aufbauen (wollen). Die Basis bietet unser **Data Lakehouse** auf Azure Databricks und der **Medaillon Architecture**.



1. Qualitätsgesichertes Datenprodukt, angelehnt an die Datenquelle
2. Core-Datenprodukt mit Data Vault im Silver-Layer
3. Datenprodukt zur Verbindung von 1:n Domänen
4. Kombiniertes Datenprodukt aus Vorstufen (in der Regel nur noch Gold-Layer)
5. Beispiel der Reporting Landschaft für die Omnichannel Geschäftssteuerung
6. Beispiel für Produkte um Customer Segmentation, Churn Forecast etc.



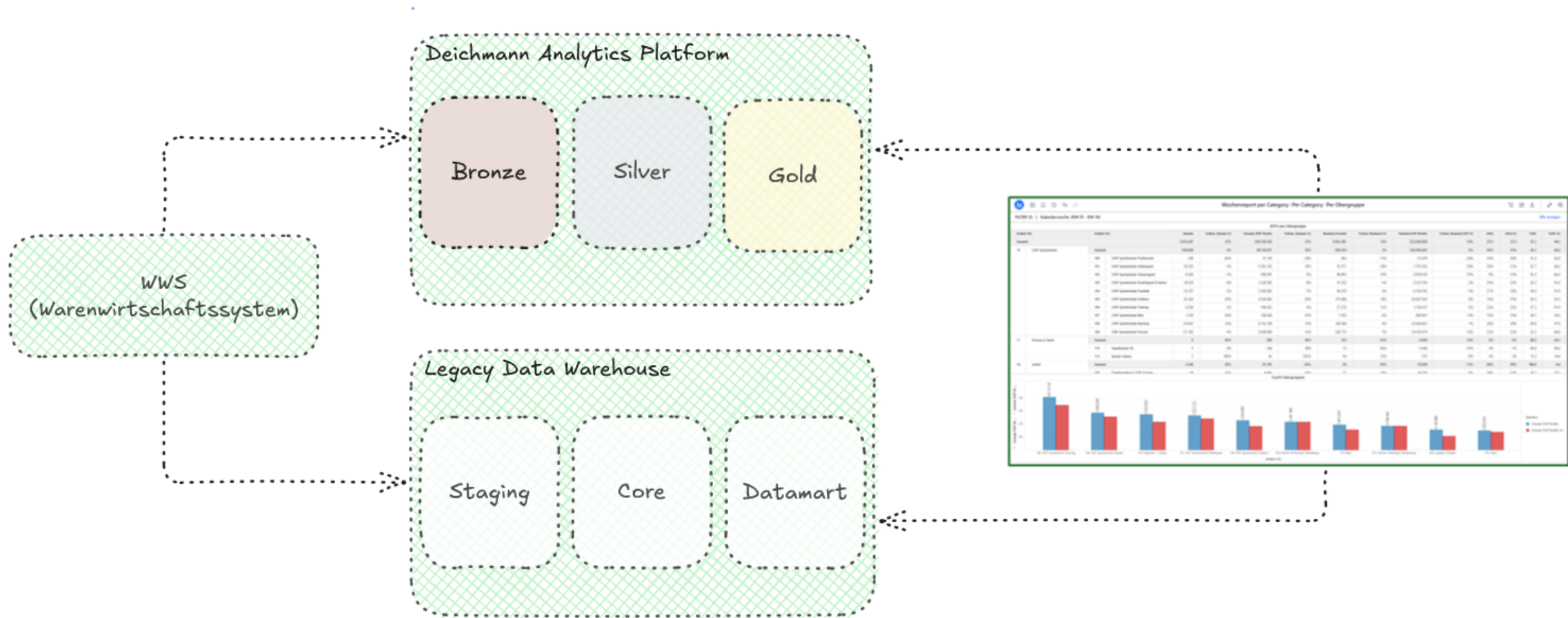
02.

Use Case “DWH Migration”



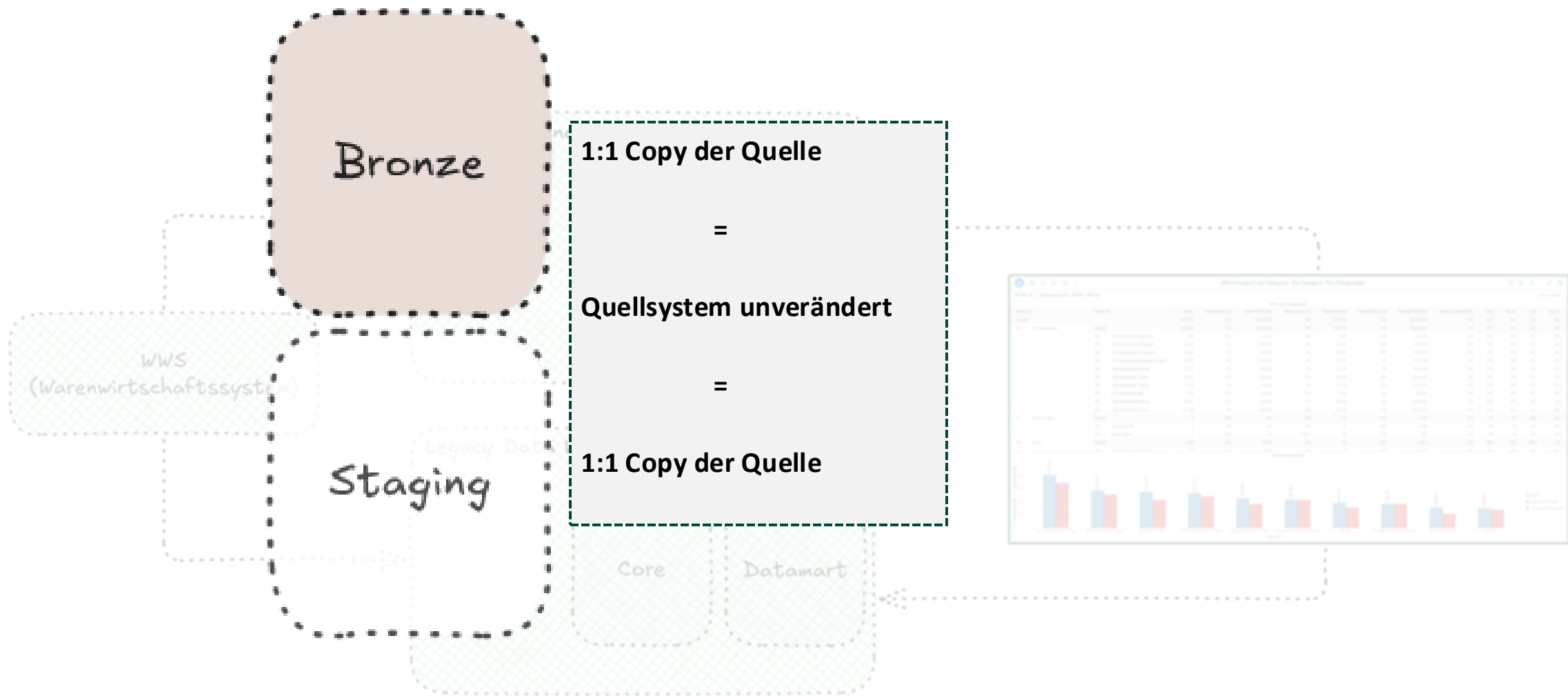
Ausgangslage

Migration der Datenbasis eines Dashboards vom Legacy DWH auf neue Analytics Platform.



