



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

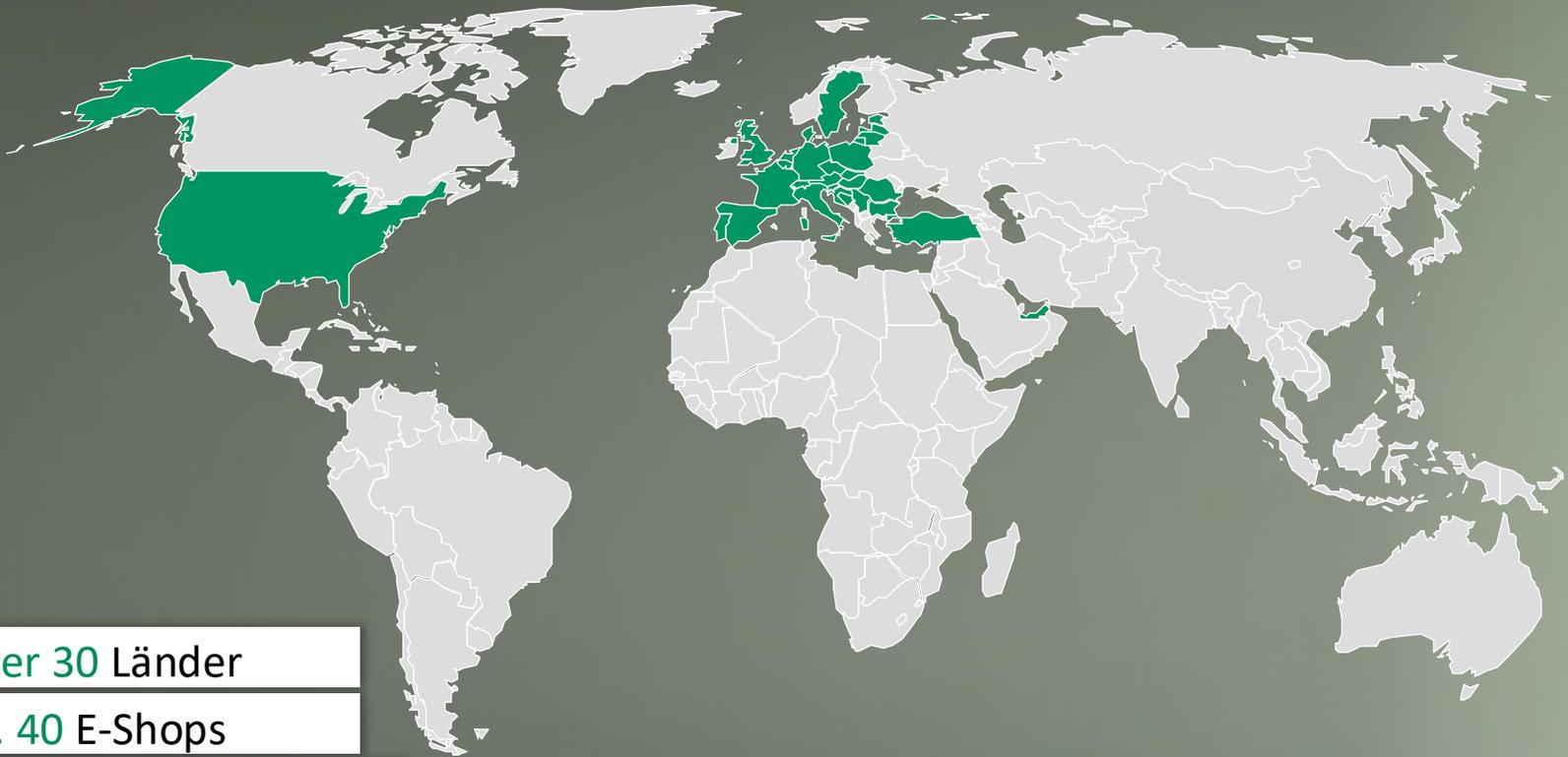
- DV2 Practitioner
- Agile Enthusiast
- Hobbies: Music, Travel, Basketball



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®

van Haren

solebox.

OCHSNER
SHOES

supremo

OCHSNER
SPORT

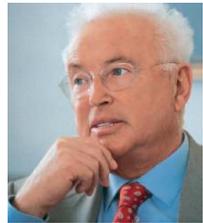
urbanstyles.

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

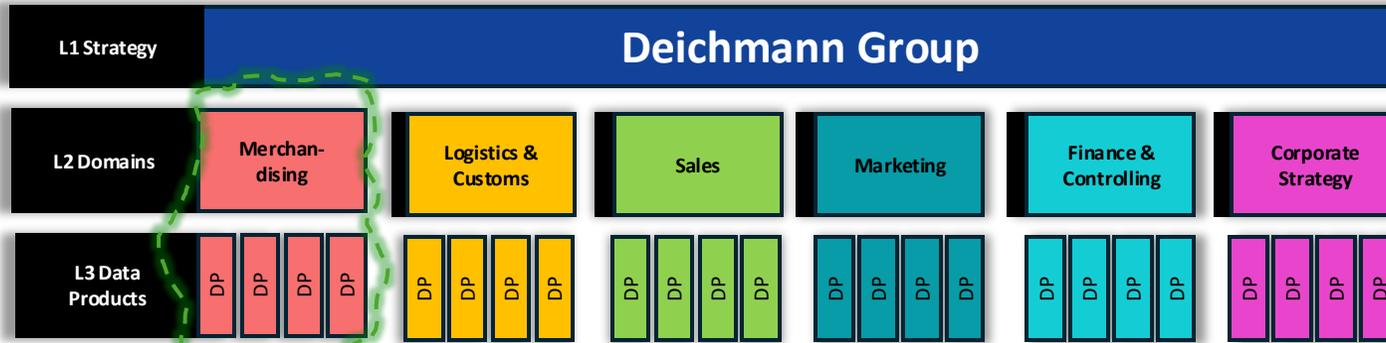


Deichmanns “Road to Data Mesh”

