

Business Vault Beispiele

Petr Beles - DDVUG 30. März 2017

petr.beles@datavault-builder.com

<http://datavault-builder.com>



DWH AUTOMATION

Lifecycle

Sandbox

Development

Testing

Depreciation

Production - Operations

- ① create environments.
- ② deploy code
- ③ test

Layers

Modelling

Starting

Core



Business Rules

- when a then b
- only xyz
-

Presentation

- Deliver all active customers

Conceptual

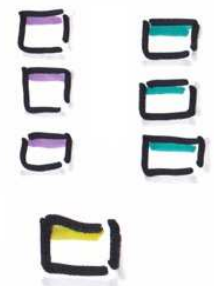
*



- if a in the range of 1-100 then b

- All customers with flag b

Logical



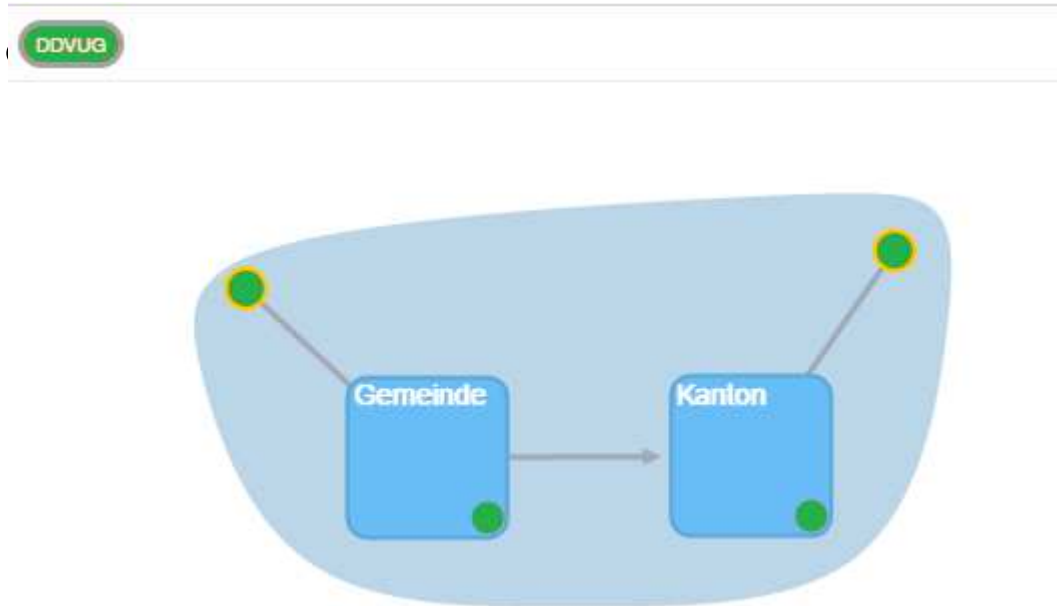
- --Decode a → b
- Case when a between 1 and 100 then b else null end as...



Technical

* optionally use industry standard model as starting point

Raw Vault



Quellsystem: DDVUG

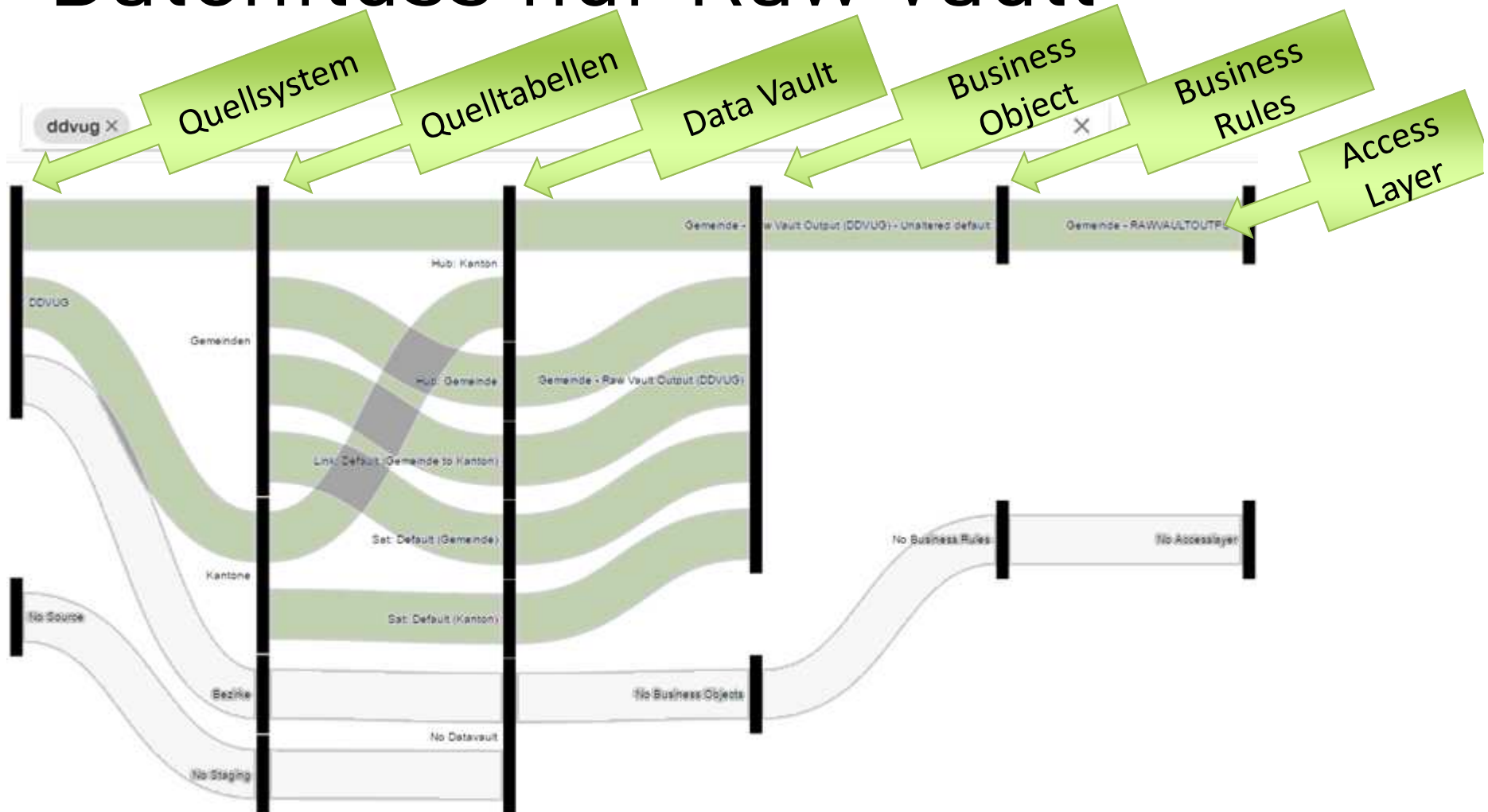
Grosse Boxen: Hubs, kleine Punkte Satelliten