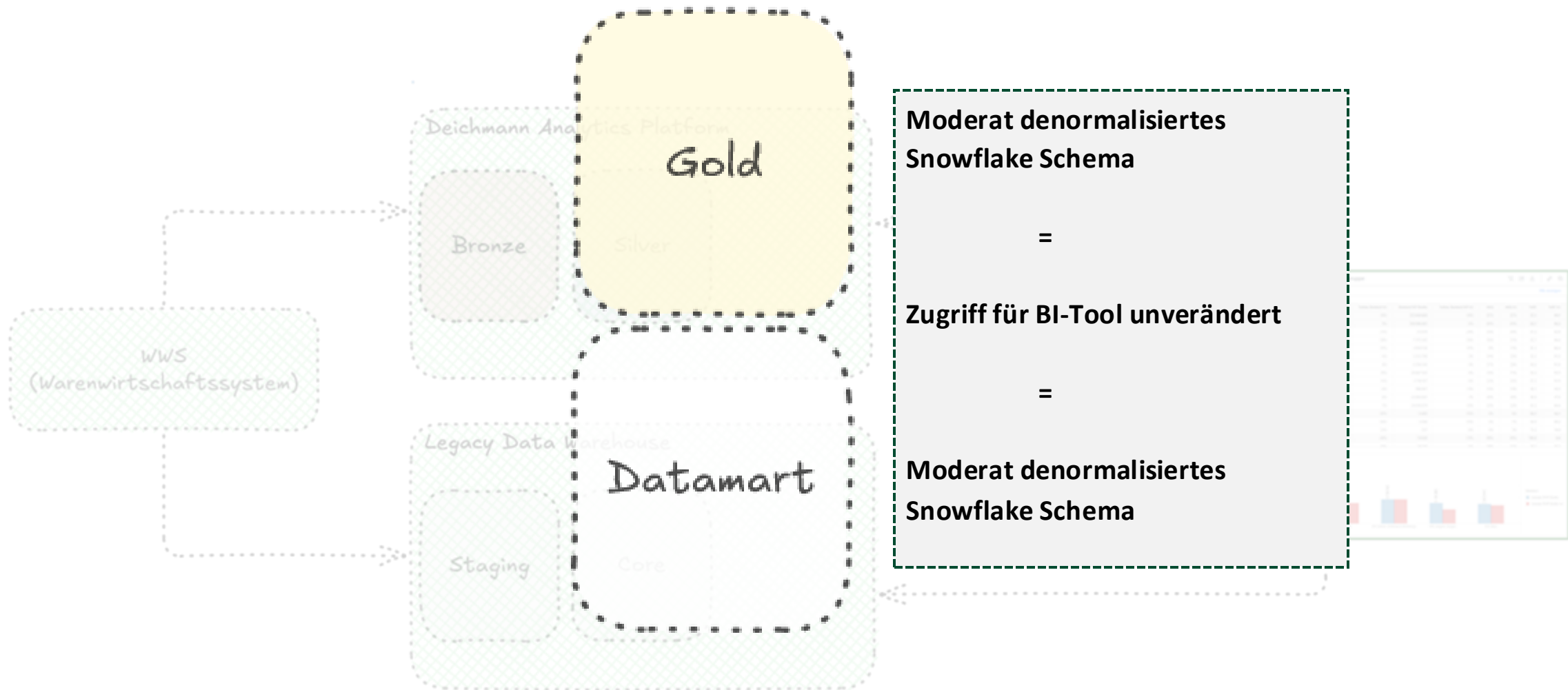
Ausgangslage

Migration der Datenbasis eines Dashboards vom Legacy DWH auf neue Analytics Plattform.



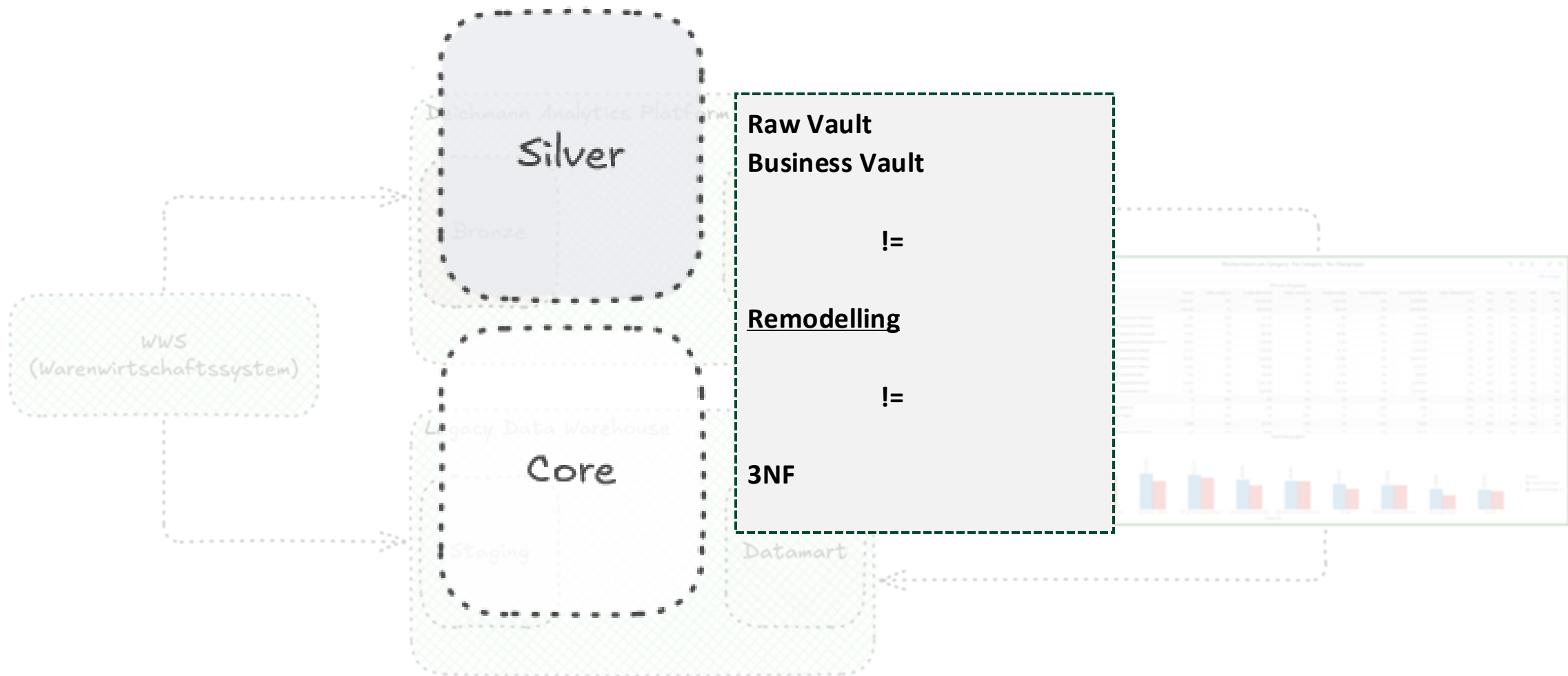
Ausgangslage

Migration der Datenbasis eines Dashboards vom Legacy DWH auf neue Analytics Plattform.