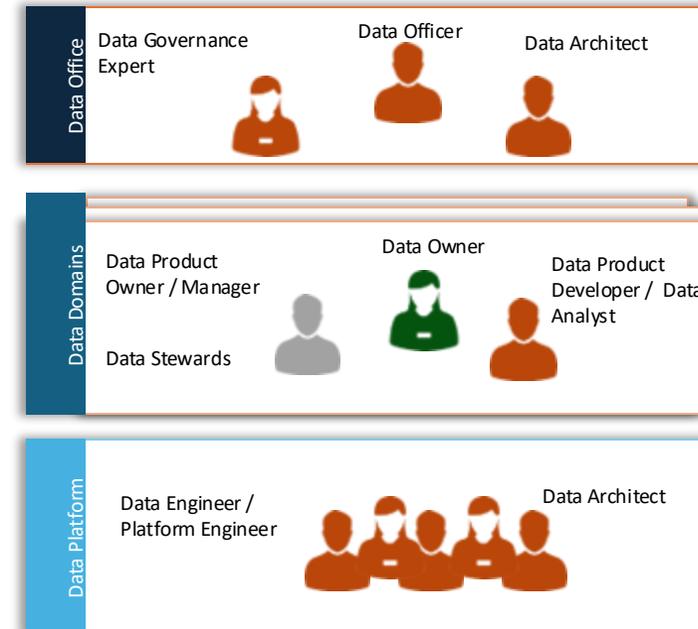


Data Domains* @ Deichmann

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

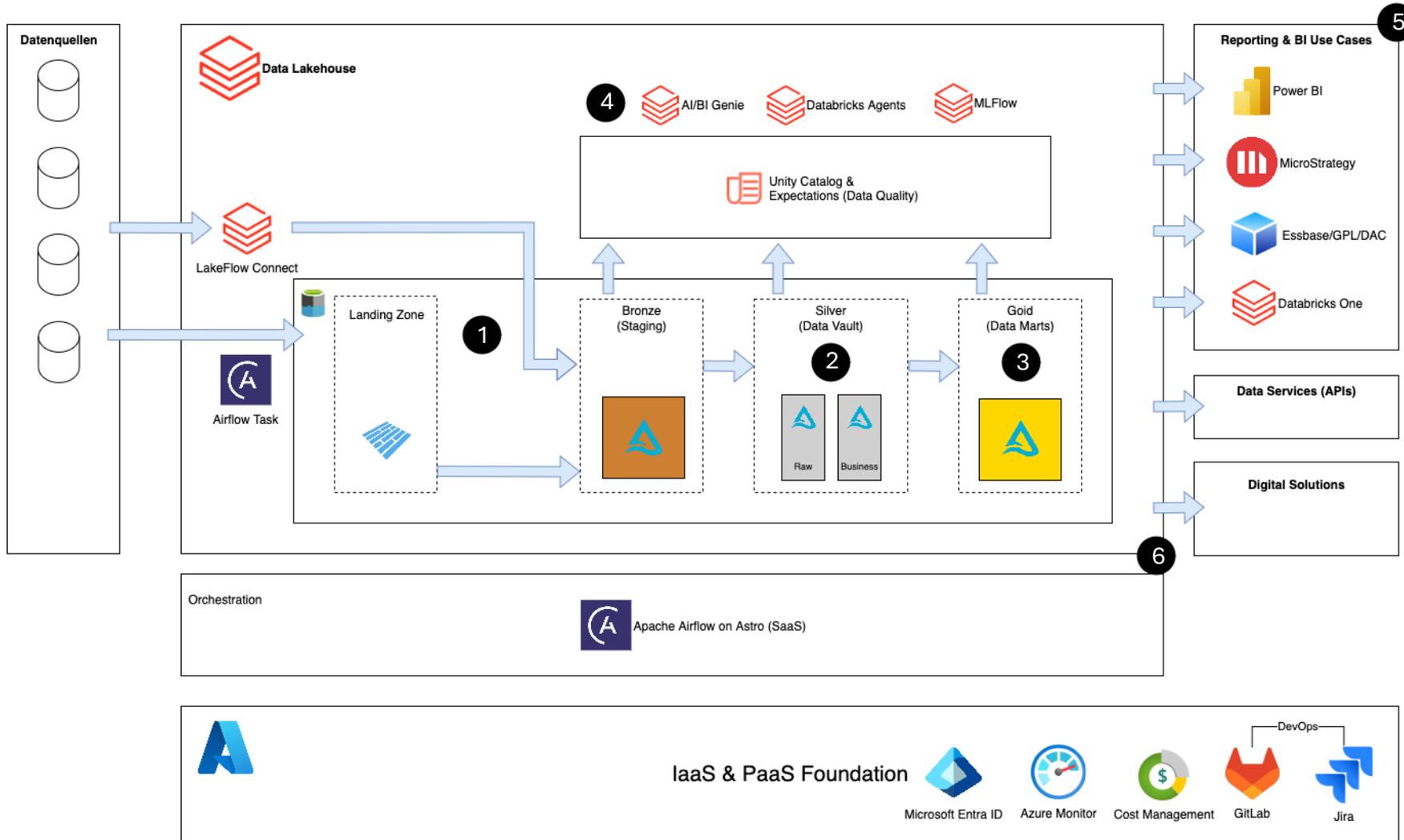


Exemplarische Rollen in einer Domäne



*unvollständig und stark vereinfacht

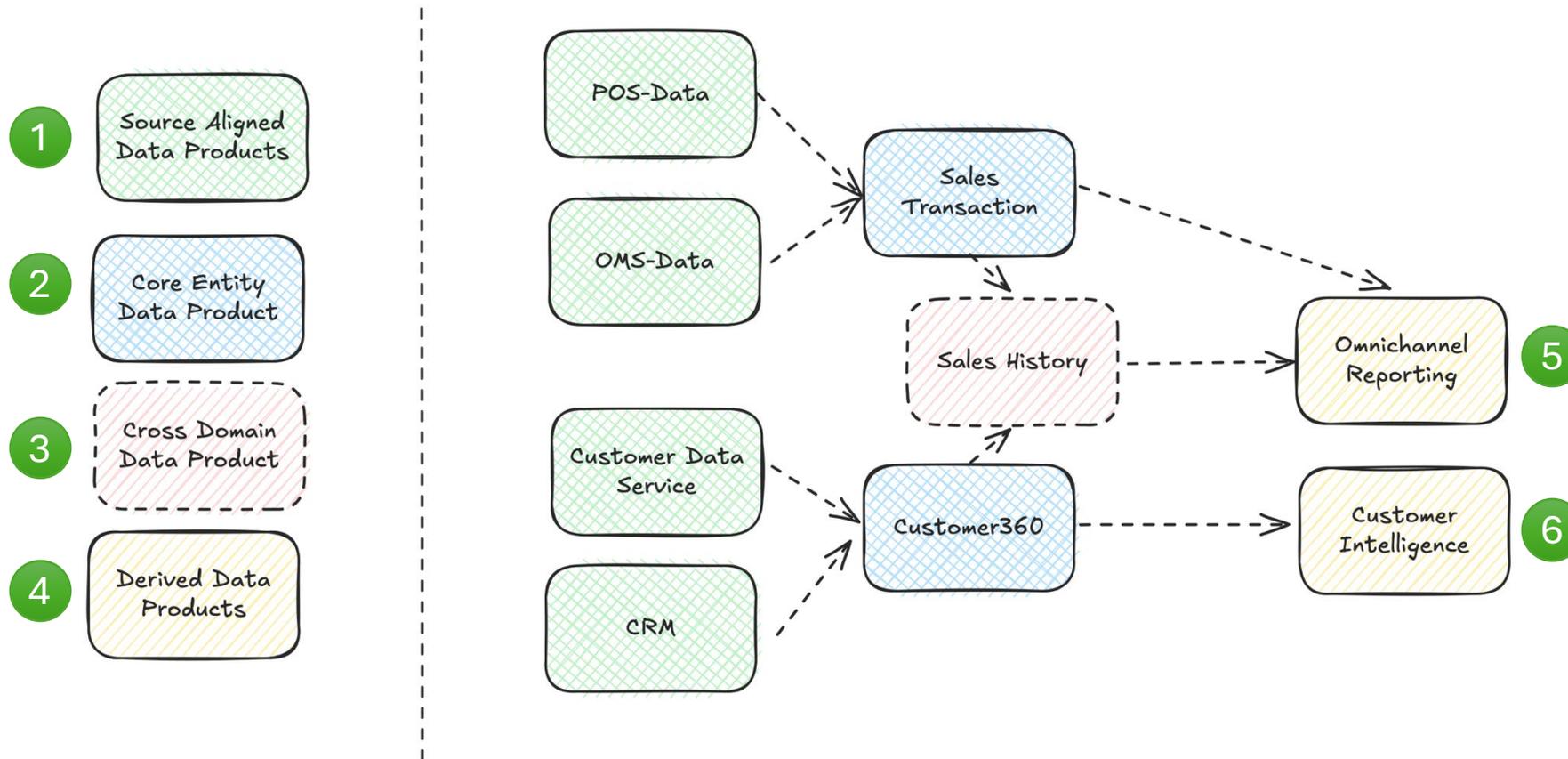
Deichmanns Data Analytics Platform (DAP)



- 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.



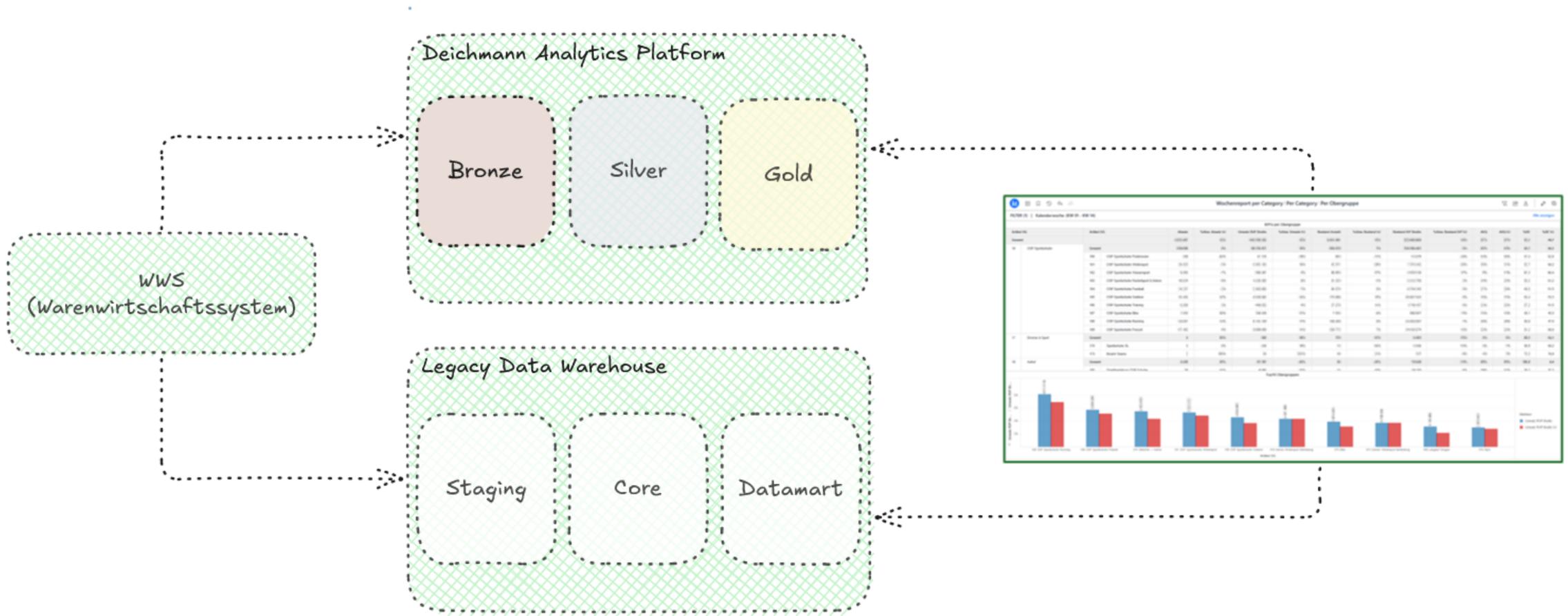
02.

