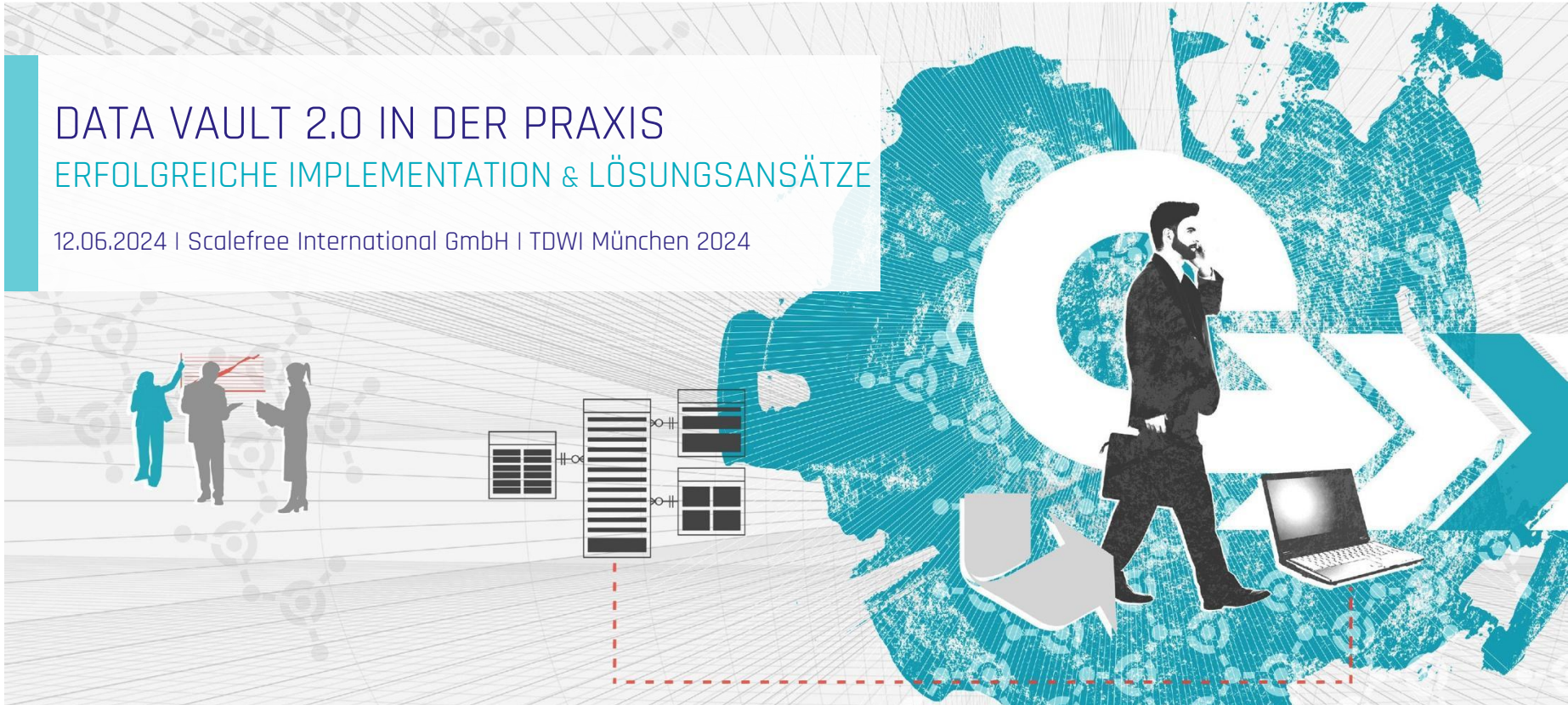


DATA VAULT 2.0 IN DER PRAXIS

ERFOLGREICHE IMPLEMENTATION & LÖSUNGSANSÄTZE

12.06.2024 | Scalefree International GmbH | TDWI München 2024



SPEAKER TRUNG TA

SENIOR BI-BERATER

Erfahrung

7 Jahre erfolgreiche Tätigkeit beim Aufbau von Scalefree International GmbH und bei der Beratung von Branchenführern in Banken- und Versicherungs-Sektoren beim Aufbau von BI-Lösungen mit Data Vault 2.0.