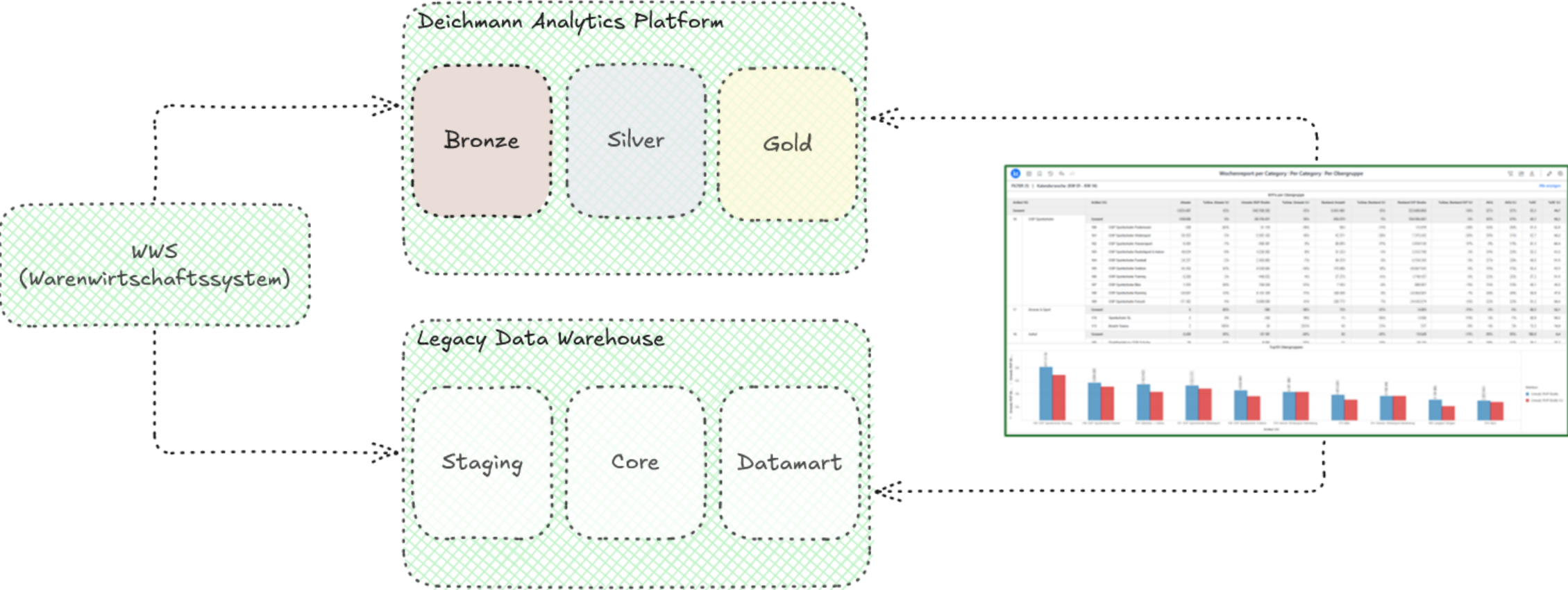
Ausgangslage

Migration der Datenbasis eines Dashboards vom Legacy DWH auf neue Analytics Plattform.



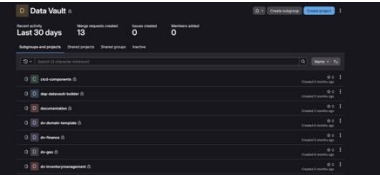
Ausgangslage

Migration der Datenbasis eines Dashboards vom Legacy DWH auf neue Analytics Platform.



Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.



```
1 orgein_company_organization:
2   source_table: wms_bronze.orgein
3   source_type: bronze
4   transformations:
5     filters:
6       - subtyp = 'F'
7     renames:
8       - [nr_benutzersicht, bk_nr_benutzersicht_company]
9       - [STAAT_NR_DEKLARANT, bk_nr_legalentity]
10      - [DAT_AKTIV_VON, eff_from]
11      - [DAT_AKTIV_BIS, eff_to]
12   entity_metadata:
13     rec_src: wms.orgein
```

```
1 company:
2   source_name: orgein_company_organization
3   business_keys:
4     - bk_nr_benutzersicht_company
5   table_properties:
6     source_system: wms
7   description: Master data for organizational units of type company
```

```
1 company:
2   - sat_name: active
3     sat_type: effectivity
4     table_properties:
5       source_system: wms
6       description: Effectivity tracking of company records to identify active or valid periods
7   - sat_name: reportctrl
8     sat_type: standard
9     columns:
10       - nr_orderung
11       - ref_cds_relevant
12     table_properties:
13       source_system: wms
14       description: Attributes controlling reporting relevance or system inclusion of companies
15   - sat_name: details
16     sat_type: standard
17     columns:
18       - nr
19       - subtyp
20       - ref_tip
```

```
1 company_currency:
2   source_name: join_company_currency
3   link_type: standard
4   hubs:
5     driver:
6       - company
7     secondary:
8       - currency
9   effectivity_satellite:
10     status: active
11   table_properties:
12     source_system: wms
13   description: Link between companies and their main currencies
```