Use Case “DWH Migration”



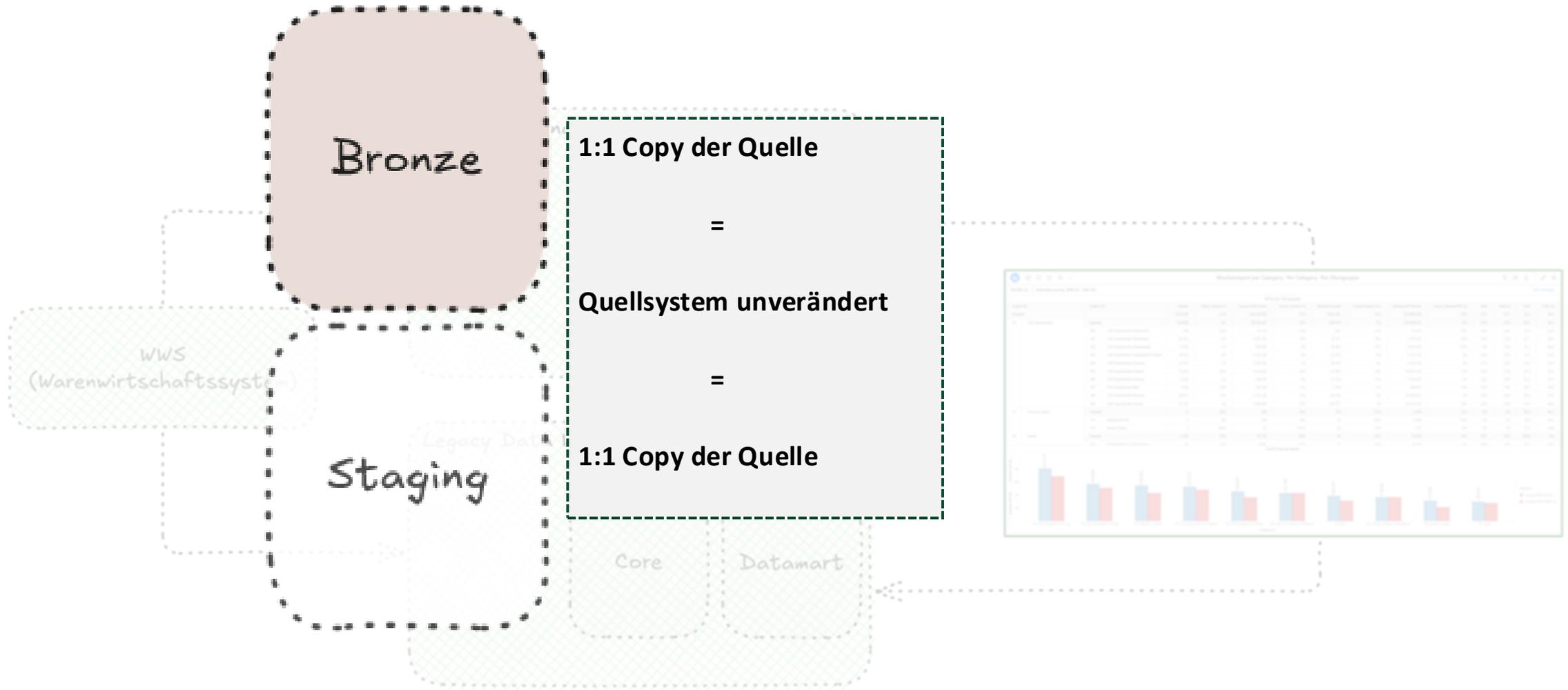
Ausgangslage

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



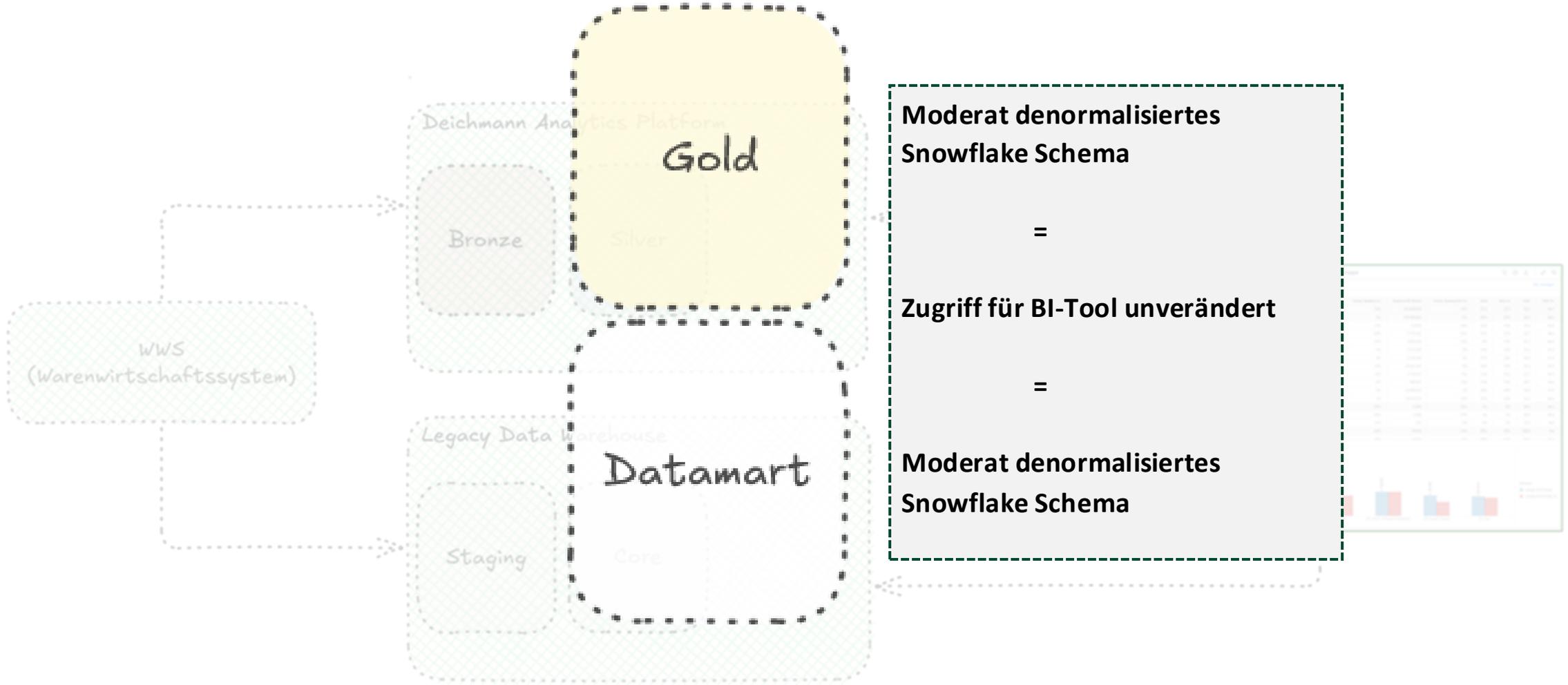
Ausgangslage

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



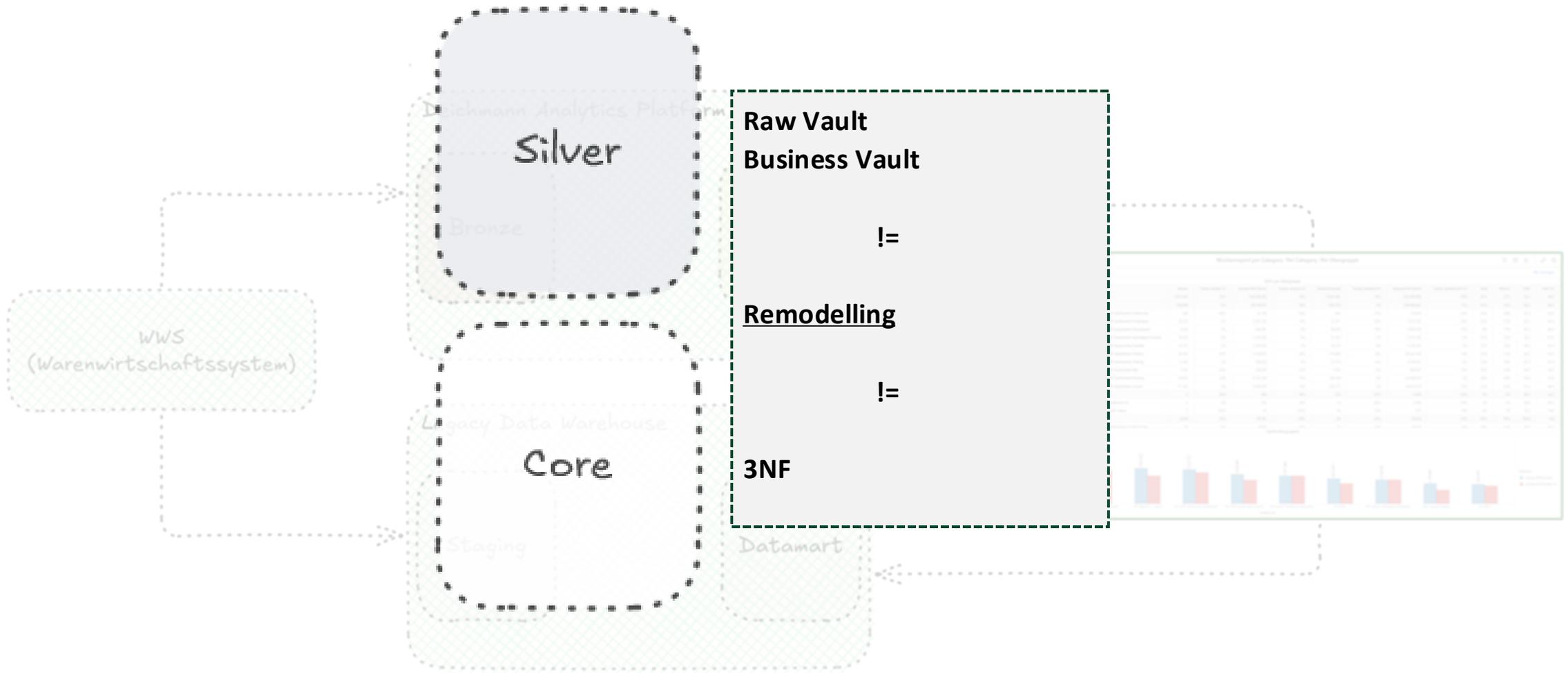
Ausgangslage

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



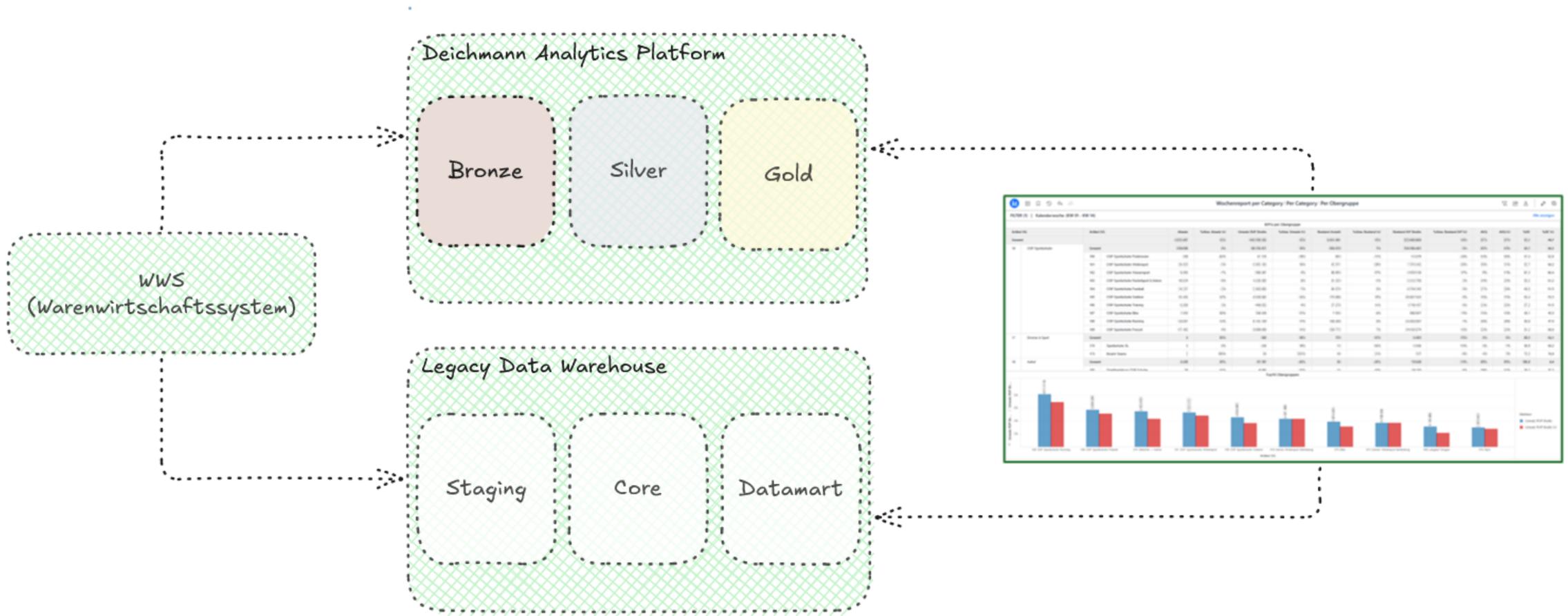
Ausgangslage

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



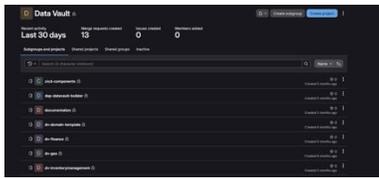
Ausgangslage

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.

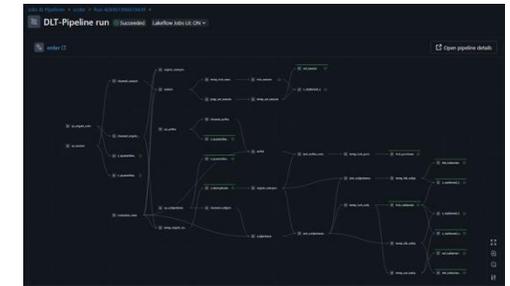
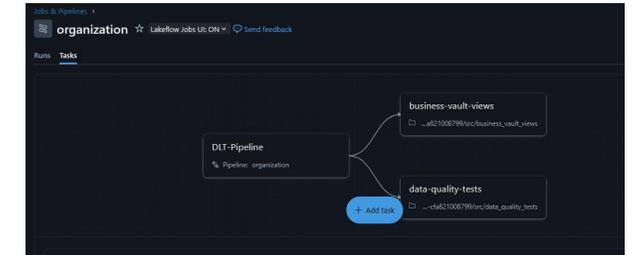
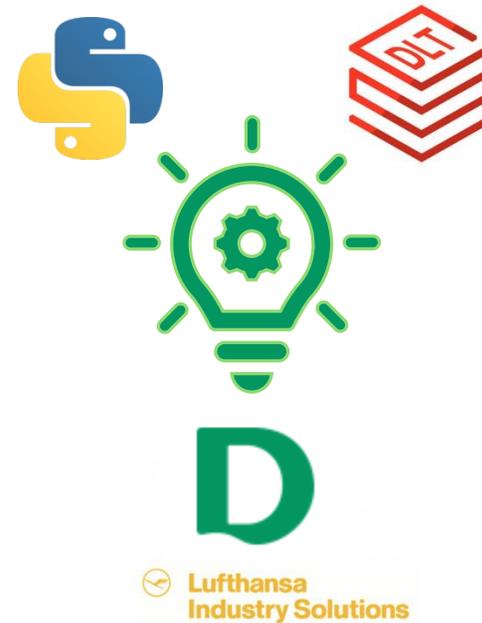
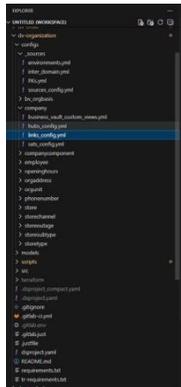


```
1 orgein_company_organization:  
2 source_table: wms_bronze.orgein  
3 source_type: bronze  
4 transformations:  
5 filters:  
6   - subtyp = 'F'  
7 renames:  
8   - [nr_benutzersicht, bk_nr_benutzersicht_company]  
9   - [STAAT_NR_DEKLARANT, bk_nr_legalentity]  
10  - [DAT_AKTIV_VON, eff_from]  
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  
8   - sat_name: reportctrl  
9   sat_type: standard  
10  columns:  
11    - nr_orbueing  
12    - ref_cols_relevant  
13  table_properties:  
14    source_system: wms  
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
```

```
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  
14
```