Ausbildung

B.Sc. Business Informatics
(Hochschule Hannover -
Schwerpunkt: Business Intelligence)

Certified Data Vault 2.0 Practitioner
(CDVP2™)

Certified SCRUM Master/Product Owner



Kompetenzen

Data Warehousing auf hybriden
Architekturen

Data Vault-Modellierung & -Implementation

Data Warehouse Automatisierung mit dbt,
WhereScape, Coalesce.

SPEAKER

TIM KIRSCHKE

SENIOR BI-BERATER

Erfahrung

5 Jahre Erfahrung mit Data Warehousing und Data Vault 2.0 Entwicklung und Verwaltung von Data Vault Automatisierungs Software

Ausbildung

B. Sc. Angewandte Mathematik (Hochschule Hannover - Schwerpunkt: Data Warehousing)

Certified Data Vault 2.0 Practitioner (CDVP2™)

Azure DP-203, Certified dbt Practitioner, Coalesce Certified



Kompetenzen

Data Warehousing auf hybriden Architekturen

Data Vault-Modellierung & -Implementation

Data Warehouse Automatisierung mit dbt, WhereScape, Coalesce.

AGENDA

1. Vorstellung Scalefree

2. Fortgeschrittene DV2.0-Implementationen

3. Vorstellung Scalefree's Open-Source-Tools

4. Q&A

EINFÜHRUNG

SCALEFREE INTERNATIONAL GMBH



BEREICHE IN DENEN WIR ARBEITEN

Business
Intelligence



Big Data



Data
Warehousing



Data Strategy



Data Analytics



Data Science



CRM Systems



Data Integration



WELTWEIT FÜHRENDE EXPERTEN IN

Data Vault 2.0



BERATUNG

DATA STRATEGY
WORKSHOPS
REVIEW



IMPLEMENTIERUNG

PROOF OF CONCEPTS
MIGRATION PROJEKTE
TOOL SETUPS
ARCHITEKTUR SETUPS



TRAINING

EINFÜHRUNGS CLASS
BOOT CAMP CLASS
MODELING CLASS
CUSTOMIZED CLASS



EVENTS

DATA DREAMLAND
WEBINARE
EXPERT SESSIONS



GRÜNDER

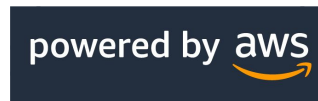


MICHAEL OLSCHIMKE



DAN LINSTEDT

PARTNERAUSWAHL



DATA VAULT 2.0 IN DER PRAXIS

ERFOLGREICHE IMPLEMENTATIONEN



FUNKTIONIERT DATA VAULT 2.0?

3000+ DV2 Bootcamp Teilnehmer

100+ DV2-Implementierungen unterstützt

DATA VAULT 2.0 STELLT (FAST) IMMER EINE GUTE LÖSUNG DAR!