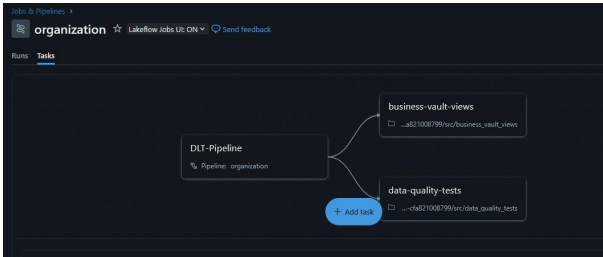


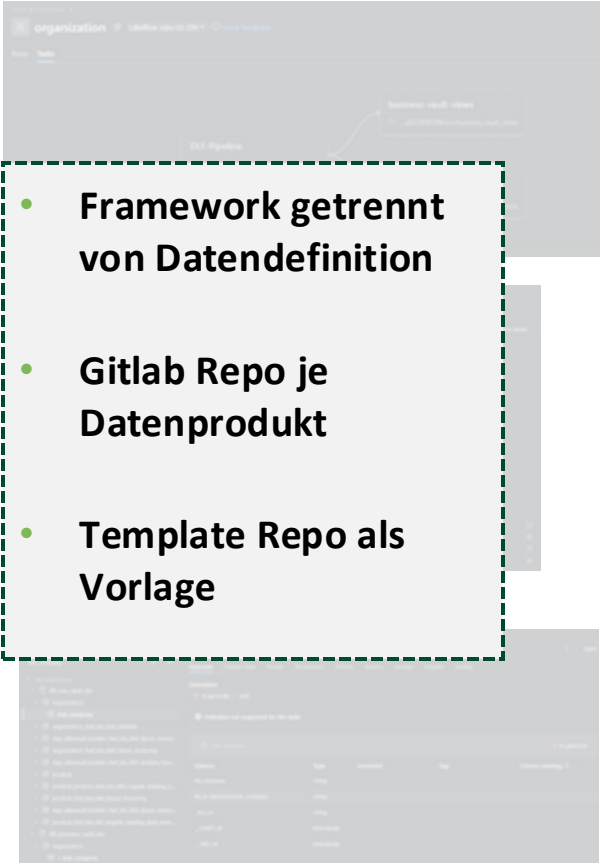
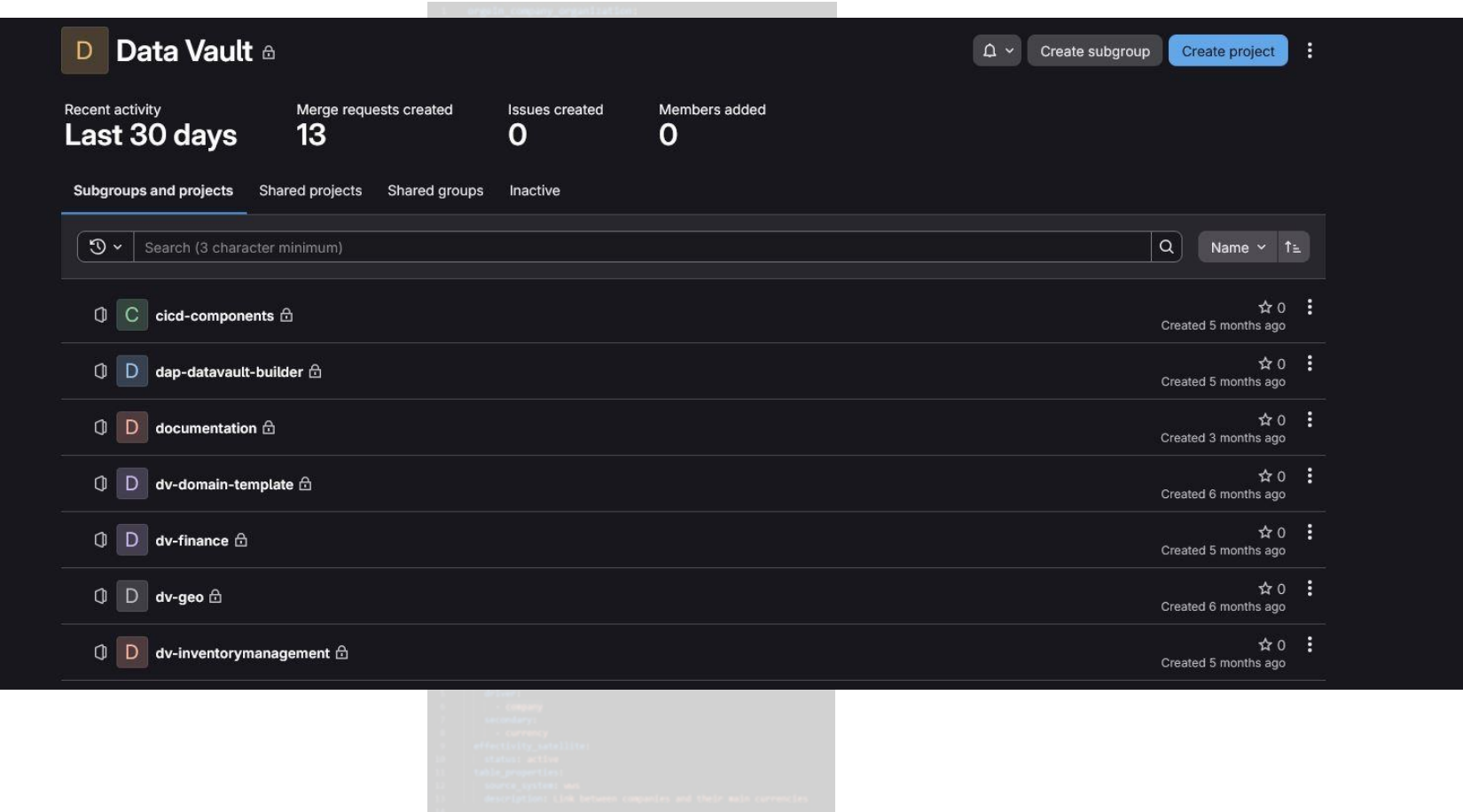
Table with 5 columns: Columns, Type, Comment, Size, Columns tracking. It lists various data sources and their properties.

Columns	Type	Comment	Size	Columns tracking
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			
all_vault_dev	string			



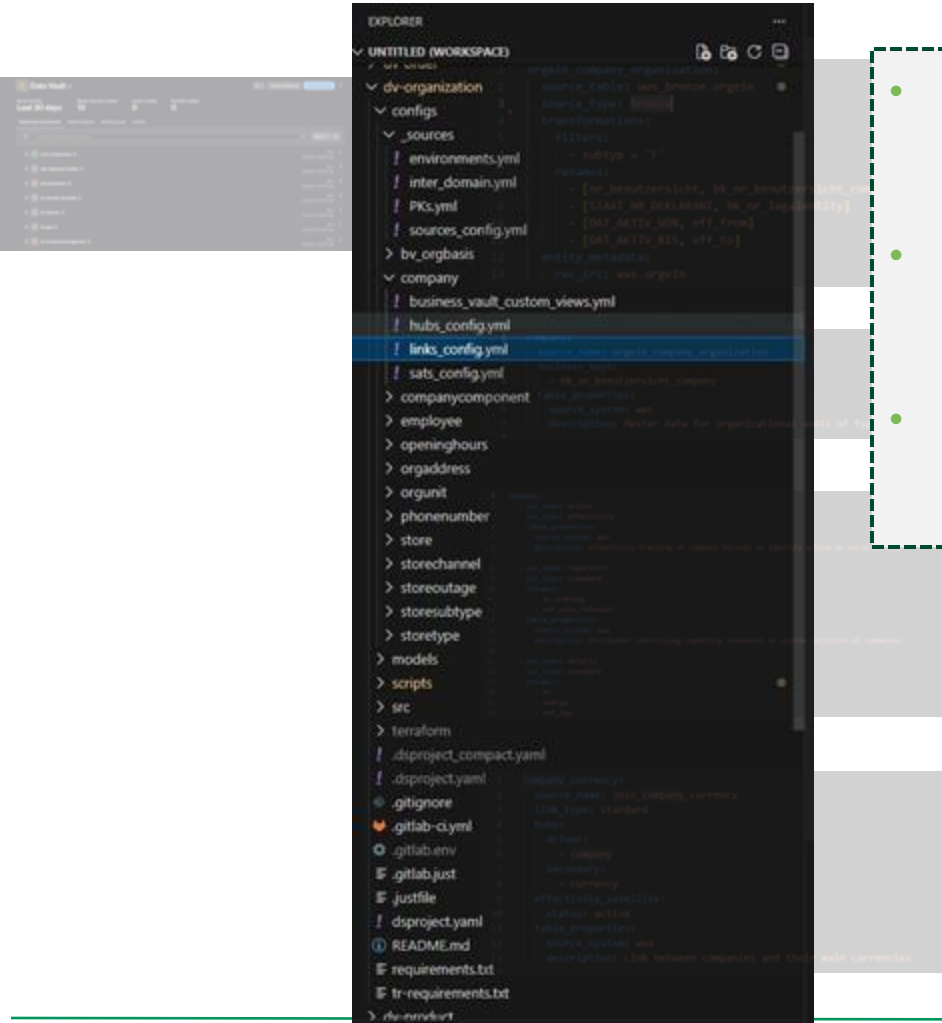
Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.



Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.



- Config Files für Tabellen
- Projekt Configs für einfaches Deployment
- Mermaid-Models




Lufthansa Industry Solutions


A screenshot of a table showing data for a DLT Pipeline. The table has columns for 'id', 'name', 'status', and 'created_at'. The data is organized into rows, with some rows highlighted in blue. The table is titled 'DLT Pipeline' and 'Data quality tests'.


Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.




```
1 orgein_company_organization:
2   source_table: web_branche_orgein
3   source_type: standard
4   transformations:
5     filters:
6       - subtype = "E"
7     renames:
8       - [nr_benutzersicht, bk_nr_benutzersicht_company]
9       - [START NR DEKLARANT, bk_nr_legalentity]
10      - [DAT AKTIV_VON, eff_from]
11      - [DAT AKTIV_BIS, eff_to]
12   entity_metadata:
13     ref_src: web_orgein
```





```
1 company:
2   source_name: orgein_company_organization
3   business_keys:
4     - bk_nr_benutzersicht_company
5   table_properties:
6     source_system: WWS
7     description: Master data for organizational units of type company
8
```

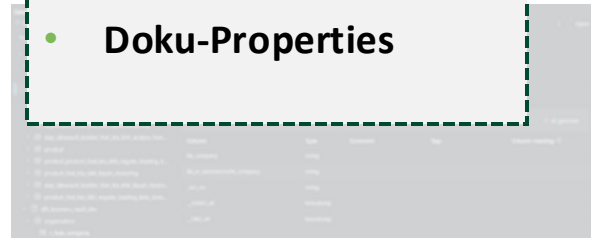


```
1 company_currency:
2   source_name: join_company_currency
3   link_type: standard
4   links:
5     - driver:
6       - company
7     - secondary:
8       - currency
9   effectivity_satellite:
10    status: active
11   table_properties:
12     source_system: WWS
13     description: Link between companies and their web currencies
```



Hub Config

- **Quelldefinition**
- **BKs**
- **Doku-Properties**





Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.