Zwei Tabellen befüllen Hub Gemeinde und Kanton



Auch wenn technisch immer N:M Tabellen eingesetzt werde, kann man auf logischen Stufe die erwarteten Granularitäten von Beziehungen definieren

Datenfluss nur Raw Vault



Stufe 1: Daten fließen in den Hub und werden nur virtuell abgefragt

Stufe 2: Aus Performance Gründen könnte man den Output materialisieren

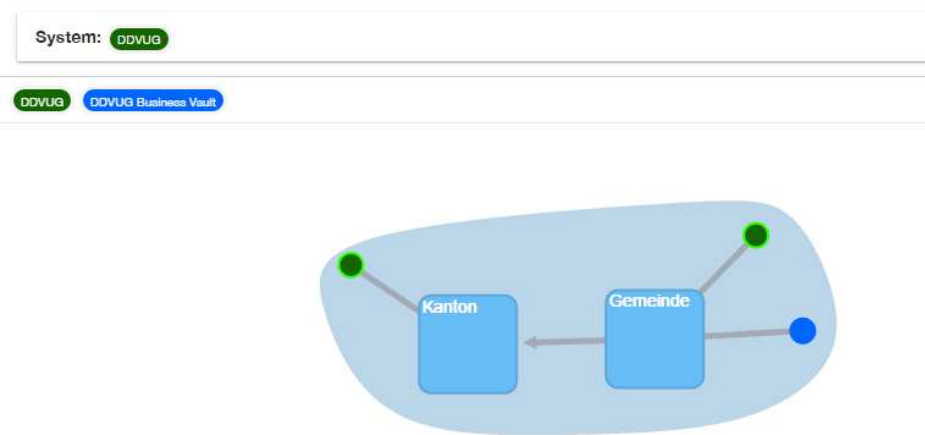
Business Vault Load: Nur ein weiteres Quellsystem*

The screenshot displays a data integration tool interface. At the top, a search bar contains 'ddvug' and a 'NEW SOURCE' button. Below is a table with columns: System, System ID, Source Table, Source Schema, and Staging Table. The table lists two systems: 'ddvug' and 'ddvugbusinessvault'. The 'ddvug' system has three source tables: 'Bezirke', 'Gemeinden', and 'Kantone', each mapped to a corresponding staging table. The 'ddvugbusinessvault' system is currently selected, and a modal window titled 'Add Table: DDVUG Business Vault' is open. The modal has tabs for 'BASE', 'COLUMNS', 'DETAILS', and 'REVIEW'. The 'COLUMNS' tab is active, showing a search for columns from 'DDVUG'. The 'Available Columns:1' list includes 'Gemeinde Name' (text). The 'Selected Columns:2' list includes 'Business Key for Hub Gemeinde' (text) and 'Kanton Name' (text). Buttons for 'ADD', 'ADD ALL', 'MOVE TO START', and 'MOVE TO END' are visible.

System	System ID	Source Table	Source Schema	Staging Table
ddvug	ddvug	Bezirke	Bezirke	Bezirke [ddvug_r_Bezirke]
ddvug	ddvug	Gemeinden	Gemeinden	Gemeinden [ddvug_r_Gemeinden]
ddvug	ddvug	Kantone	Kantone	Kantone [ddvug_r_Kantone]
DDVUG Business Vault	ddvugbusinessvault			

* **Martijn Evers:** A business rule is just a small (source) data model that happens to derive its data from the Data Vault and not externally. It is (normally) loaded in the same manner and into the same kind of objects as with a raw Data Vault.