ÜBER DEN TELLERRAND HINAUS

Data Mesh & Data Vault 2.0

Realtime-Daten mit Data Vault 2.0

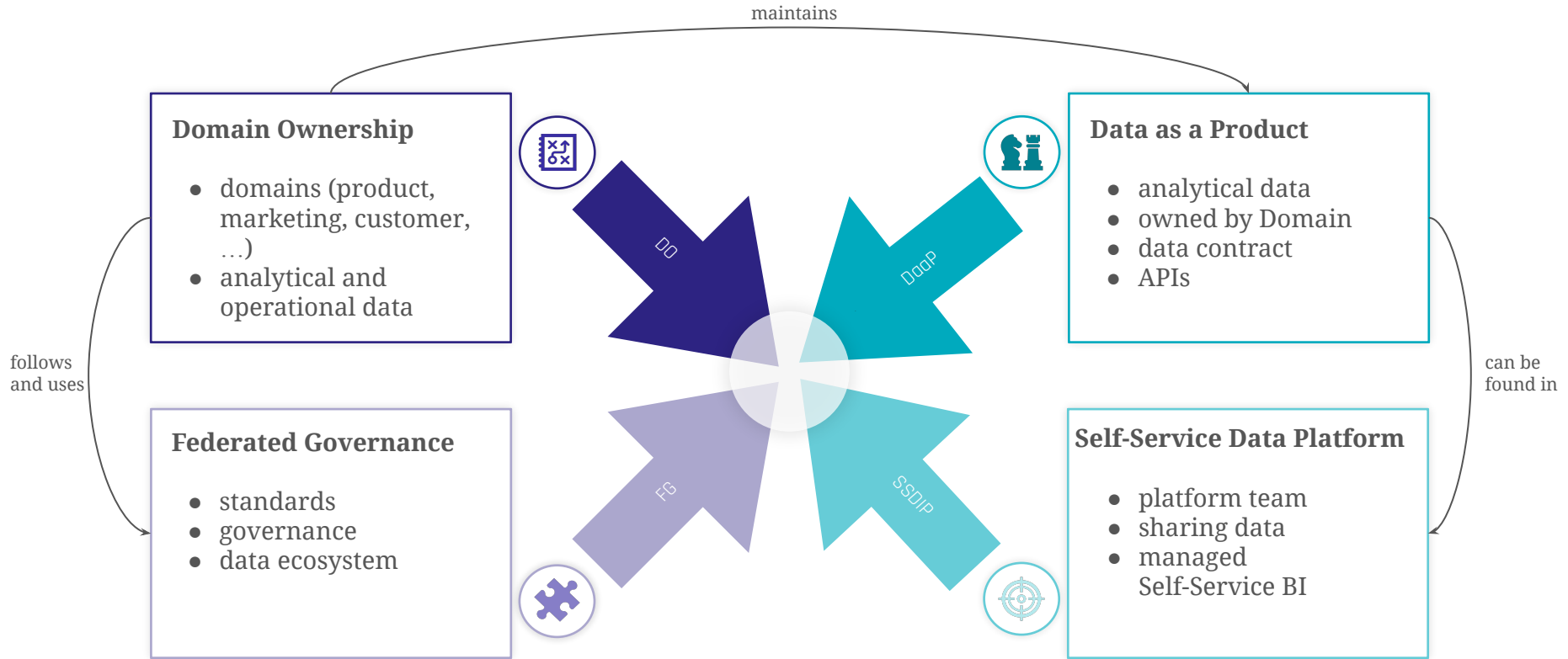
Fehlende Business Keys

Handhabung semistrukturierter Daten

DATA MESH x DATA VAULT 2.0
IT'S A MATCH!

