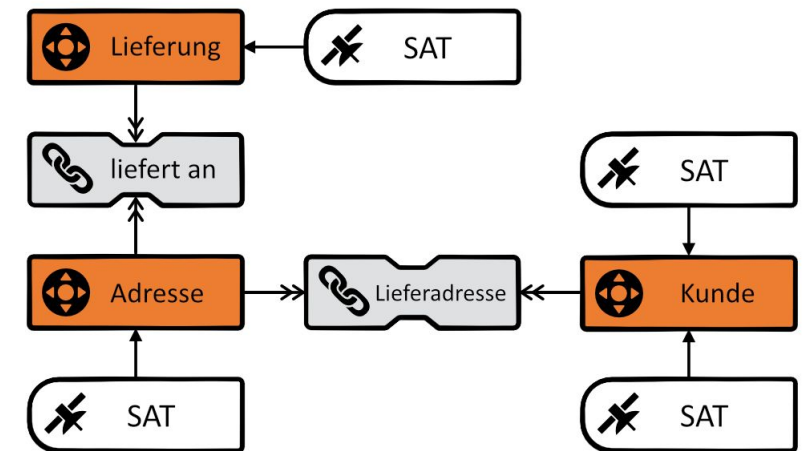




Dataspot (DataCatalog++) und Vaultspeed

Eine Automation der Automation



Dataspot (DataCatalog++) und Vaultspeed



- Dataspot - Wir schaffen einen 4-dimensionalen Überblick über dein Business: Daten, Geschäftsfunktionen & Prozesse, Systeme und Verantwortliche
- Vaultspeed - Ultra-automated data transformation, Build superior data products in less than 2 sprints in your Cloud Data Warehouse, Lakehouse, or Mesh.
- Alligator Company - Metadatenbrücke - Model-driven Data Vault
- Willibald Beispieldatensatz für Data Vault unter Lizenz CC-4.0-BY

dataspot.