Data Vault: Raw Vault + Business Vault

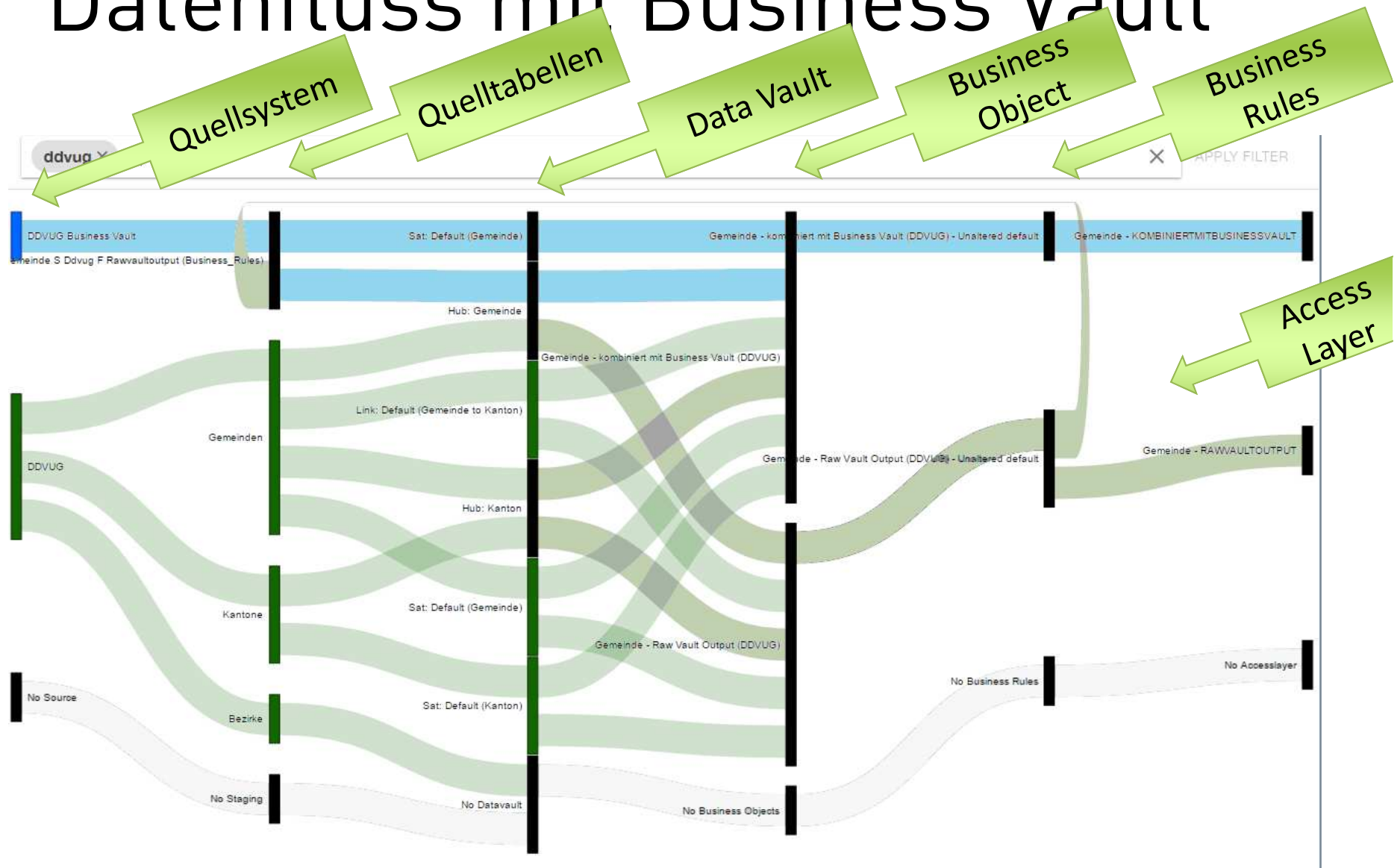


Quellsystem: DDVUG + DDVUG Business Vault
Business Vault erzeugt weitere Satelliten / Links im gleichen Modell



Es gibt auch Ansätze, welche Raw und Business Vault trennen

Datenfluss mit Business Vault



Stufe 3: Aus Auditability Gründen oder Performance werden die Daten in den BV geladen

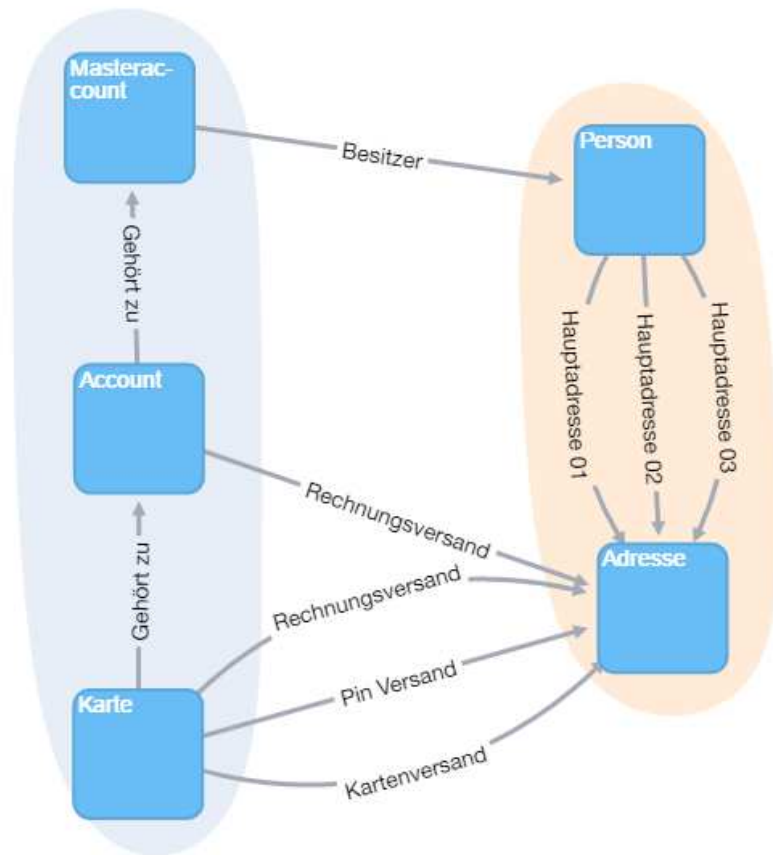
Beispiel Auditability:

- Kreditkarten Unternehmen
- Im Core wird die Beziehung zwischen Karte, Account, Masteraccount und Person gespeichert
- Es gibt pro Entität verschiedene Adressen je nach Funktion
- Ziel: Datenqualität und Nachverfolgbarkeit verbessern nach dem Versand von Dokumenten

Beispiel Auditability:

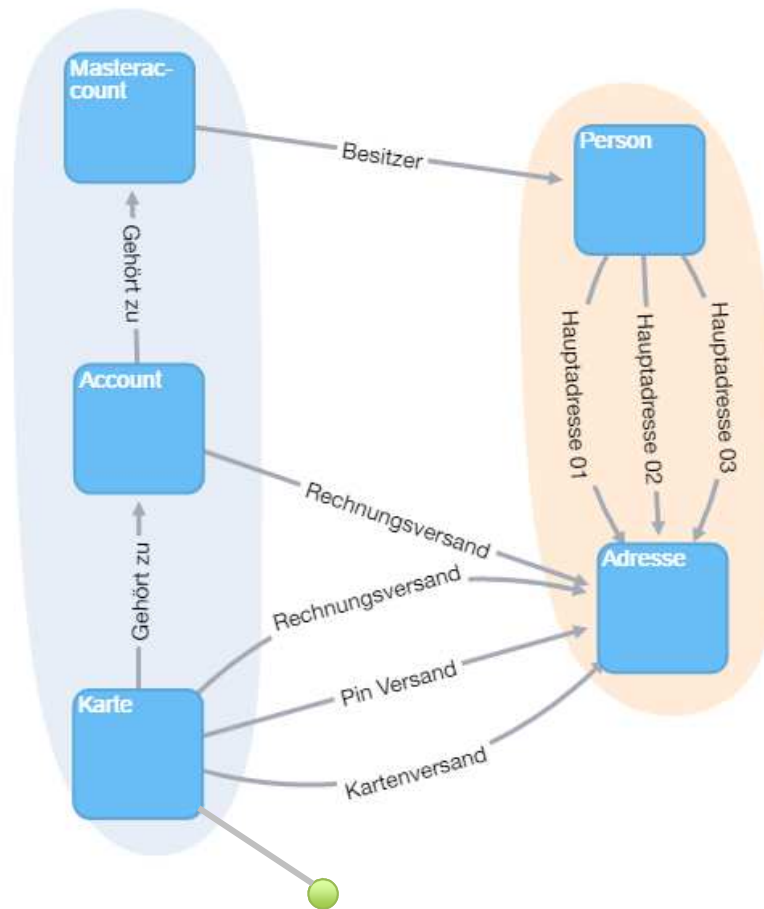


Beispiel Auditability:



- Verschiedene Adressen pro Stufe
- Vererbung von Funktionen
- Erstellung einer View / Granularität und Funktion
- Erstellung der Vererbungslogik als View
- Testing
- Materialisierung der fertigen Logik als Business Vault

Beispiel Auditability:

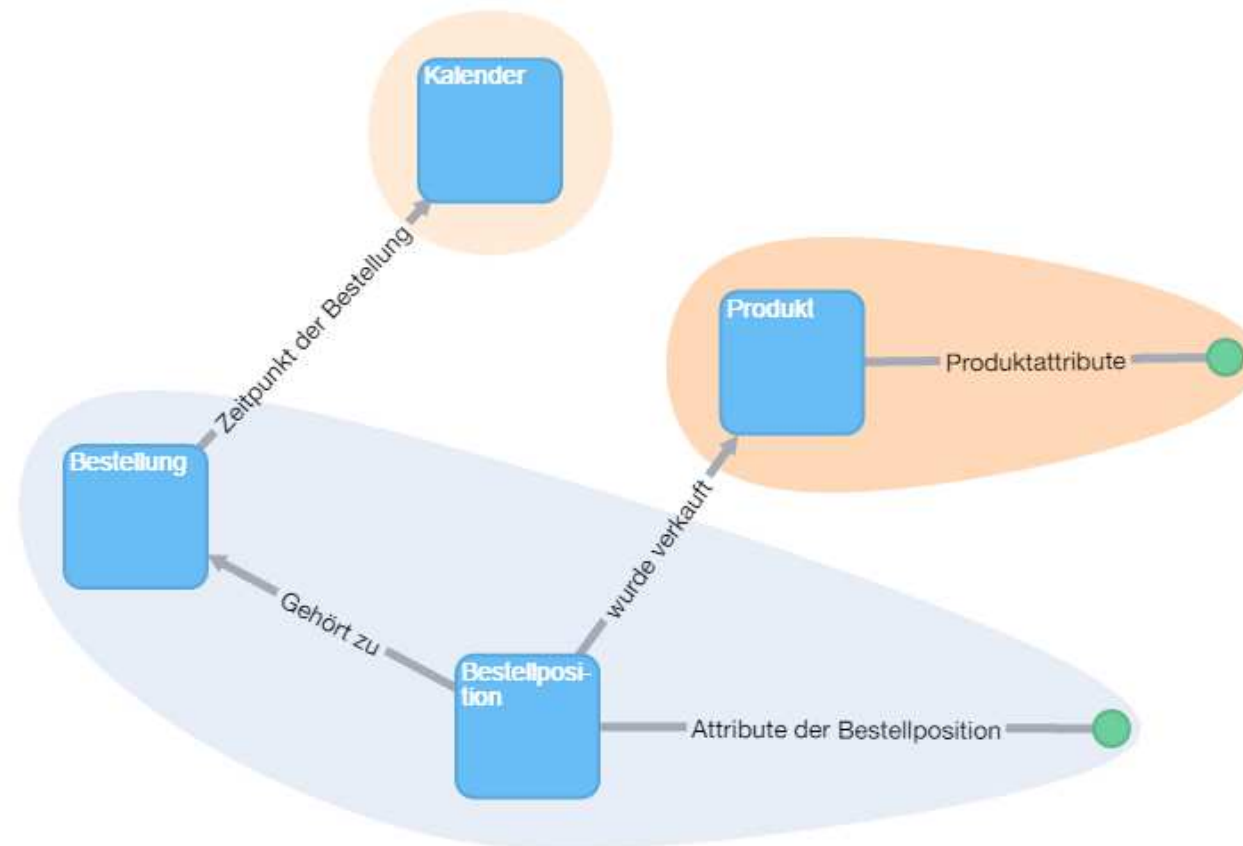


- Direkt auf der tiefsten Granularität wird die nach Vererbungslogik hergeleitete Adresse abgelegt
- Somit kann auch nachvollzogen werden, wann welche Dokumente wohin adressiert wurden