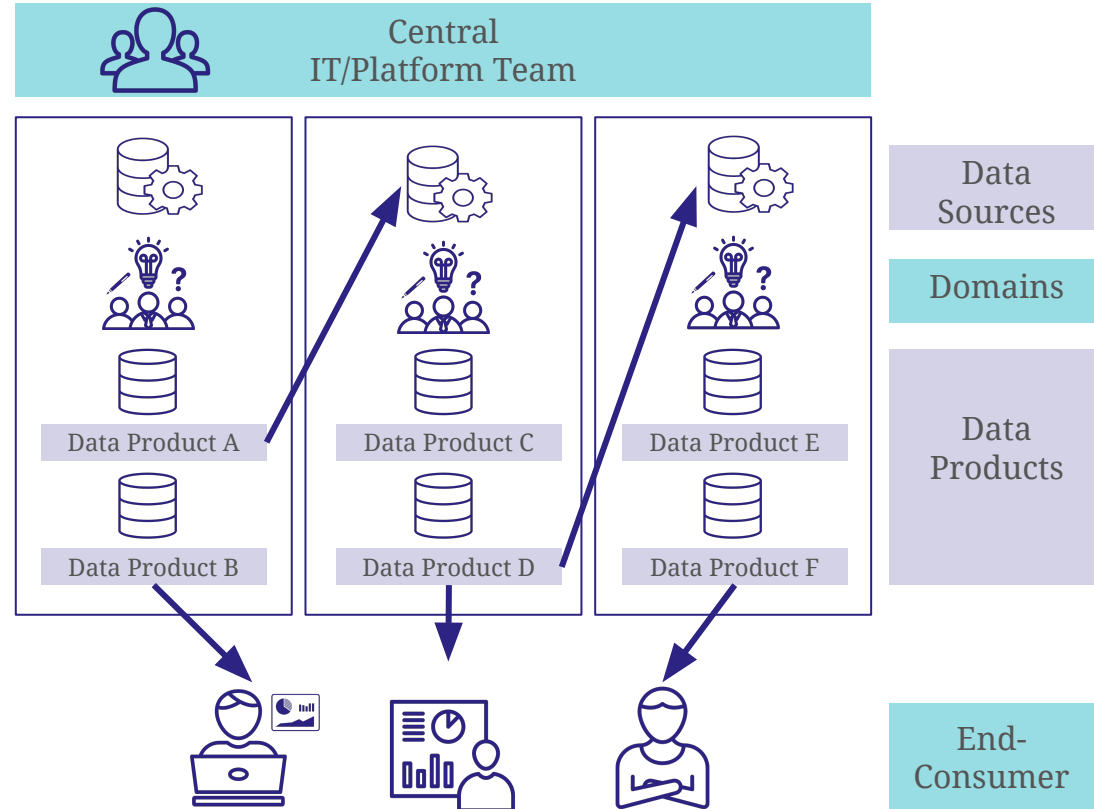


WHAT IS DATA MESH?