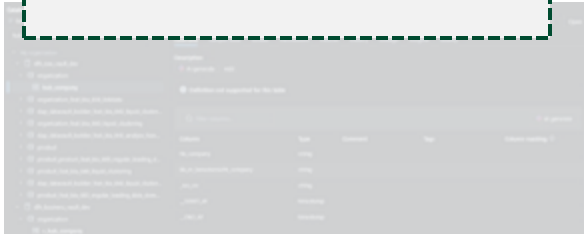


```
1 orgein_company_organization:
2   source_table: wws_bronze.orgein
3   source_type: bronze
4   transformations:
5     filters:
6       - subtyp = 'F'
7     renames:
8       - [nr_benutzersicht, bk_nr_benutzersicht_company]
9       - [STAAT_NR_DEKLARANT, bk_nr_legalentity]
10      - [DAT_AKTIV_VON, eff_from]
11      - [DAT_AKTIV_BIS, eff_to]
12   entity_metadata:
13     rec_src: wws.orgein
```

```
1 company_currency:
2   source_table: join_company_currency
3   link_type: standard
4   links:
5     - primary:
6       - company
7     - secondary:
8       - currency
9   effectivity: satellite
10   status: active
11   table_properties:
12     source_system: wws
13   description: Link between companies and their main currencies
```

Source Config

- Einfache Transformationen in YAML
- Komplexere Definitionen als SQL
- Wiederverwendbar



Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datastrukturen als Databricks Lakeflow
Declarativ

```
1 company:
2   - sat_name: active
3     sat_type: effectivity
4     table_properties:
5       source_system: wws
6       description: Effectivity tracking of company records to identify active or valid periods
7
8   - sat_name: reportctrl
9     sat_type: standard
10    columns:
11      - nr_ordnung
12      - ref_cois_relevant
13    table_properties:
14      source_system: wws
15      description: Attributes controlling reporting relevance or system inclusion of companies
16
17   - sat_name: details
18     sat_type: standard
19     columns:
20       - nr
21       - subtyp
22       - ref_typ
```

Sat Config

- Spalten können angegeben werden
- BKs des Hub müssen in Quelle sein
- Abweichende Quellen möglich

```
1 company_currency:
2   source_name: join_company_currency
3   link_type: standard
4   hubs:
5     driver:
6       - company
7     secondary:
8       - currency
9   effectivity_satellite:
10    status: active
11   table_properties:
12     source_system: wws
13     description: Link between companies and their main currencies
14
```

Link Config

- BKs aller Hubs müssen in Quelle sein

Data Vault Toolkit

Entwicklung eines Frameworks für
Declarative Pipelines.

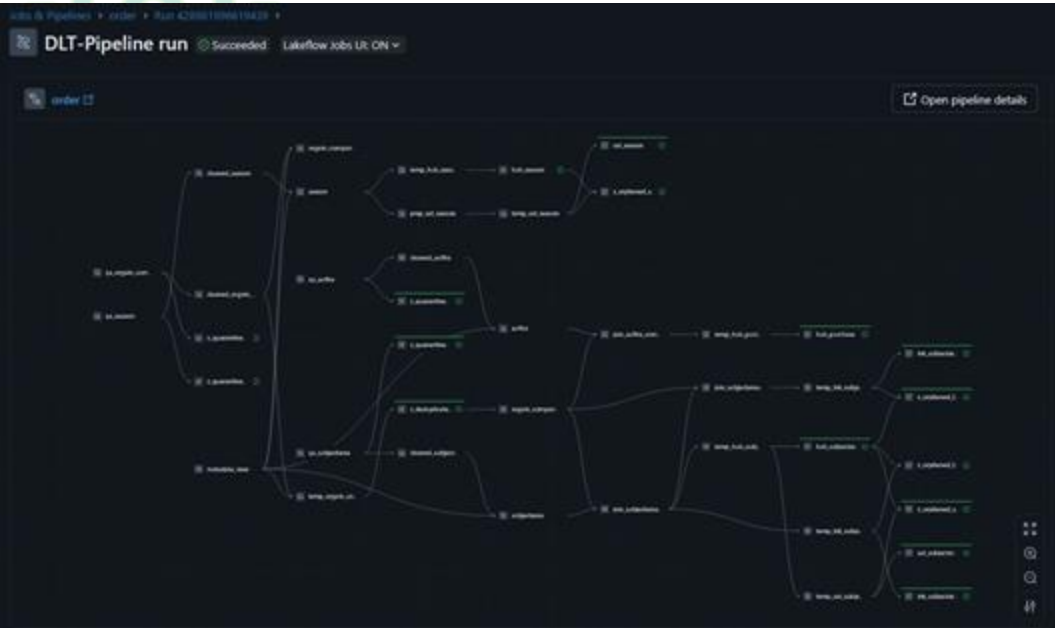
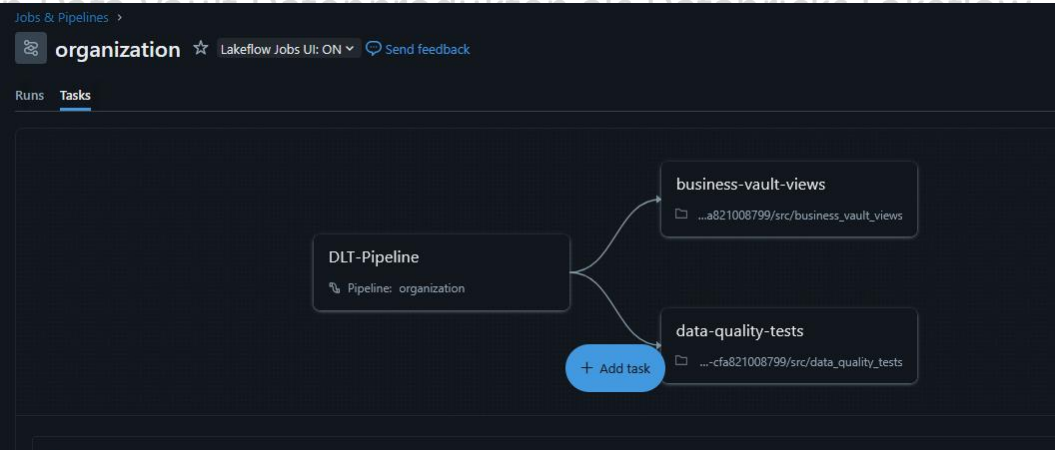