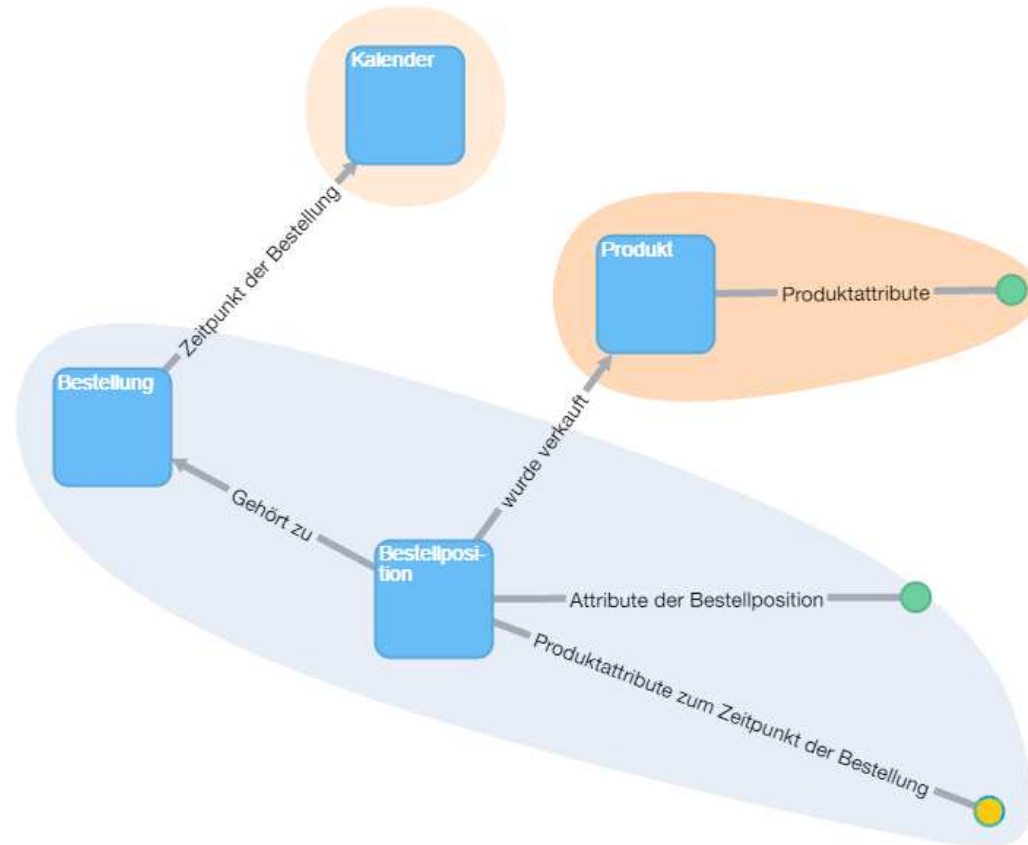
Beispiel Auflösung von SCD Attributen

- Industrie Unternehmen
- Produziert einen Teil der Artikel selbst, aber lässt auch gewisse Teile produzieren
- Dies ändert aber über Zeit, wie auch die Produktverantwortlichen