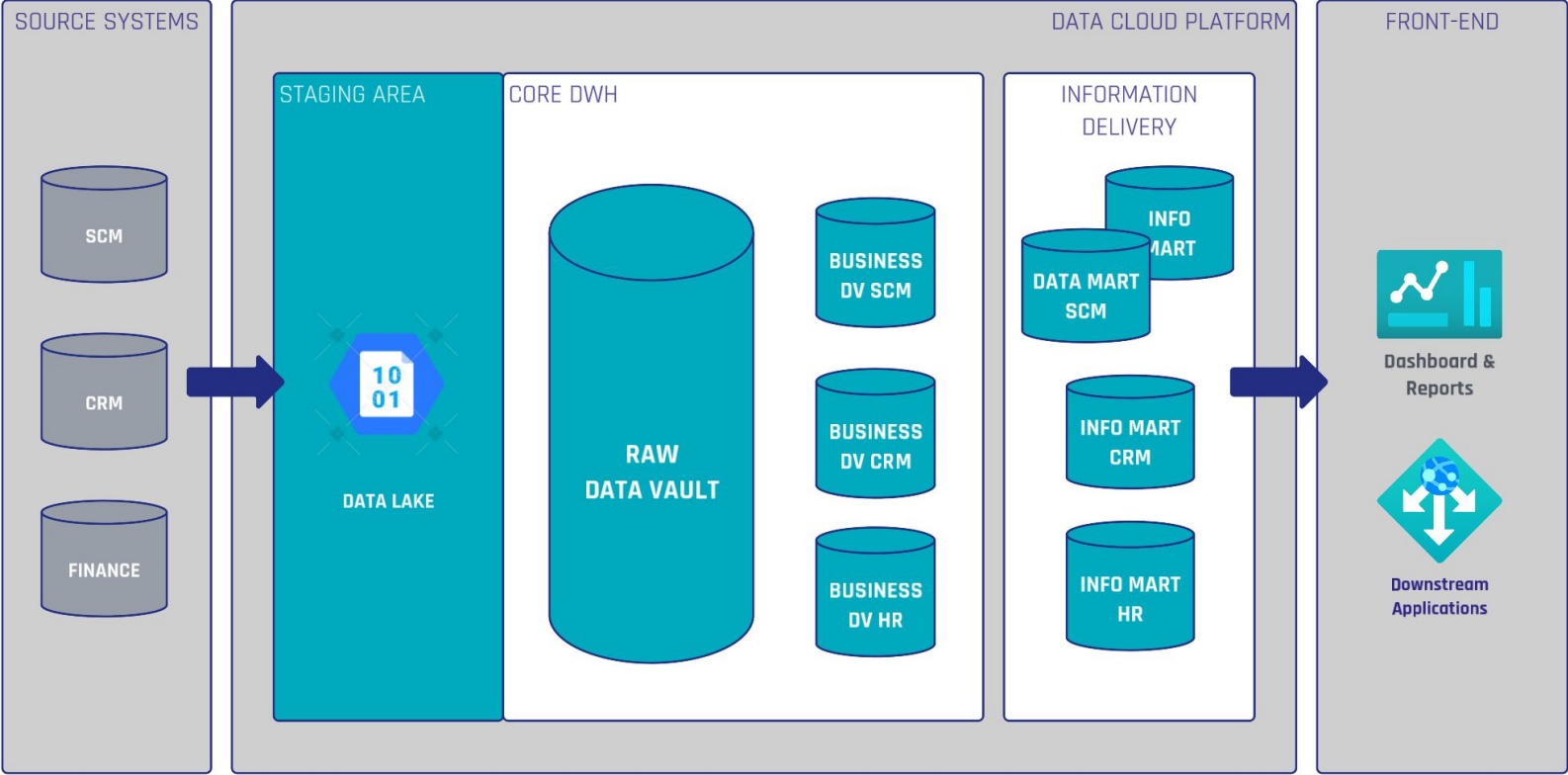


DEZENTRALISIERTER DWH-ANSATZ

- Zentrales IT-Team: Plattform, Standards, Data Governance,...
- Domänen besitzen Wissen über ihre Daten
- Gemeinsame Nutzung von Daten über Data Contracts,
- “Shopping” für Datenkonsumenten