<https://github.com/m2data/Willibald-Data>



Samen- und Pflanzenhandel
WILLIBALD



<https://dwa-compare.info/>

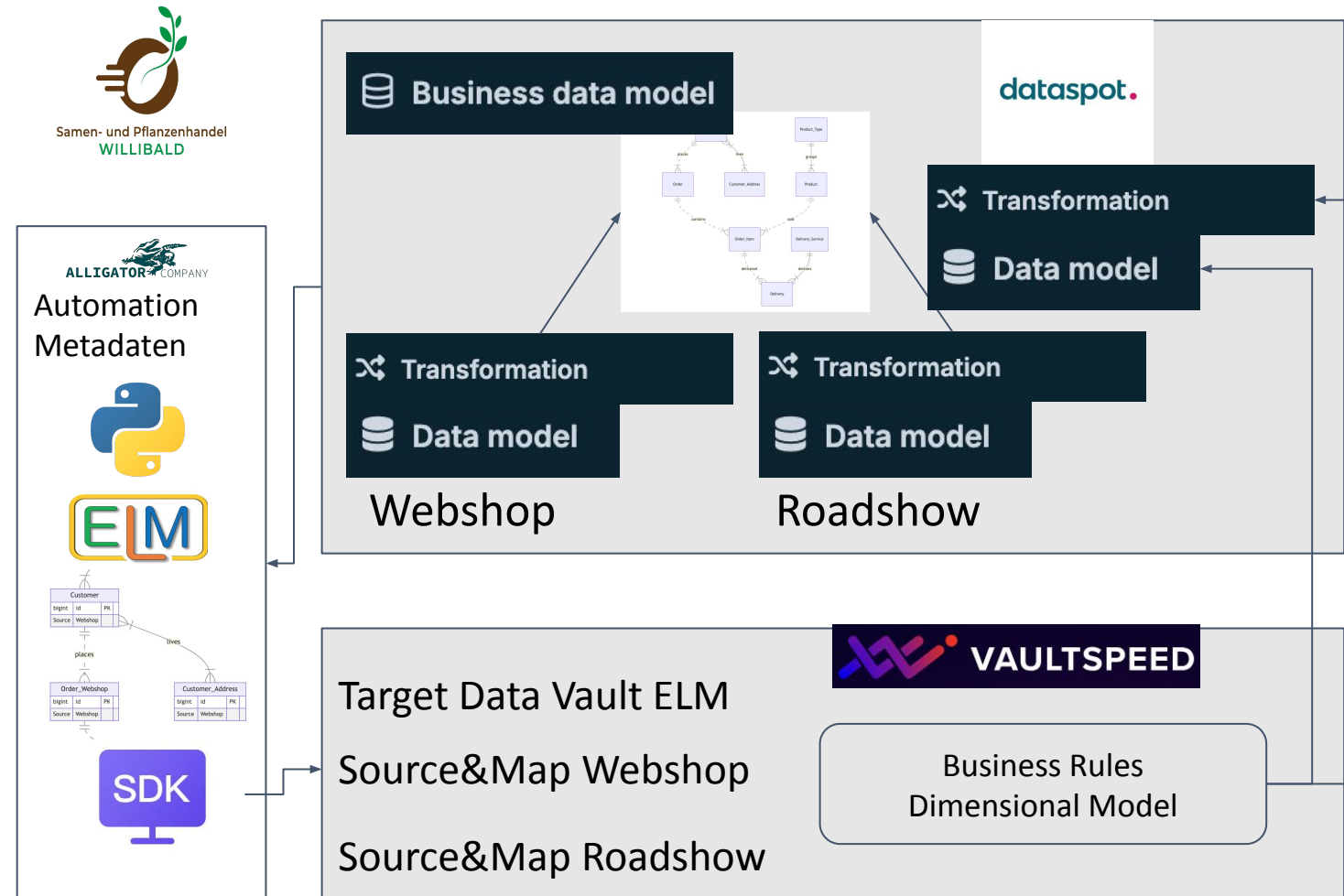
Herausforderungen bei der Einführung von Data Vault