Beispiel Auflösung von SCD Attributen



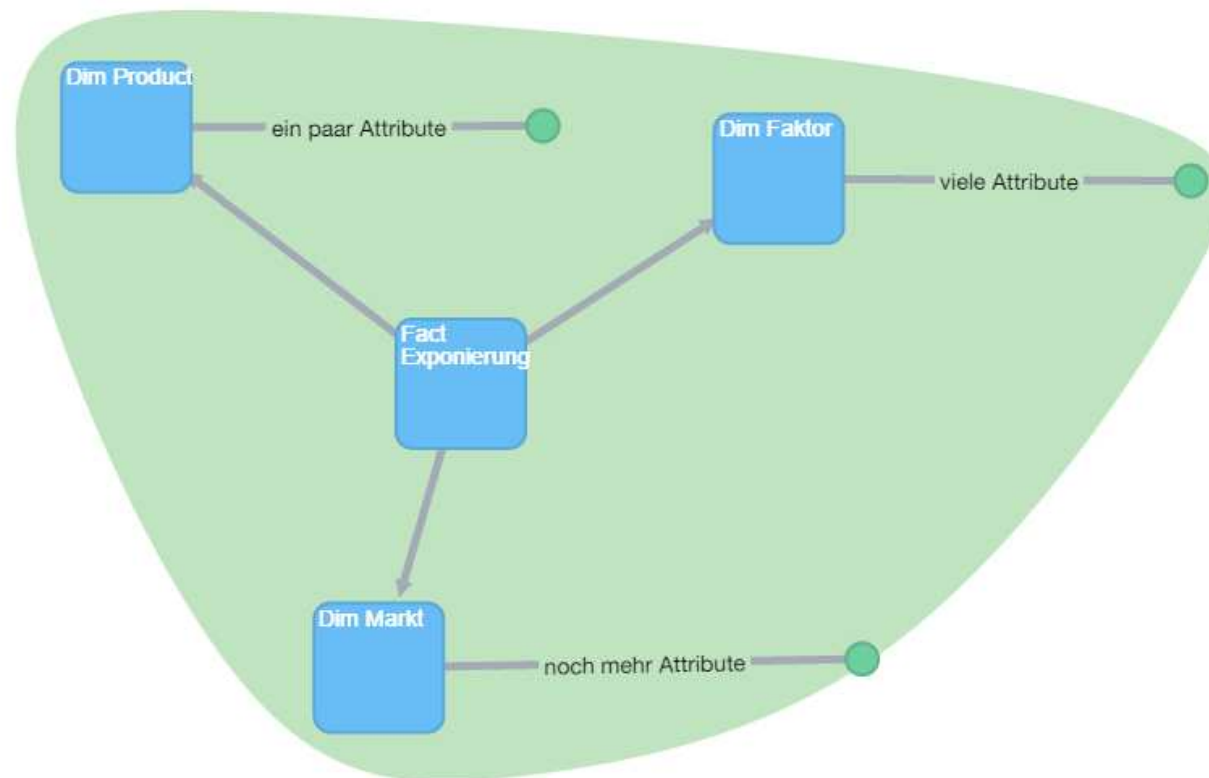
Beispiel Auflösung von SCD Attributen



Beispiel Point in Time Table

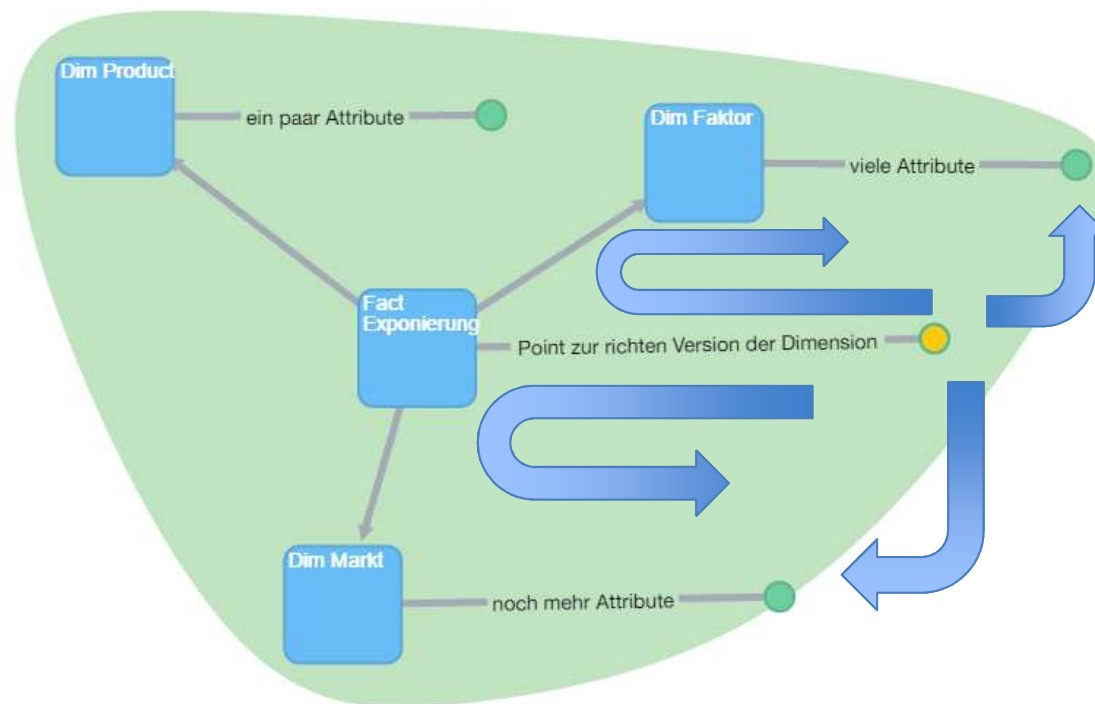
- Rückversicherung
- Konsolidiert Risiko und Finanzdaten
- Quelle ist ein operationell historisiertes Daten-Modell

Beispiel Point in Time Table



Data Vault würde die Modellierung der Fact auch als Link anstatt Hub erlauben

Beispiel Point in Time Table

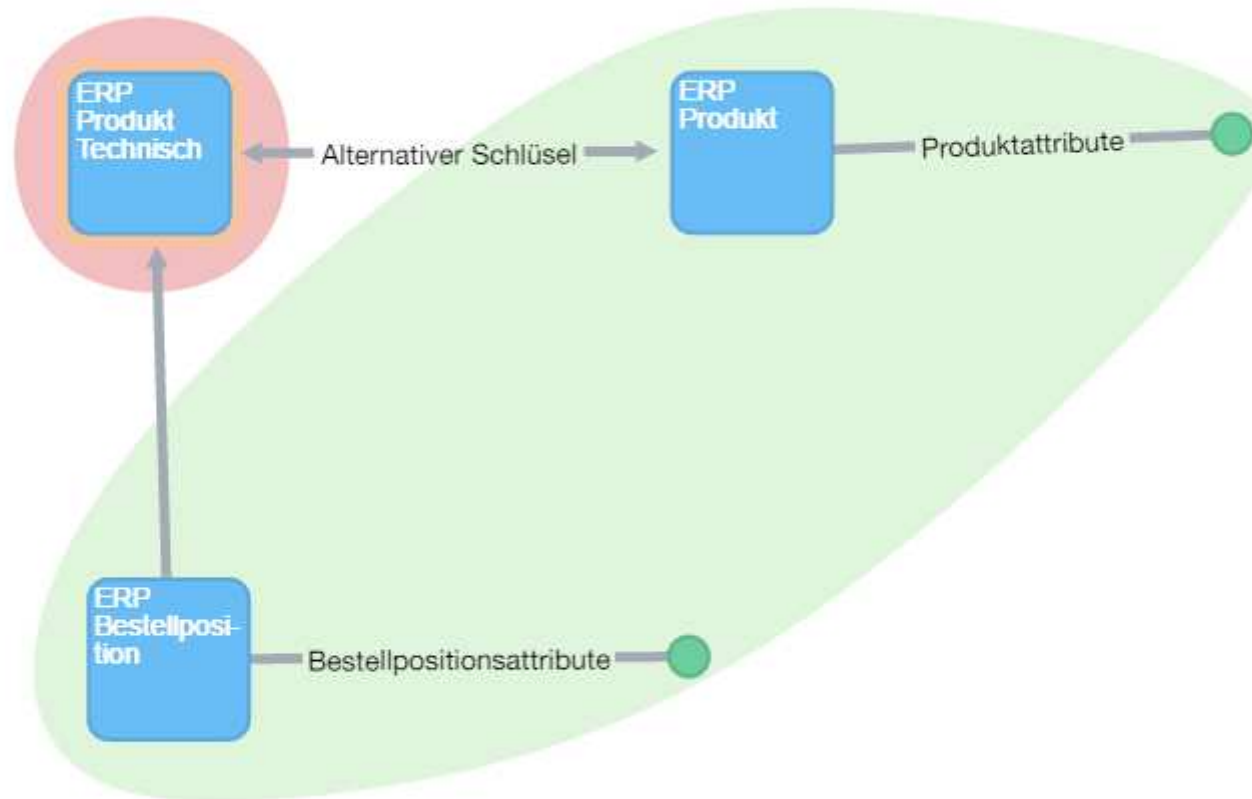


Diskussion: sind PIT Tabellen nur für einen Hub mit allen seinen Satelliten gedacht oder soll das hier dargestellte Szenario auch unterstützt werden