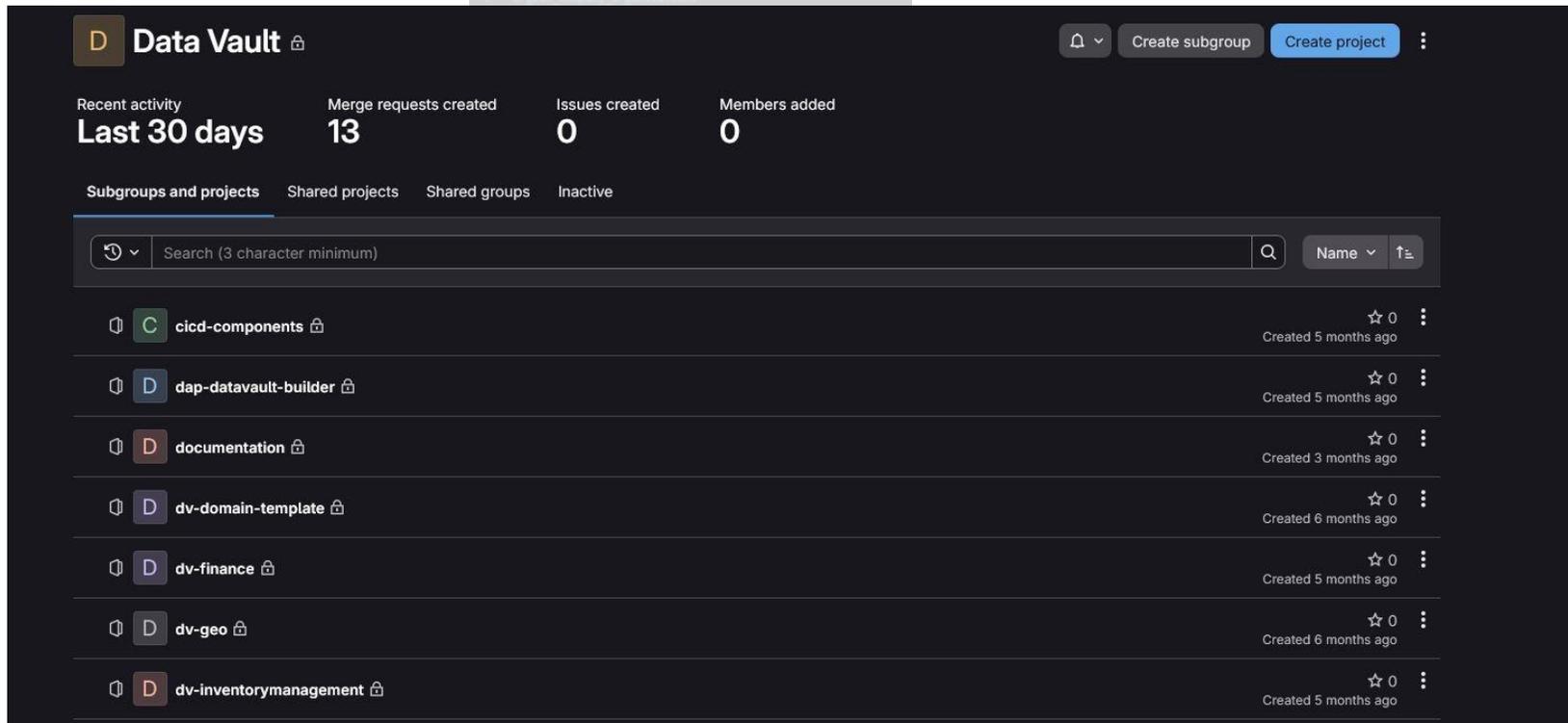


Columns	Type	Comment	Tags	Columns missing
nr	string			
subtyp	string			
ref_typ	string			
nr_orbueing	string			
ref_cols_relevant	string			
nr_benutzersicht_company	string			
bk_nr_legalentity	string			
eff_from	timestamp			
eff_to	timestamp			



Data Vault Toolkit

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



```
1 | company  
2 | secondary  
3 | currency  
4 | effective_to_date  
5 | status: active  
6 | table_properties  
7 | source_system: web  
8 | description: link between companies and their web currencies
```

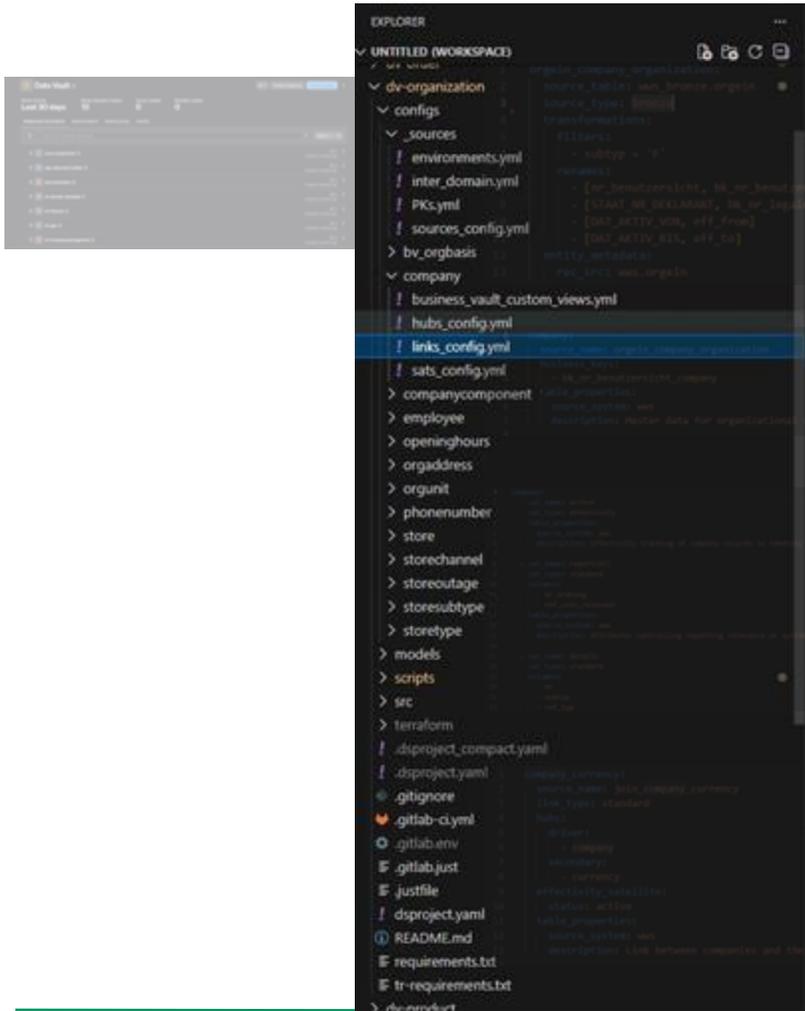
The diagram illustrates the architecture of the Data Vault Toolkit. It shows a 'Framework' box on the left, a 'Data Definition' box in the middle, and a 'Data Product' box on the right. Arrows indicate the flow of data from the framework to the data definition, and from the data definition to the data product. A dashed green box highlights the 'Framework' and 'Data Definition' components, with a list of key features:

- Framework getrennt von Datendefinition
- Gitlab Repo je Datenprodukt
- Template Repo als Vorlage



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**

A screenshot of a table showing data for 'hubs_company'. The table has several columns and rows of data. The columns include 'id', 'name', 'description', 'status', 'created_at', and 'updated_at'. The rows contain various data points, including '1', '2', '3', '4', '5', '6', '7', '8', '9', '10', '11', '12', '13', '14', '15', '16', '17', '18', '19', '20', '21', '22', '23', '24', '25', '26', '27', '28', '29', '30', '31', '32', '33', '34', '35', '36', '37', '38', '39', '40', '41', '42', '43', '44', '45', '46', '47', '48', '49', '50', '51', '52', '53', '54', '55', '56', '57', '58', '59', '60', '61', '62', '63', '64', '65', '66', '67', '68', '69', '70', '71', '72', '73', '74', '75', '76', '77', '78', '79', '80', '81', '82', '83', '84', '85', '86', '87', '88', '89', '90', '91', '92', '93', '94', '95', '96', '97', '98', '99', '100'.