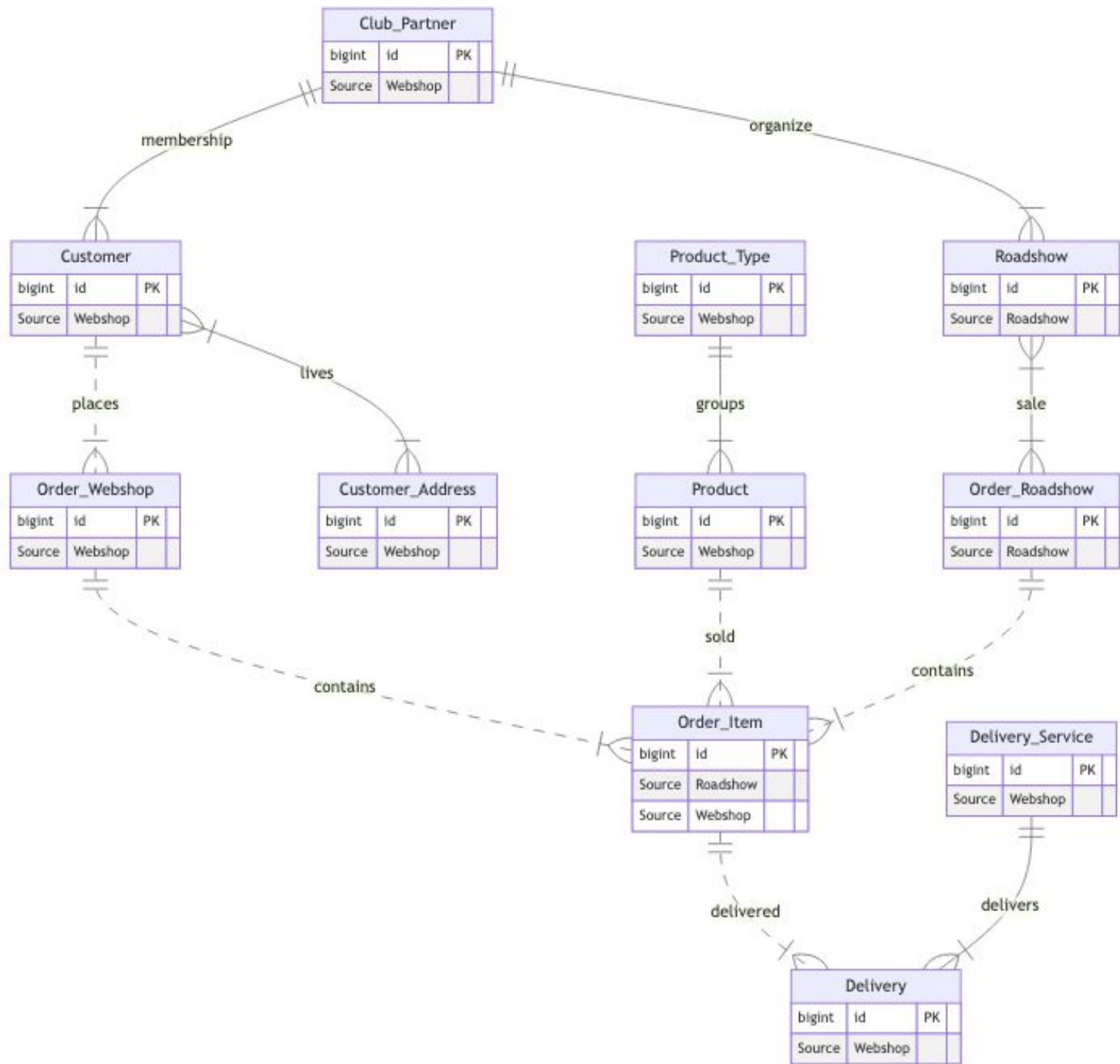


- Symptome für schlecht laufende Automatisierungen
 - sehr viele Hubs
 - komplexe Geschäftsregeln mit vielen Abhängigkeiten - zwischen Datavault und Report
 - Bezeichnungen der Hubs ergeben keinen Zusammenhang zum Geschäftsmodell des Kunden
 - Bezeichnungen aus dem dimensional Modell stimmen nicht mit den Bezeichnungen der Hubs überein
- Probleme
 - Laufzeiten für Entwicklung nehmen stark zu, da die Bedienung der Automation langsam wird
 - Laufzeiten der Beladung steigt an
 - Kosten im Cloud Data Warehouse steigen
- Ursachen
 - Anforderungen werden aus den Datenmodellen der Quellen abgeleitet
 - Lernkurve steil, benötigtes Know-How zur Modellierung wird unterschätzt
 - Fachbereich ist zu wenig involviert