ARCHITEKTURBEISPIEL FÜR DATA MESH MIT DATA VAULT



REAL-TIME DATA WAREHOUSE
MIT DATA VAULT 2.0

HERAUSFORDERUNGEN

Teilmenge an Quellobjekten kommt als Realtime Stream

Mehrere Geschäftsobjekte teilen sich einen Stream

REALTIME DATEN

KAPPA ARCHITEKTUR

INTEGRIERTE DATENBANK TRIGGER



Nachricht

Stage

Raw Vault

**Business
Vault**

**Information
Marts**

Nicht auf jeder Datenbank verfügbar

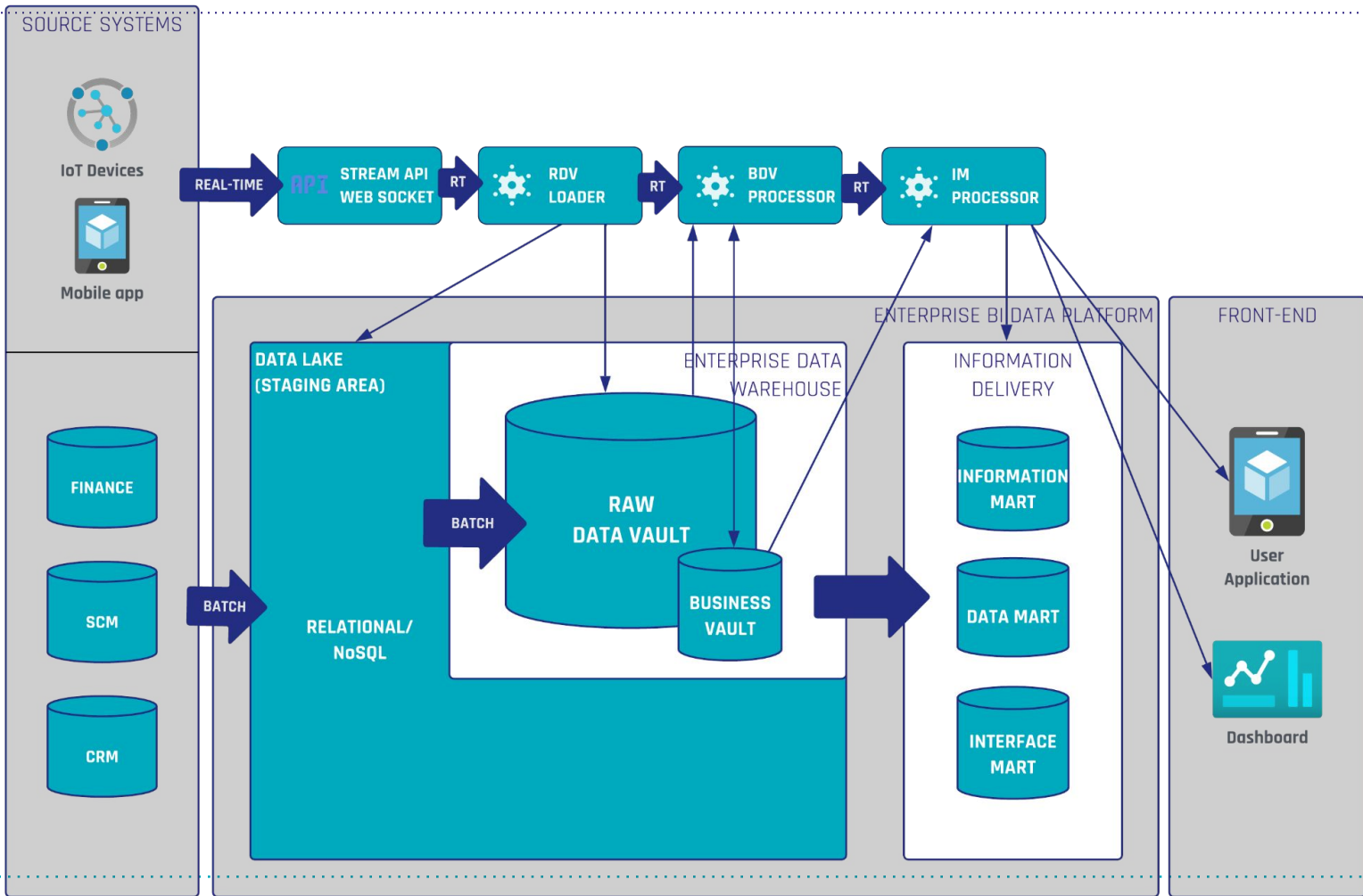
Eingeschränkt auf DB Operationen

LAMBDA ARCHITEKTUR

Eigenständige Realtime Pipeline

Triggert reguläre Beladungslogiken oder führt eigene aus

Kann andere Ziele integrieren



HANDHABUNG SEMI-STRUKTURIERTER DATEN MIT DATA VAULT 2.0



VERTEILTE LÖSUNG: DATENSTRUKTUREN

Verwenden Sie die richtige Technologie für die richtigen Daten!

- Strukturierte Daten: relationale Datenbank-Engine (nicht unbedingt eine relationale Datenbank)
- Unstrukturierte Daten: Data Lakes (HDFS, S3, Azure Blob Storage, Google Cloud Storage, etc.)
- Semi-strukturierte Daten: ?

BEST PRACTICES

Satellite mit JSON-Payload (~Schema-on-Read)

- Datentypen-Unterstützung abhängig von Zielplattform, Beispiele:
 - Snowflake → VARIANT,
 - BigQuery, Oracle (v.18c+), MS SQL Server (2016+) → JSON,
 - Postgres → JSONB
- Tipp: Beachte Limitation des zu verwendenden Datentypen!
 - z.B. 16MB max. Volumen für Snowflake's VARIANT-Datenfeld

FEHLENDE BUSINESS KEYS



BUSINESS KEYS

Unternehmensweit

Mitarbeiternummer,
AssetID



Systemweit

Bestellnummer,
Kundennummer

Global

VIN, ISBN, IBAN, ISIN,
WKN

BUSINESS KEYS

Keine Business Keys?

→ **Technische Keys!**

Keine Technischen Keys?

→ **Sequences!**

SEQUENCES

ROW_NUMBER() OVER (ORDER BY)

“BK”



					1
					2
					3
					4

Keine Delta Erkennung möglich!

SCALEFREE OPEN SOURCE TOOL STACK

EINFACH EINSTEIGEN



OPEN SOURCE TOOLS VON SCALEFREE

01 DATAVAULT4dbt
POWERED BY SCALEFREE

DV2.0 Beladungsmuster für
Automatisierungstool dbt

02 TURBOVAULT4dbt
POWERED BY SCALEFREE

Front-End Automatisierung für
DATAVAULT4dbt

03 DATAVAULT4coalesce
POWERED BY SCALEFREE

DV2.0 Beladungsmuster für
Automatisierungstool Coalesce

DATA VAULT 2.0 SCHNELLSTART



Kostenloses Trial dbt
Cloud

Open Source Version als
Python Package



Kostenloses Trial via
Snowflake Partner Connect

DATA DREAMLAND 2024

UNLOCK THE POWER OF DATA-
DRIVEN BUSINESS

SEPTEMBER 23RD & 24TH, 2024
HANNOVER

GET YOUR TICKET!