Databricks Job

- DLT-Pipeline
- Business Vault View-Aufbau bei Änderungen
- Data Quality Checks

DLT-Pipeline

- Abhängigkeiten automatisch
- Automatische Increment-Erkennung

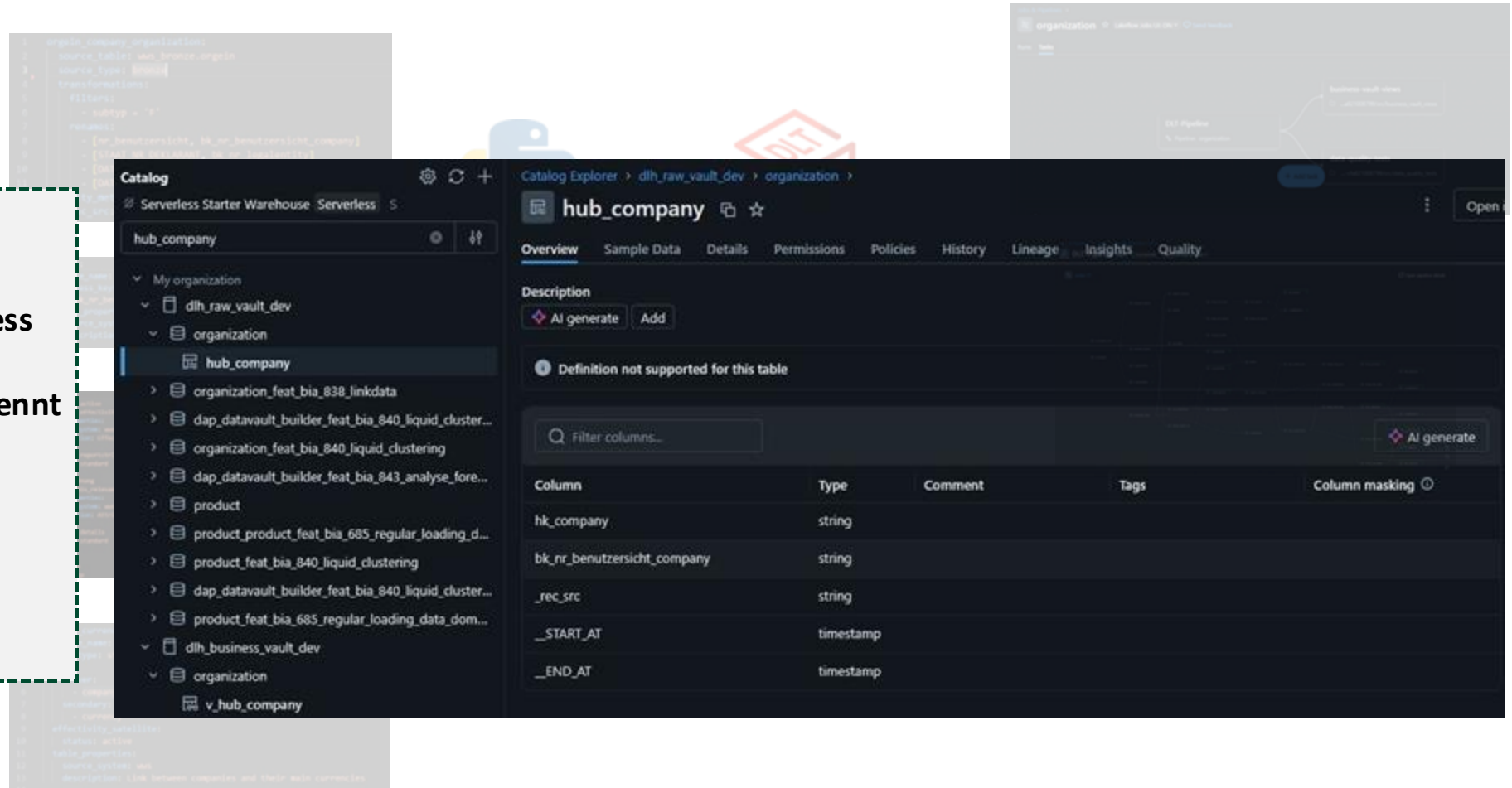


Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.

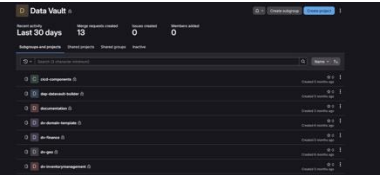
Unity Catalog

- Raw Vault / Business Vault auf Katalogebene getrennt
- Domäne / Datenprodukt als Schema



Data Vault Toolkit

Entwicklung eines Frameworks zur Implementierung von Data Vault Datenprodukten als Databricks Lakeflow Declarative Pipelines.

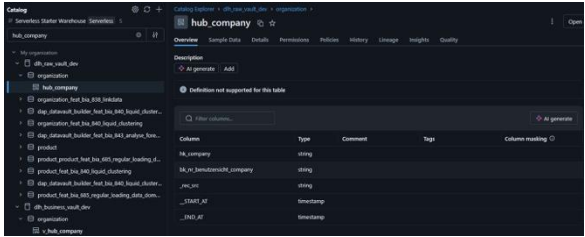
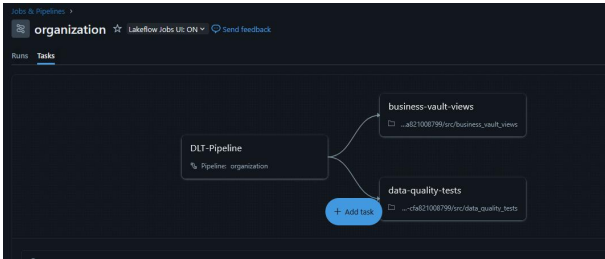


```
1 orgein_company_organization:
2   source_table: wms_bronze.orgein
3   source_type: bronze
4   transformations:
5     filters:
6       - subtyp = 'F'
7     renames:
8       - [nr_benutzersicht, bk_nr_benutzersicht_company]
9       - [STAAT_NR_DEKLARANT, bk_nr_legalentity]
10      - [DAT_AKTIV_VON, eff_from]
11      - [DAT_AKTIV_BIS, eff_to]
12   entity_metadata:
13     rec_src: wms.orgein
```

```
1 company:
2   source_name: orgein_company_organization
3   business_keys:
4     - bk_nr_benutzersicht_company
5   table_properties:
6     source_system: wms
7   description: Master data for organizational units of type company
```

```
1 company:
2   - sat_name: active
3     sat_type: effectivity
4     table_properties:
5       source_system: wms
6       description: Effectivity tracking of company records to identify active or valid periods
7   - sat_name: reportctrl
8     sat_type: standard
9     columns:
10       - nr_orderung
11       - ref_cds_relevant
12     table_properties:
13       source_system: wms
14       description: Attributes controlling reporting relevance or system inclusion of companies
15   - sat_name: details
16     sat_type: standard
17     columns:
18       - nr
19       - subtyp
20       - ref_tip
```

```
1 company_currency:
2   source_name: join_company_currency
3   link_type: standard
4   hubs:
5     driver:
6       - company
7     secondary:
8       - currency
9   effectivity_satellite:
10     status: active
11   table_properties:
12     source_system: wms
13   description: Link between companies and their main currencies
```



Columns	Type	Comment	Tags	Columns tracking
id	string			
organization	string			
product	string			
status	string			



Challenges DWH Prototyp



Keys

Legacy DWH stark auf Surrogate Keys ausgerichtet.
Business Keys müssen identifiziert werden.



Abgrenzung

Wie schneiden wir die Datendomänen?
Was gehört noch in den Business Vault, was schon in den Serving Layer?