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: [SOURCE]  
4 transformations:  
5 filters:  
6   - subtyp = "X"  
7  
8 remarks:  
9   - [nr_benutzersicht, bk_nr_benutzersicht_company]  
10  - [START_WK_DEKLARANT, bk_nr_legalentity]  
11  - [DAT_AKTIV_WK, eff_from]  
12  - [DAT_AKTIV_BIS, eff_to]  
13  
14 entity_metadata:  
15 ref_sct: wws_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
```

Hub Config

- Quelldefinition
- BKs
- Doku-Properties



```
1 company_currency:  
2 source_table: data_company_currency  
3 link_type: standard  
4 links:  
5   - company  
6     - currency  
7     - effectivity_date/letter  
8   status: active  
9  
10 table_properties:  
11 source_system: wws  
12 description: Link between companies and their base currencies
```

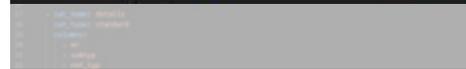


Data Vault Toolkit

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
```

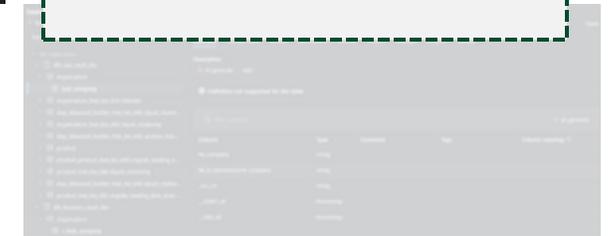


```
1 company_currency:
2   source_table: data_company_currency
3   link_type: standard
4   table:
5     columns:
6       - company
7       - secondary
8       - currency
9     effective_to: date(yyyy)
10    status: active
11    table_properties:
12      source_system: wws
13    description: link between companies and their wws currencies
```



Source Config

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



Entwicklung eines Data Vault Data Modells 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.

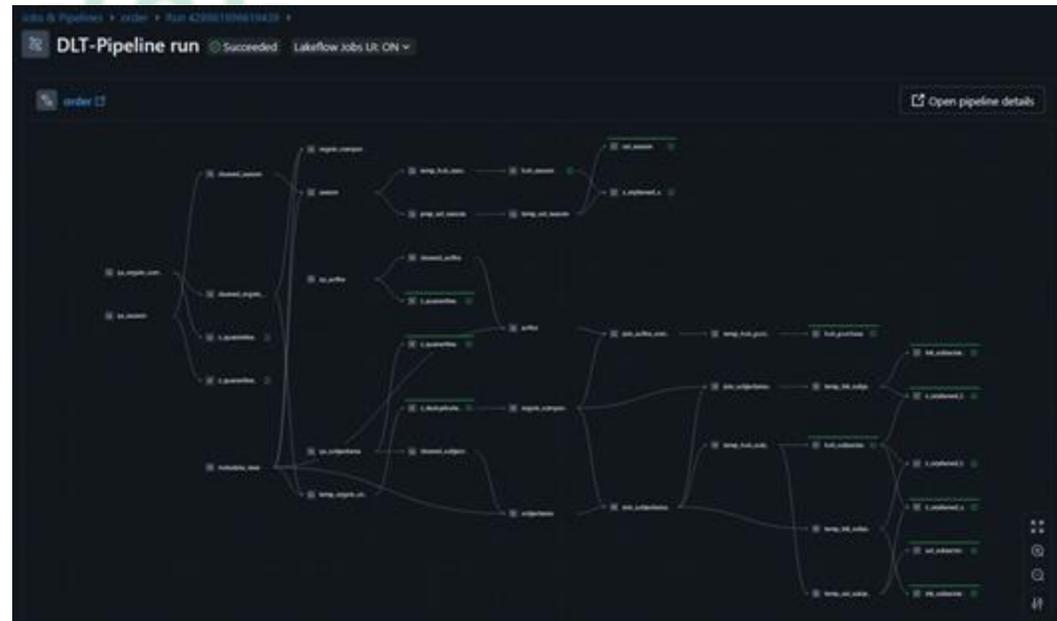
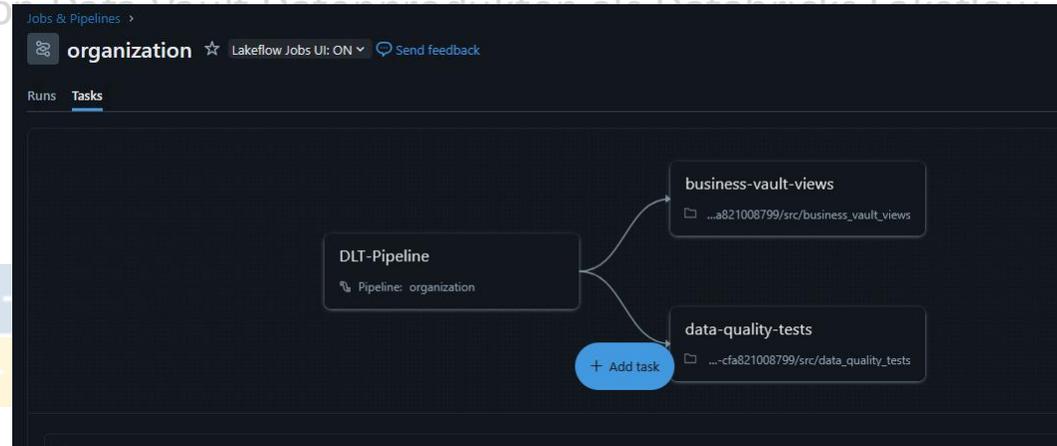


Databricks Job

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

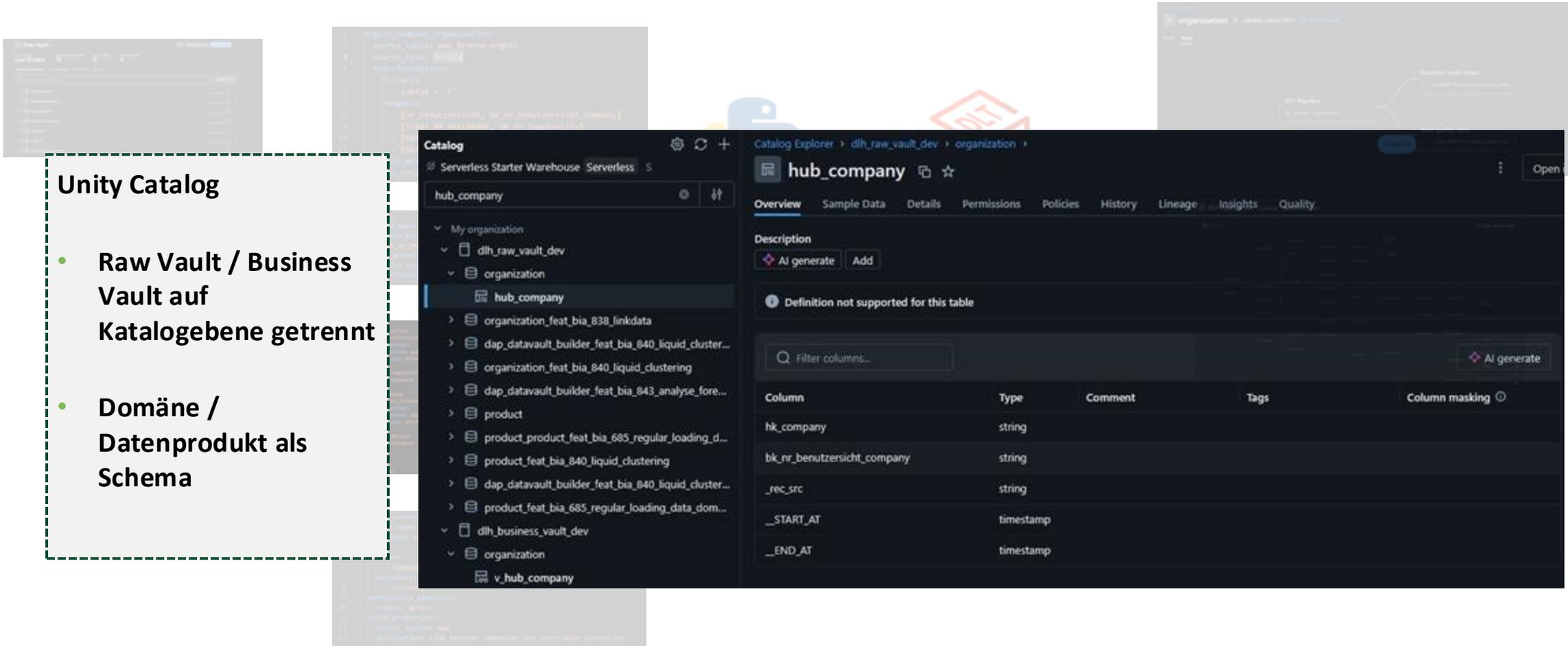
DLT-Pipeline

- Abhängigkeiten automatisch
- Automatische Increment-Erkennung



Data Vault Toolkit

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



The screenshot displays the Unity Catalog interface. On the left, a 'Catalog Explorer' shows a hierarchical view of the 'hub_company' catalog, including sub-catalogs like 'dlh_raw_vault_dev' and 'dlh_business_vault_dev', and various tables and views. On the right, the 'hub_company' table is selected, showing its 'Overview' page. The 'Description' section includes an 'AI generate' button. Below, a table lists the columns and their types:

Column	Type	Comment	Tags	Column masking
hk_company	string			
bk_nr_benutzersicht_company	string			
_rec_src	string			
_START_AT	timestamp			
_END_AT	timestamp			

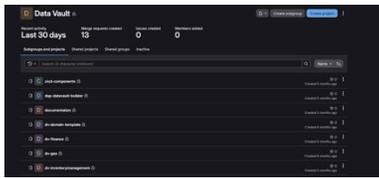
Unity Catalog

- 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.