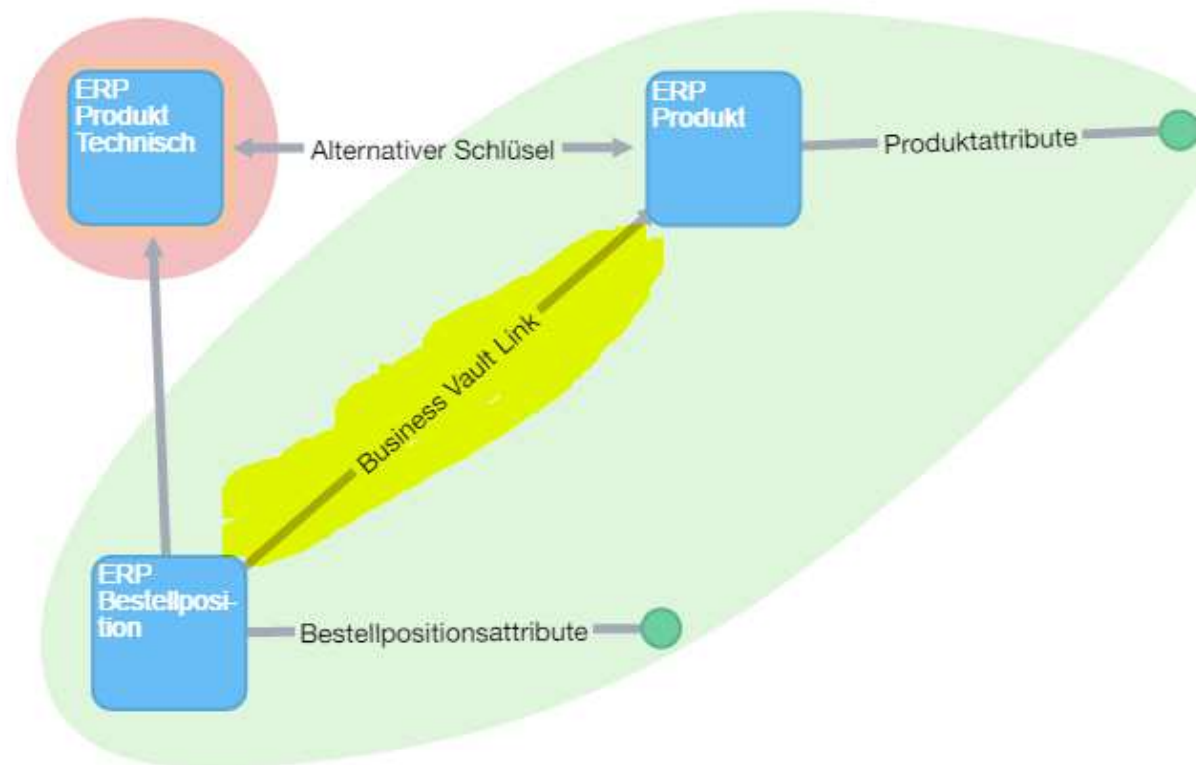
Beispiel Delta Loads

- Industrie
- Quellsysteme liefern nur Änderungen
- Wo geschieht die Auflösung von technischen zu Business Keys?
- Frage stellt sich vor allem bei hoch normalisierten Quellsystemen