Stammdaten

Keine abgeschlossenen Datenprodukte, Nutzung über Domänengrenzen hinweg



Technische Restriktionen

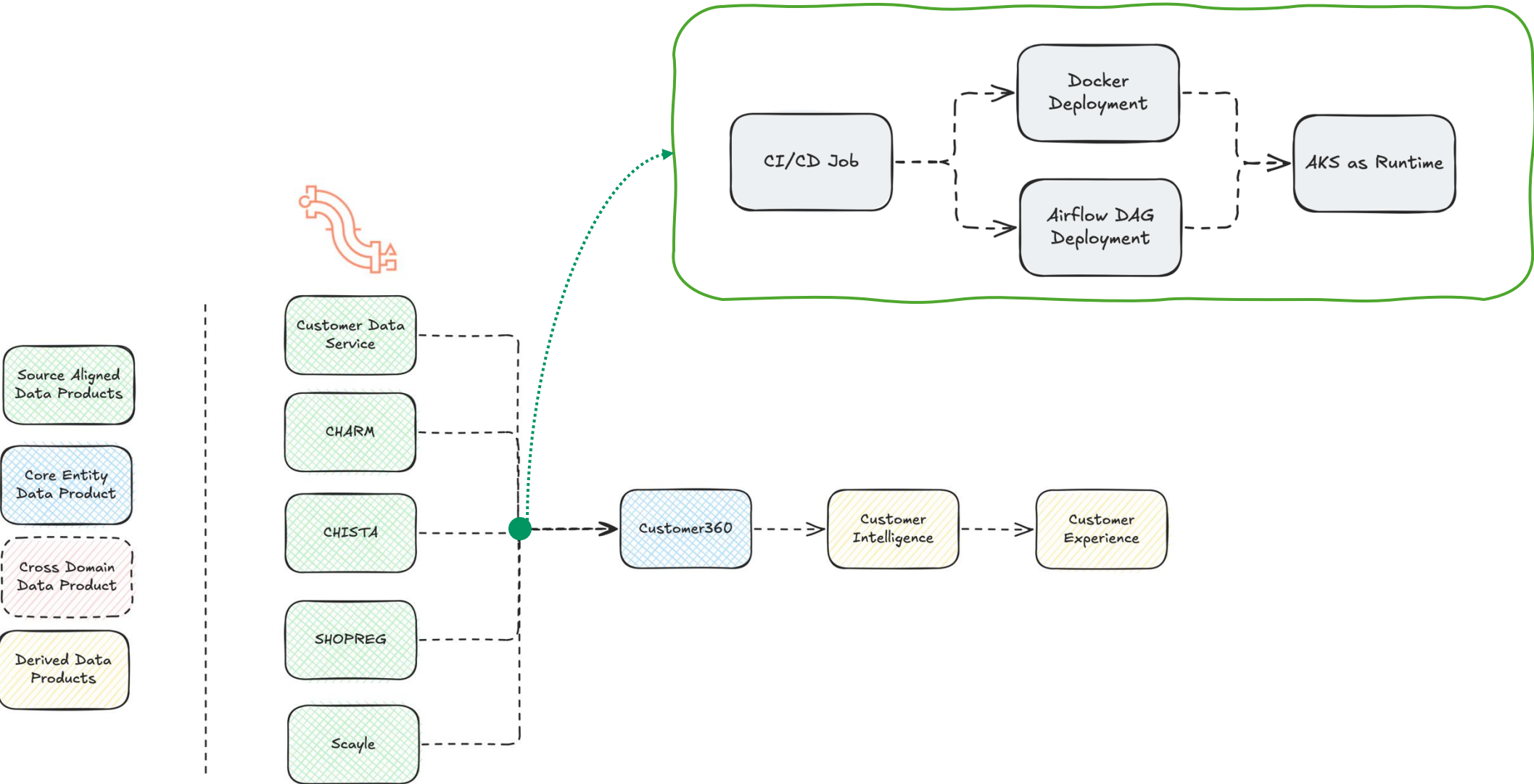
Nicht alles in DLT umsetzbar. Es mussten Workarounds gefunden und kleinere Abweichungen vom Data Vault Standard in Kauf genommen werden.



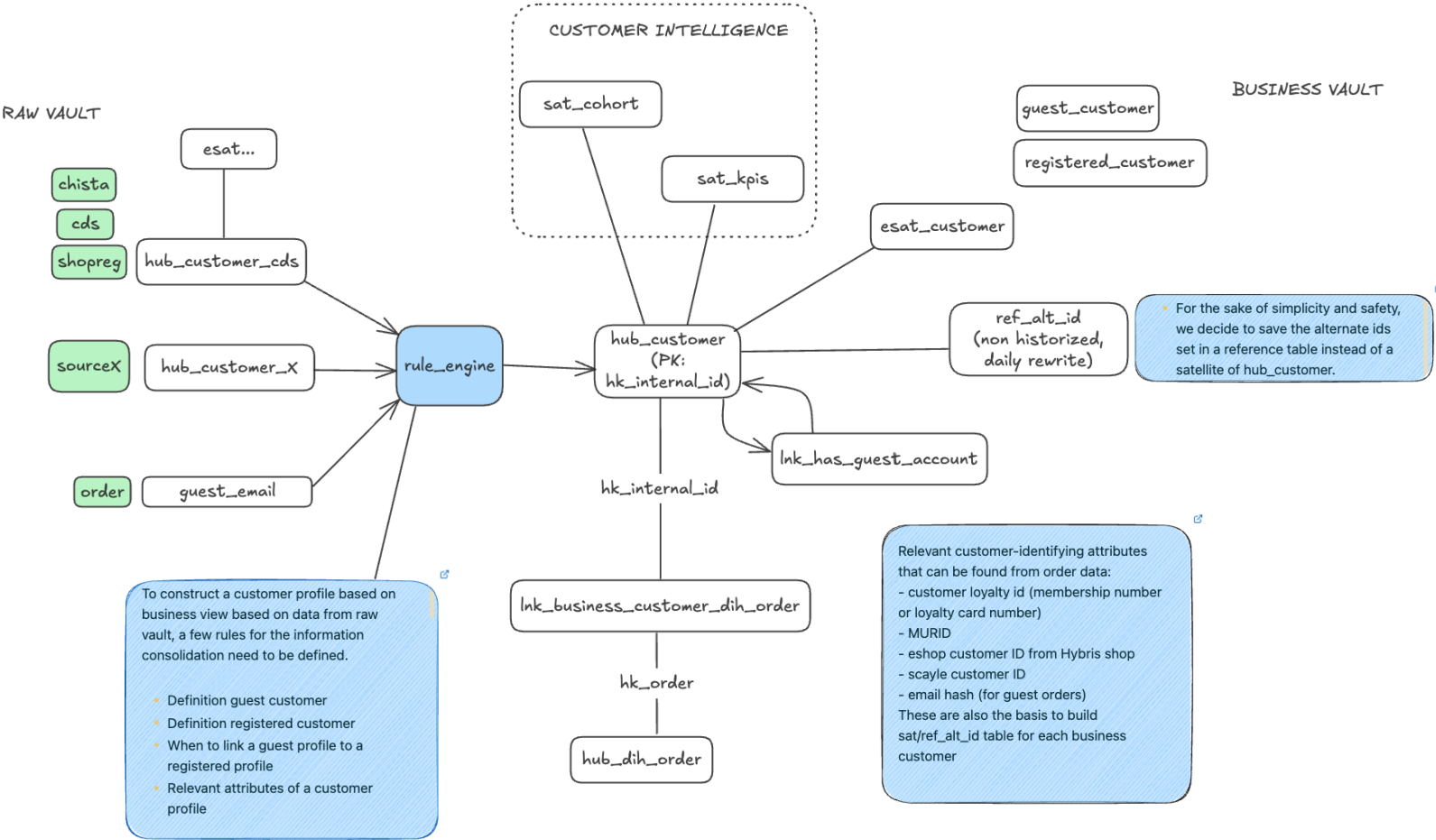
03.