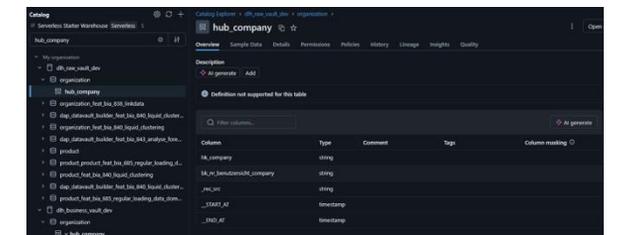
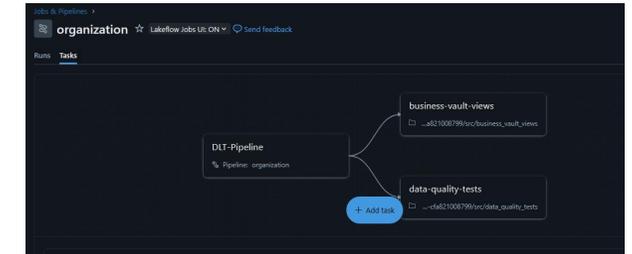
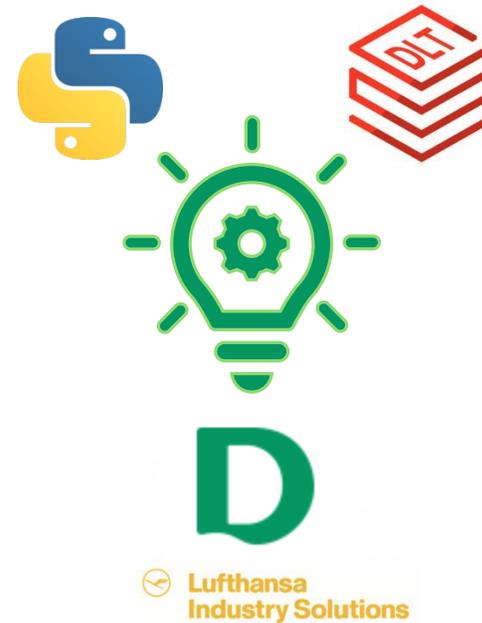
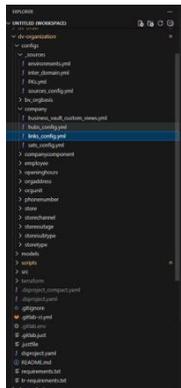


```
1 orgein_company_organization:  
2 source_table: wms_bronze.orgein  
3 source_type: bronze  
4 transformations:  
5 filters:  
6   - subtyp = 'F'  
7 renames:  
8   - [nr_benutzersicht, bk_nr_benutzersicht_company]  
9   - [STAAT_NR_DEKLARANT, bk_nr_legalentity]  
10  - [DAT_AKTIV_VON, eff_from]  
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  
8   - sat_name: reportctrl  
9   sat_type: standard  
10  columns:  
11    - nr_orbung  
12    - ref_cols_relevant  
13 table_properties:  
14   source_system: wms  
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
```

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



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

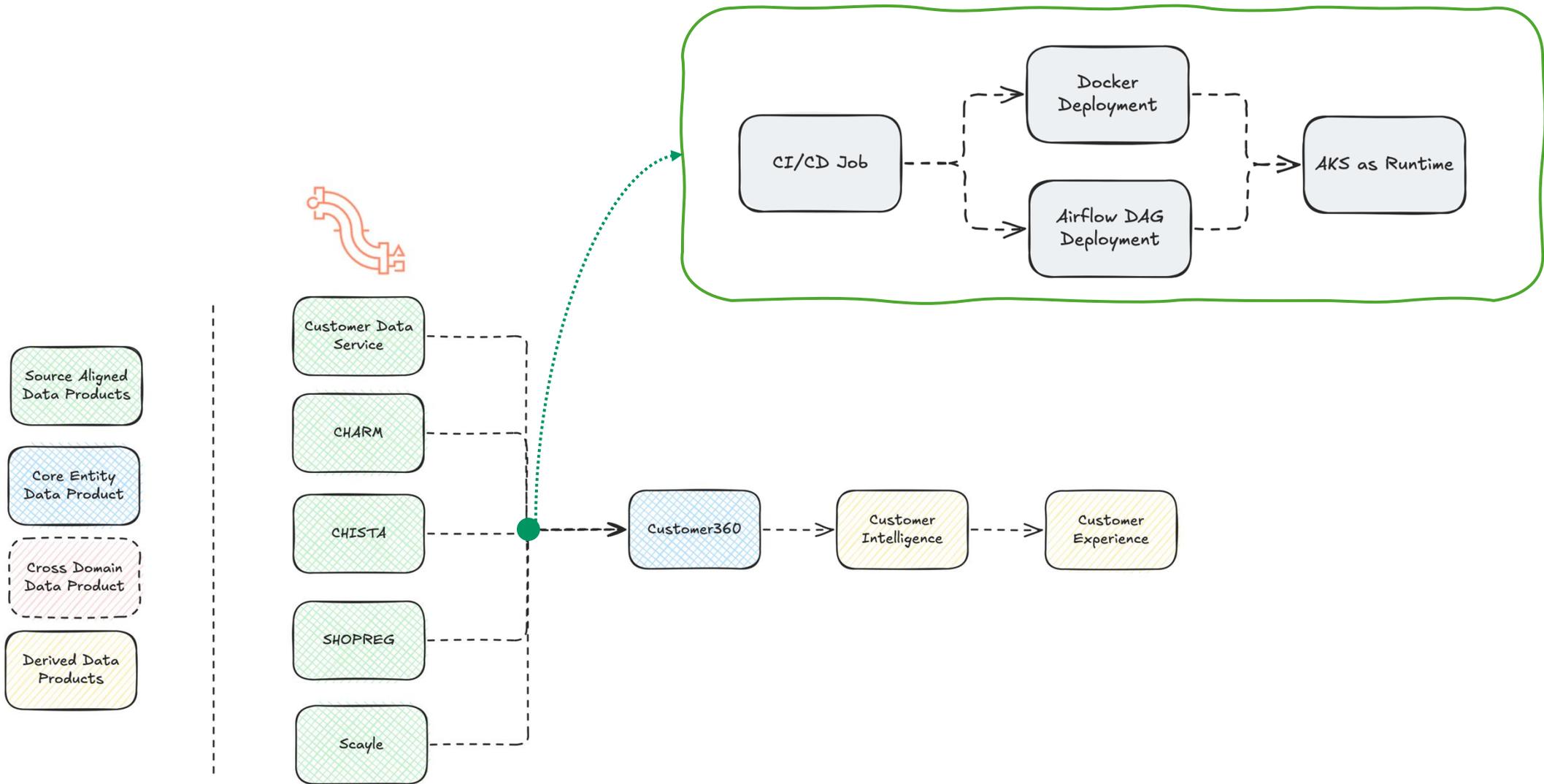
Nicht alles in DLT umsetzbar. Es mussten Workarounds gefunden und kleinere Abweichungen vom Data Vault Standard in Kauf genommen werden.