Vorgehensweise - Think differently

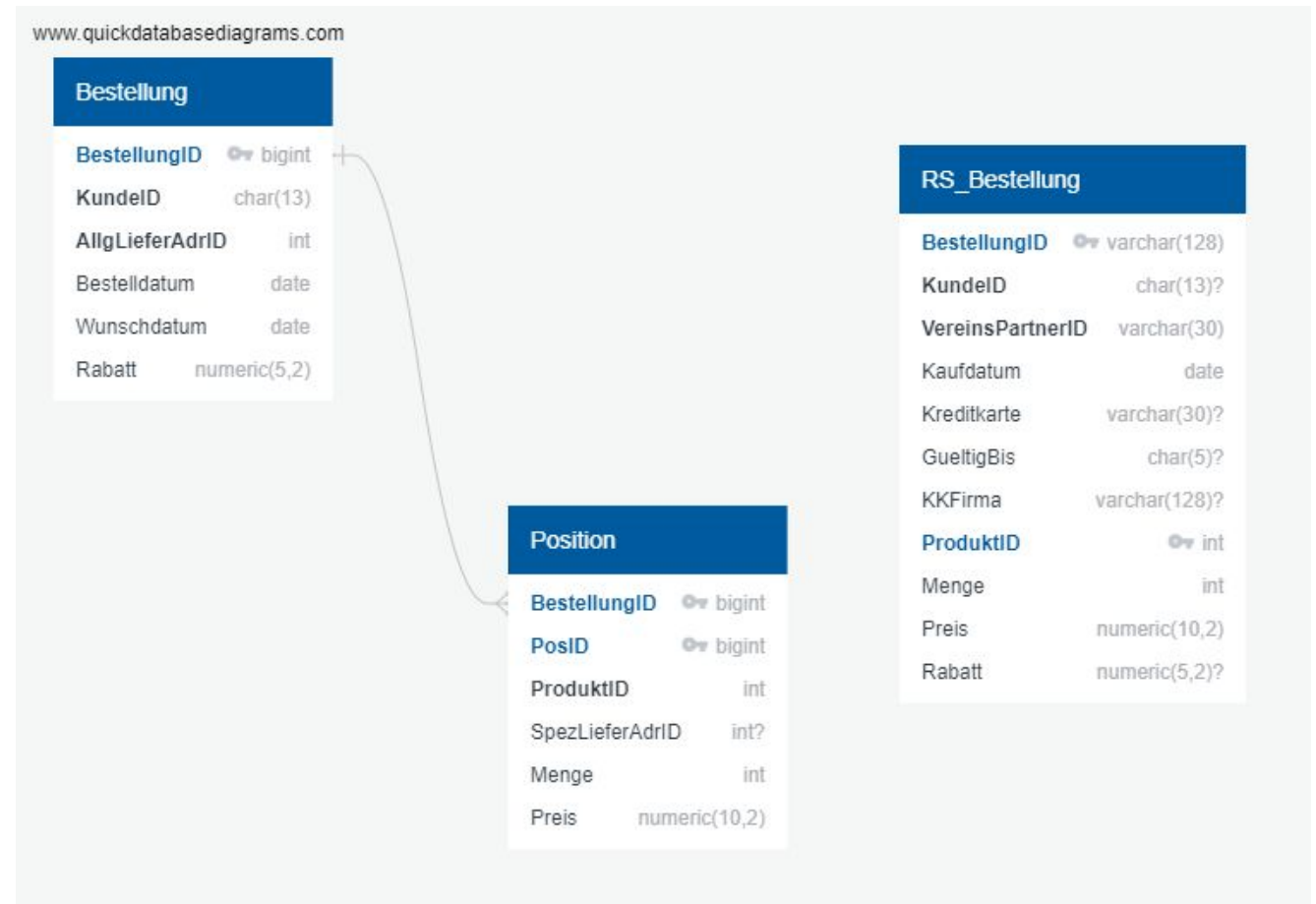
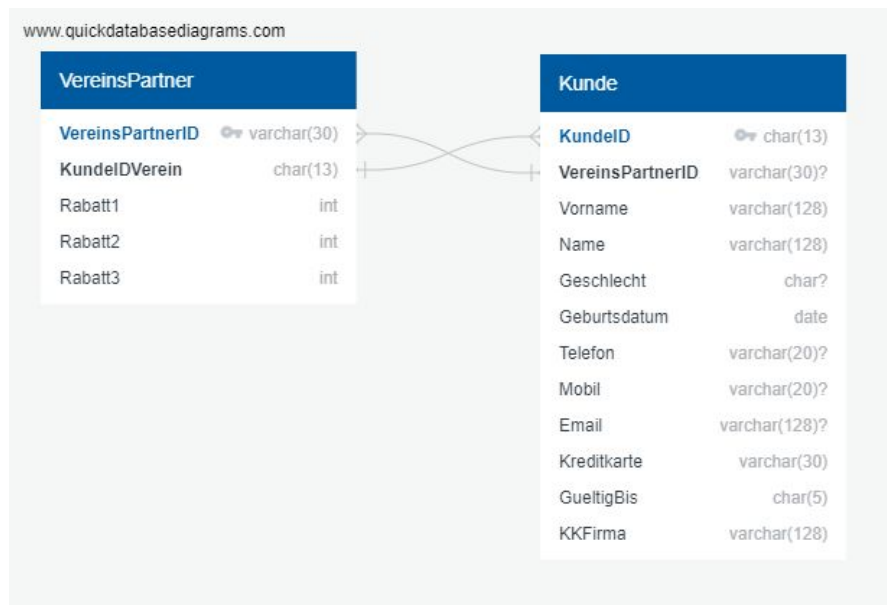
1. Conceptual Model / BTM
Align & Refine / ELM
2. Generate ELM / ER vom BTM
3. Generate Mappings der
Quell-Attribute und Transformationen
(Hardrules)
4. Apply ELM to Vaultspeer
5. Apply Mappings for all Source Data
Models
6. Deliver Target Schema & Business
Rules back to Dataspot