Use Case “Customer Analytics”

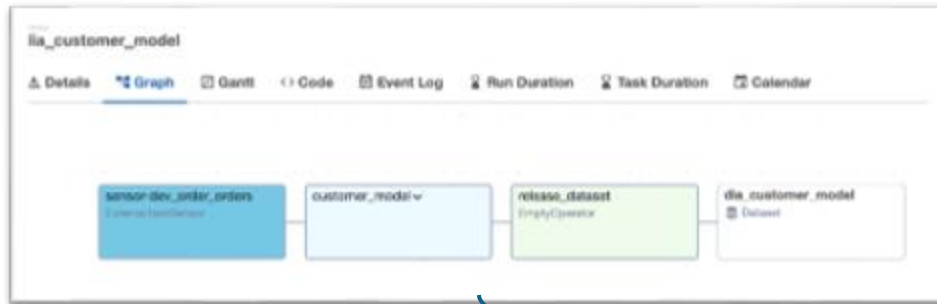




Auszug aus dem Customer Data Model



datavault4dbt & Astronomer Cosmos im DataHub Apache Airflow



Airflow DAG führt dbt Projekt aus und liefert das Data Asset “dia_customer_model” zur weiteren Orchestrierung aus.



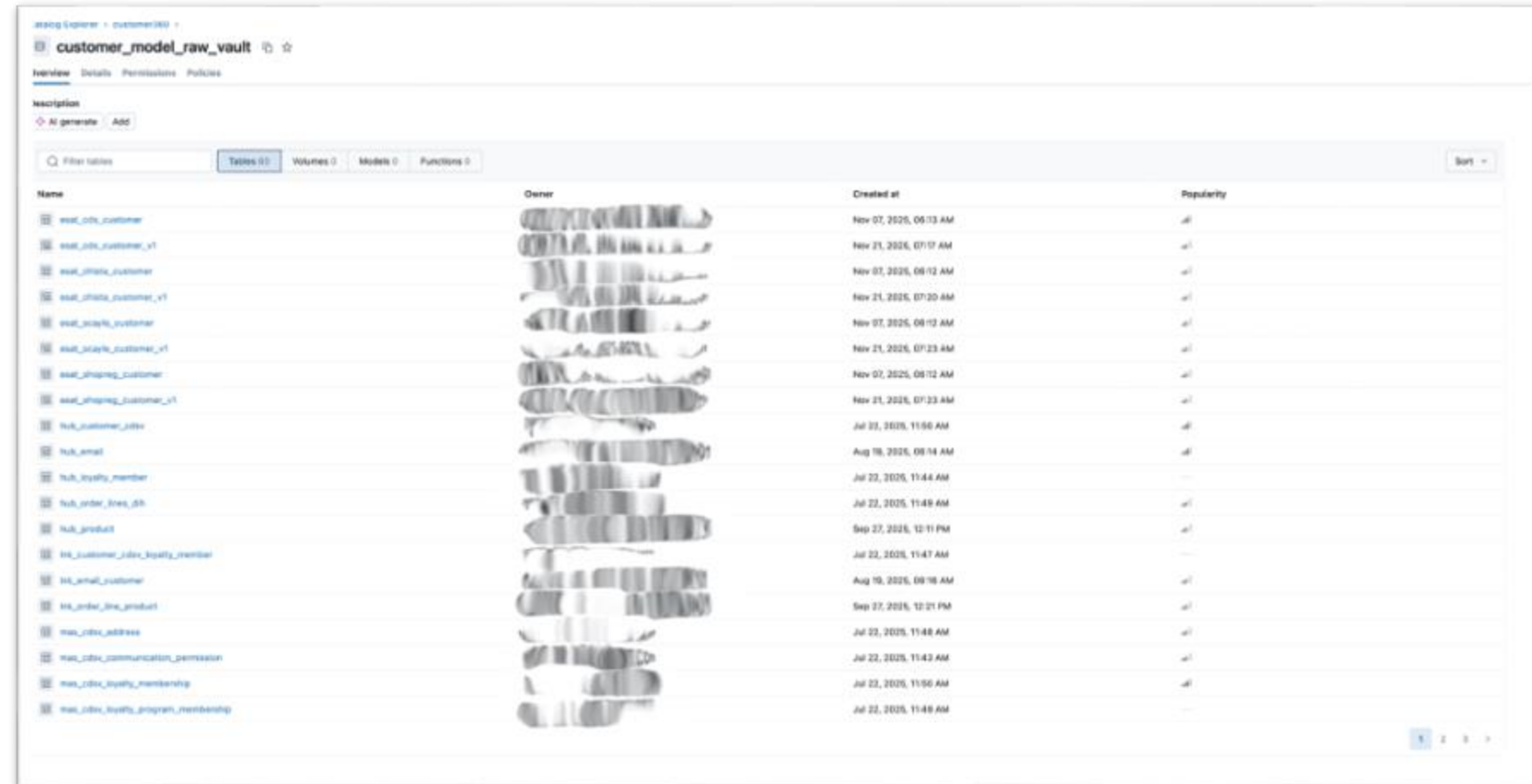
Expandiert: Die Lineage aus dbt im DAG via Cosmos. Ausgeführt mittels deferred Operators auf unserem AKS.

Das (Teil-)Ergebnis – es lebt!

Auszug aus der Raw Vault – der Weg zum goldenen Record in der Business Vault

Disclaimer: **GDPR**

- ✓ Separater Storage
- ✓ Zugriff nur auditiert
- ✓ Data Masking
- ✓ Löschkonzept implementiert
- ✓ Trennung auf PII-Satelliten



The screenshot shows the Snowflake web interface for a schema named 'customer_model_raw_vault'. It displays a list of tables, all of which have their data masked (blurred). The tables are organized into categories like 'raw_customer', 'raw_order_line', and 'raw_product'. The interface includes filters for Tables (20), Volumes (0), Models (0), and Functions (0). A 'Sort' button is visible on the right.

Name	Owner	Created at	Popularity
raw_customer		Nov 07, 2025, 06:13 AM	all
raw_customer_v1		Nov 21, 2025, 07:17 AM	all
raw_order_line		Nov 07, 2025, 06:12 AM	all
raw_order_line_v1		Nov 21, 2025, 07:20 AM	all
raw_product		Nov 07, 2025, 06:12 AM	all
raw_product_v1		Nov 21, 2025, 07:23 AM	all
raw_customer_order_line		Nov 07, 2025, 06:12 AM	all
raw_customer_order_line_v1		Nov 21, 2025, 07:23 AM	all
raw_customer_order_line		Jul 22, 2025, 11:56 AM	all
raw_email		Aug 18, 2025, 08:14 AM	all
raw_order_line		Jul 22, 2025, 11:44 AM	all
raw_order_line_v1		Jul 22, 2025, 11:49 AM	all
raw_product		Sep 27, 2025, 12:11 PM	all
raw_customer_order_line		Jul 22, 2025, 11:47 AM	all
raw_email		Aug 18, 2025, 08:18 AM	all
raw_order_line		Sep 27, 2025, 12:21 PM	all
raw_order_line_v1		Jul 22, 2025, 11:48 AM	all
raw_order_line_v2		Jul 22, 2025, 11:43 AM	all
raw_order_line_v3		Jul 22, 2025, 11:50 AM	all
raw_order_line_v4		Jul 22, 2025, 11:49 AM	all



04.

Was haben wir gelernt?



Lessons Learned



Gemeinsame BK Definitionen sind der Schlüssel zum Erfolg

1

2

Automation als Skalierungsfaktor



Einheitliche Definitionen und Styleguides (trotz unterschiedlicher Ansätze)

3

4

Test-Strategie frühzeitig überlegen



SCRUM als Treiber

5

von Kopf bis Schuh

Überall wo's Podcast gibt.



So erreichst du uns:



Sebastian Ennigkeit

sebastian_ennigkeit@deichmann.com



Kevin Haferkamp

kevin_haferkamp@deichmann.com



Wir suchen dich!

WERKSTUDENT (m/w/d)

PRAKTIKANT (m/w/d)

PROFESSIONAL (m/w/d)

Bewirb dich jetzt!



deichmann-karriere.de

D