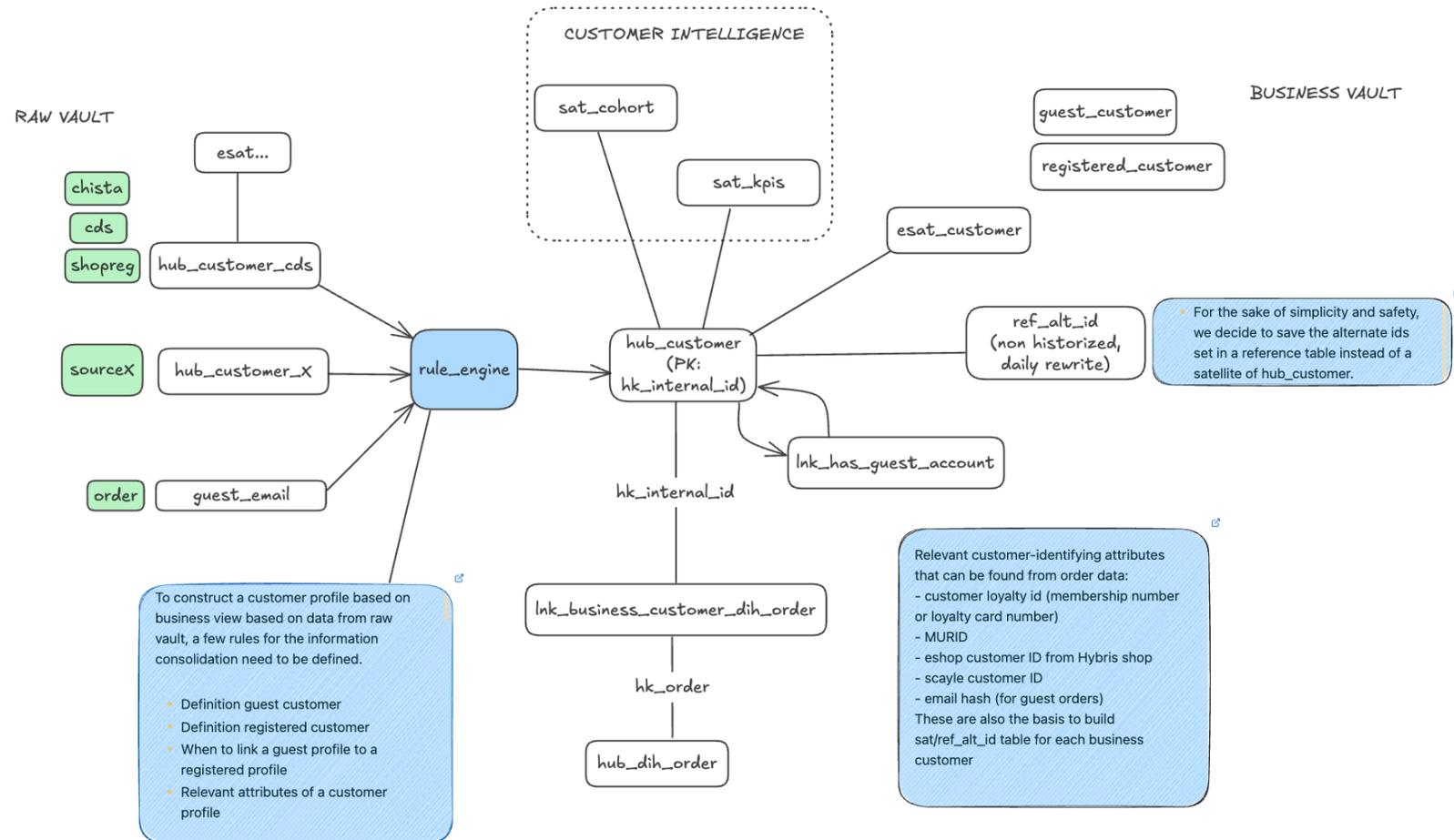
03.

Use Case “Customer Analytics”

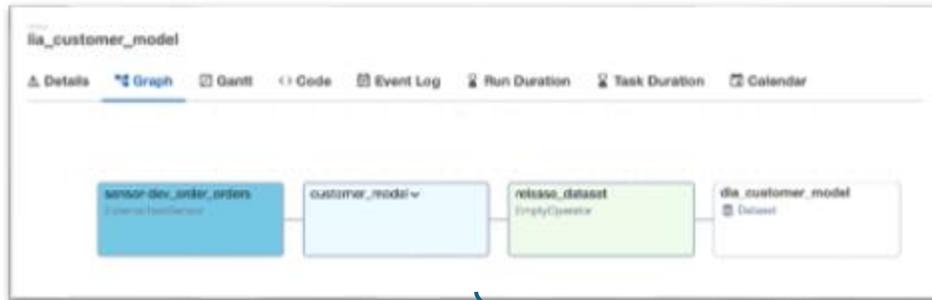




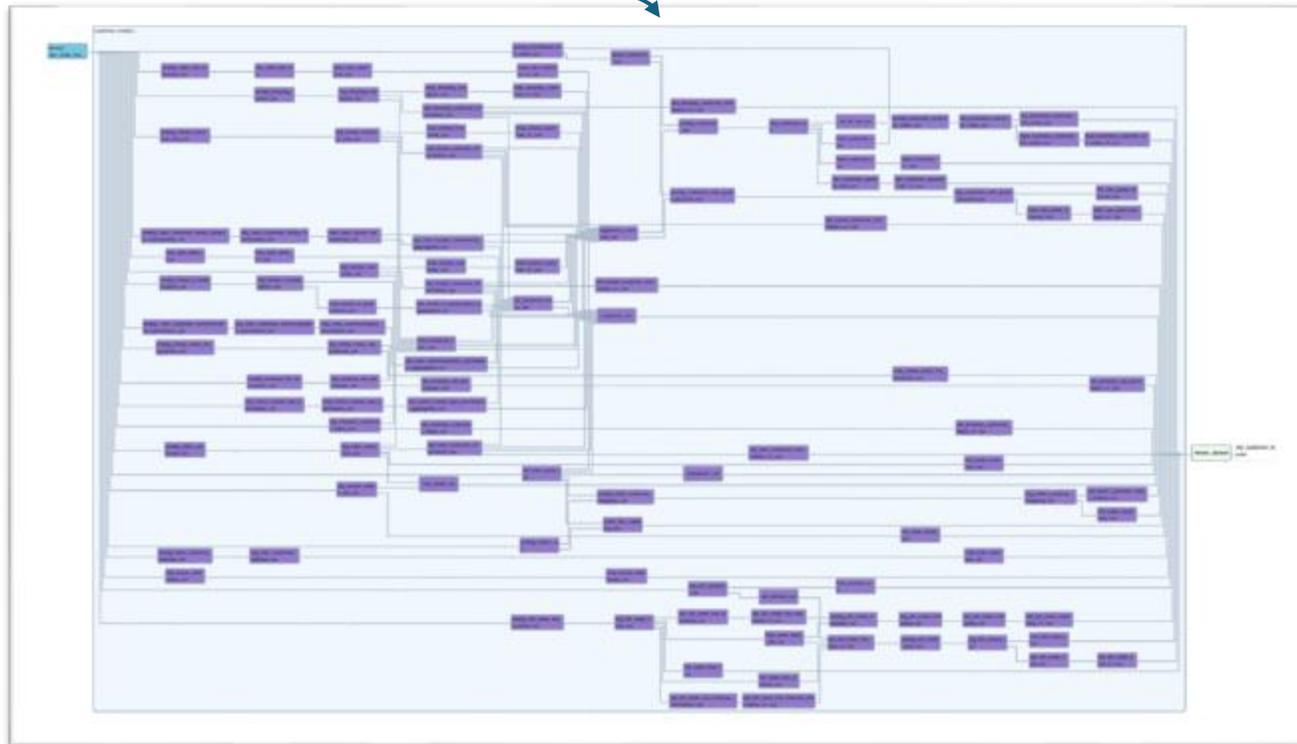
Auszug aus dem Customer Data Model



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 AWS Glue console interface for a 'customer_model_raw_vault'. It displays a table with the following columns: Name, Owner, Created at, and Popularity. The Owner column is redacted with a grey box. The table lists various data sources and their creation dates.

Name	Owner	Created at	Popularity
raw_cds_customer	[REDACTED]	Nov 07, 2025, 06:13 AM	0%
raw_cds_customer_v1	[REDACTED]	Nov 21, 2025, 07:17 AM	0%
raw_cds_customer	[REDACTED]	Nov 07, 2025, 06:12 AM	0%
raw_cds_customer_v1	[REDACTED]	Nov 21, 2025, 07:20 AM	0%
raw_cds_customer	[REDACTED]	Nov 07, 2025, 06:12 AM	0%
raw_cds_customer_v1	[REDACTED]	Nov 21, 2025, 07:23 AM	0%
raw_shipping_customer	[REDACTED]	Nov 07, 2025, 06:12 AM	0%
raw_shipping_customer_v1	[REDACTED]	Nov 21, 2025, 07:23 AM	0%
tbl_customer_idr	[REDACTED]	Jul 22, 2025, 11:50 AM	0%
tbl_email	[REDACTED]	Aug 19, 2025, 09:14 AM	0%
tbl_loyalty_member	[REDACTED]	Jul 22, 2025, 11:44 AM	0%
tbl_order_line_dr	[REDACTED]	Jul 22, 2025, 11:49 AM	0%
tbl_product	[REDACTED]	Sep 27, 2025, 12:11 PM	0%
tbl_customer_idr_loyalty_member	[REDACTED]	Jul 22, 2025, 11:47 AM	0%
tbl_email_customer	[REDACTED]	Aug 19, 2025, 09:18 AM	0%
tbl_order_line_product	[REDACTED]	Sep 27, 2025, 12:21 PM	0%
raw_cds_address	[REDACTED]	Jul 22, 2025, 11:48 AM	0%
raw_cds_communication_permission	[REDACTED]	Jul 22, 2025, 11:43 AM	0%
raw_cds_loyalty_membership	[REDACTED]	Jul 22, 2025, 11:50 AM	0%
raw_cds_loyalty_program_membership	[REDACTED]	Jul 22, 2025, 11:49 AM	0%



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