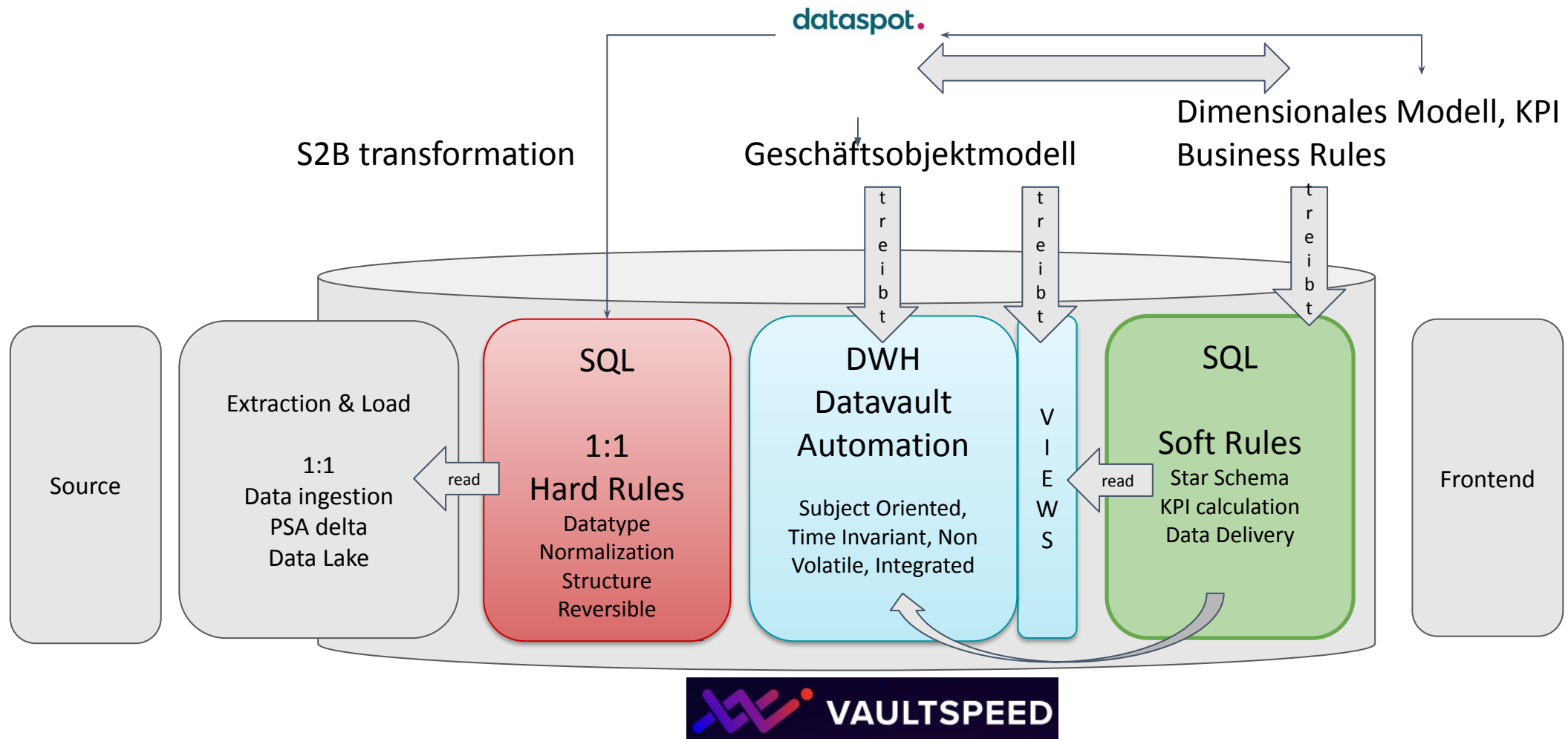


[Mermaid Live](#)

Frühe Integration - Was sind für Euch gültige Hardrules? Was ist mit Jedi?



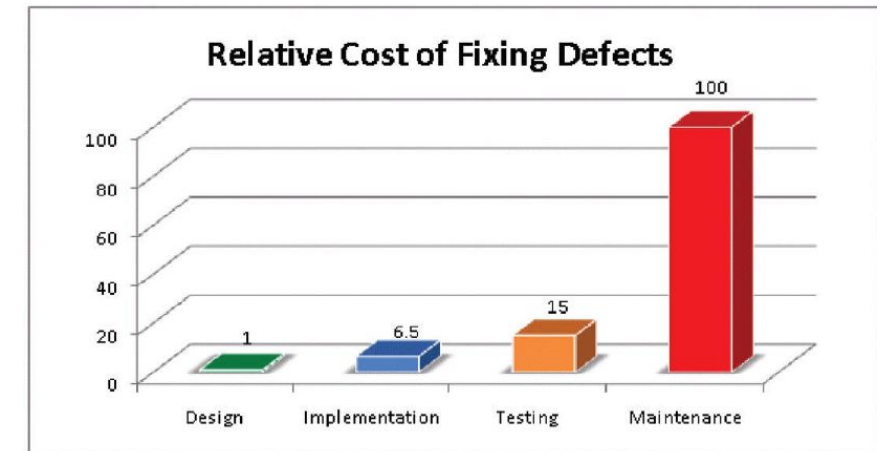
Datavault / DWA - Automation der Automation



Welche Verbesserungen erhoffen wir uns durch modellgetriebene Automation ?



- Fachbereiche direkt beteiligen
 - Verantwortlichkeit übernehmen
 - gemeinsames Lernen
 - neue Erkenntnisse durch Kommunikation
 - Fachbereiche verstehen das Modell im Datawarehouse
 - **Vertrauen**
 - **Self-Service**
- Divide & Conquer
 - Hardrules / DV mapping / Softrules
 - nur vom Fachbereich definierte CBC/Hub implementieren
 - Einzelschritte werden einfacher, Logik aufgeteilt (Automation übernimmt Koordination bei Beladung und Maintenance)
 - **geringere Laufzeiten verringern die Kosten im Cloud Data Warehouse**
- MDA / modellgetriebene Implementierung für Data Vault
 - Frühes finden und beheben von Fehlern
 - Standardisierung der Metadaten / Herstellerunabhängigkeit
 - Reduktion des Quellcodes (Python/SQL)
 - **Erhöhte Wartbarkeit**
 - **schnellere Entwicklungszyklen**



Danke!



- Webinar?
- Workshop auf der Frühjahrstagung?



Prompt: generiere ein witziges
abschlussbild für meine schlussfolie
in einer präsentation mit dem motto
"ende gut alles gut"