POWERED BY
SCALEFREE



www.scalefr.ee/ddl2024



**DATA VAULT
INNOVATORS**

SCALEFREE 

JOIN THE DATA VAULT
INNOVATORS COMMUNITY
DVIC

COMMUNITY SUMMIT
11.07.2024
MÜNCHEN
KOSTENLOS

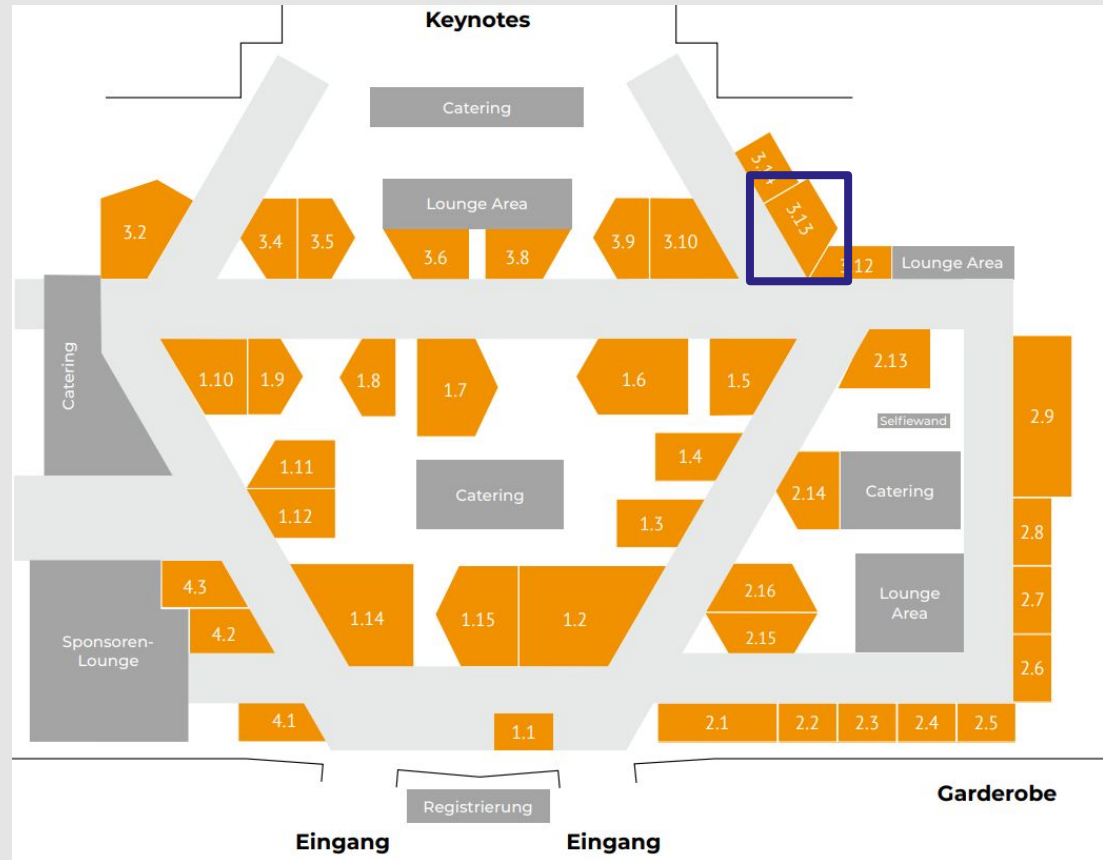


scalefr.ee/dvic-signup



NOCH FRAGEN?

SCALEFREE STAND 3.13



11. - 13. Juni 2024

Wie hat Ihnen die Session gefallen?
Geben Sie uns gerne Ihr Feedback



Mi 8.3

12.06.2024

11:25-12:10

Data Vault 2.0 in der Praxis: Erfolgreiche Implementierung und Lösungsansätze