Beispiel Delta Loads



Beispiel Delta Loads

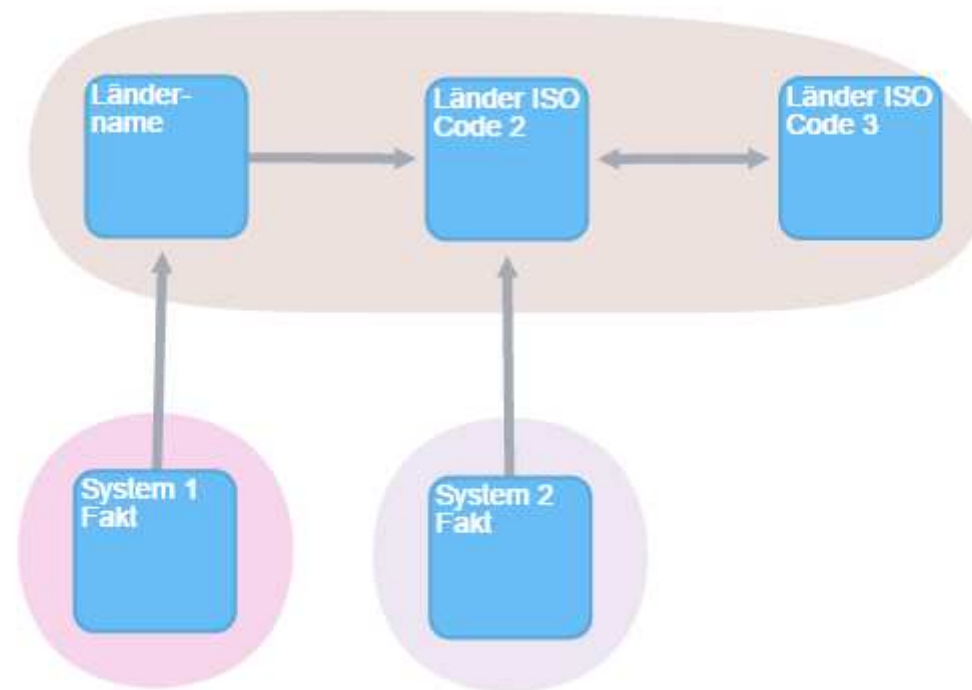


i Alternativ kann man auch eine Input Historisierung in Betracht ziehen

Verschiedene Keys pro Quelle

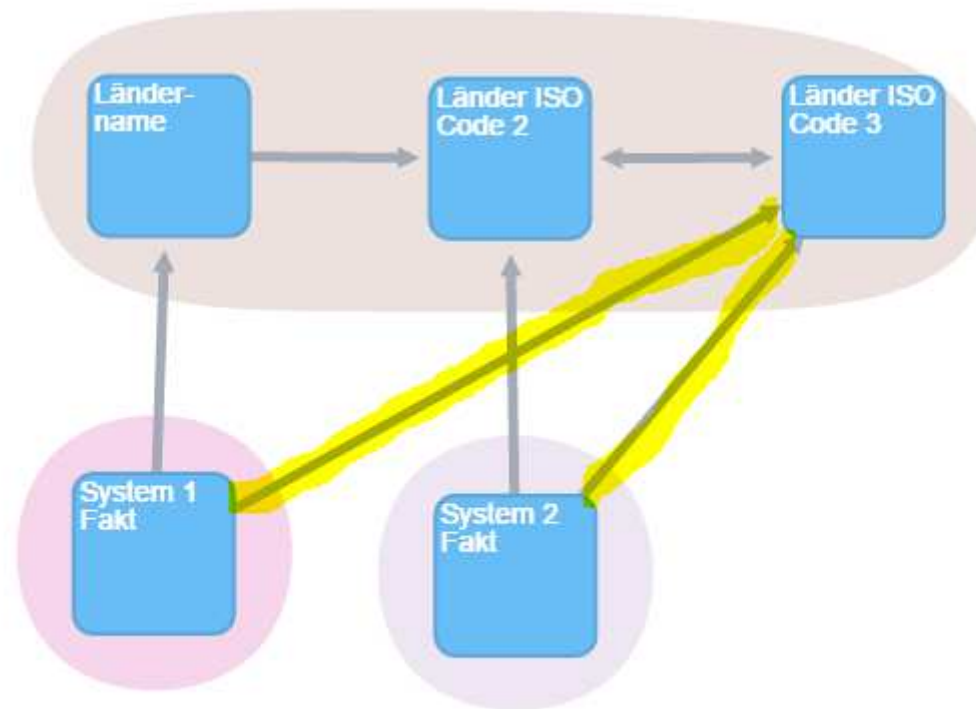
- Finanzindustrie
- Verschiedene Quellsysteme sind nicht harmonisiert und liefern unterschiedliche Codes für das Gleiche Objekt

Verschiedene Keys pro Quelle



? Darf man für das Gleiche Objekt mehrere Hubs anlegen

Verschiedene Keys pro Quelle

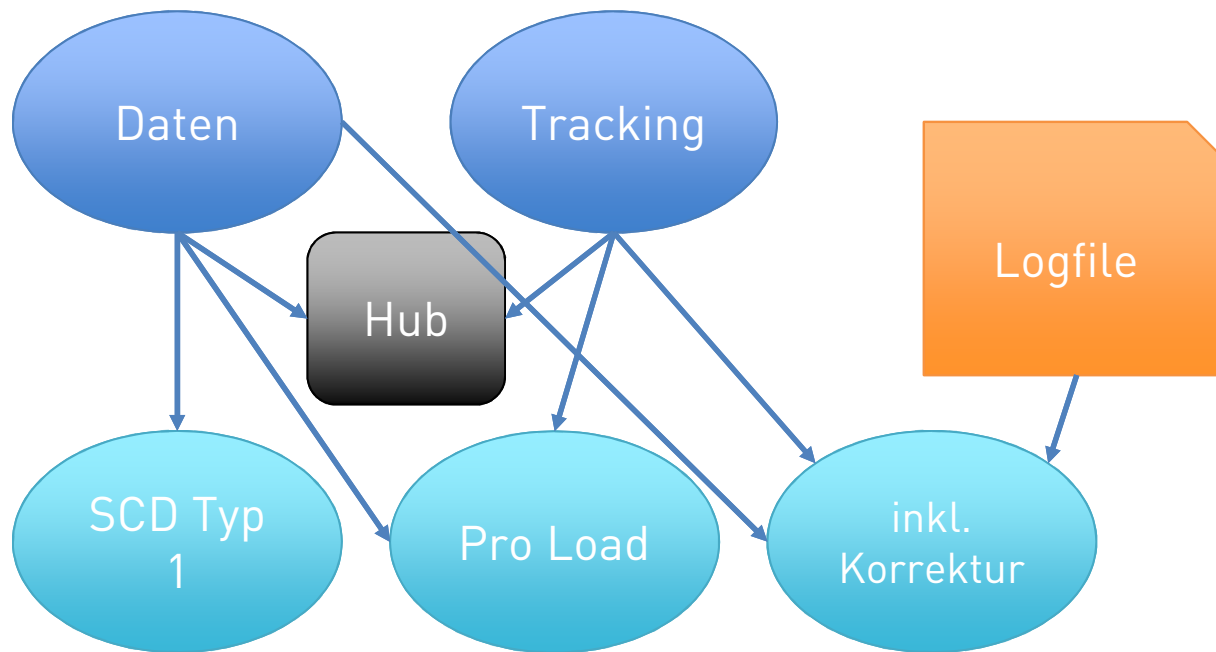


i Dies entspricht am ehesten den Bridge Tables aus dem Data Vault Konzept

Risk Warehouse

- Rückversicherung
- Risiko Modellierer haben andere Sichten als Legal Reporting
 - Sicht pro Datenload
 - What If Szenarion
 - Correction Loads
 - Legal Reporting Sicht

Virtuelle Satelliten



Virtuelle Satelliten

- Hub oder Link besitzt Daten und Tracking Satellit
 - Es gibt einen SCD Typ 1 Pointer auf den aktuellen Dateneintrag (MatView)
 - Es gibt einen SCD Typ 1 Sicht (Join MatView und Daten)
 - Es gibt eine Sicht pro Datenload (Join Tracking Sat mit Daten)
 - Es gibt eine Sicht pro Datenload mit Korrekturen (Join Tracking Sat mit Daten und Logfile)

Fragen?



- petr.beles@datavault-builder.com