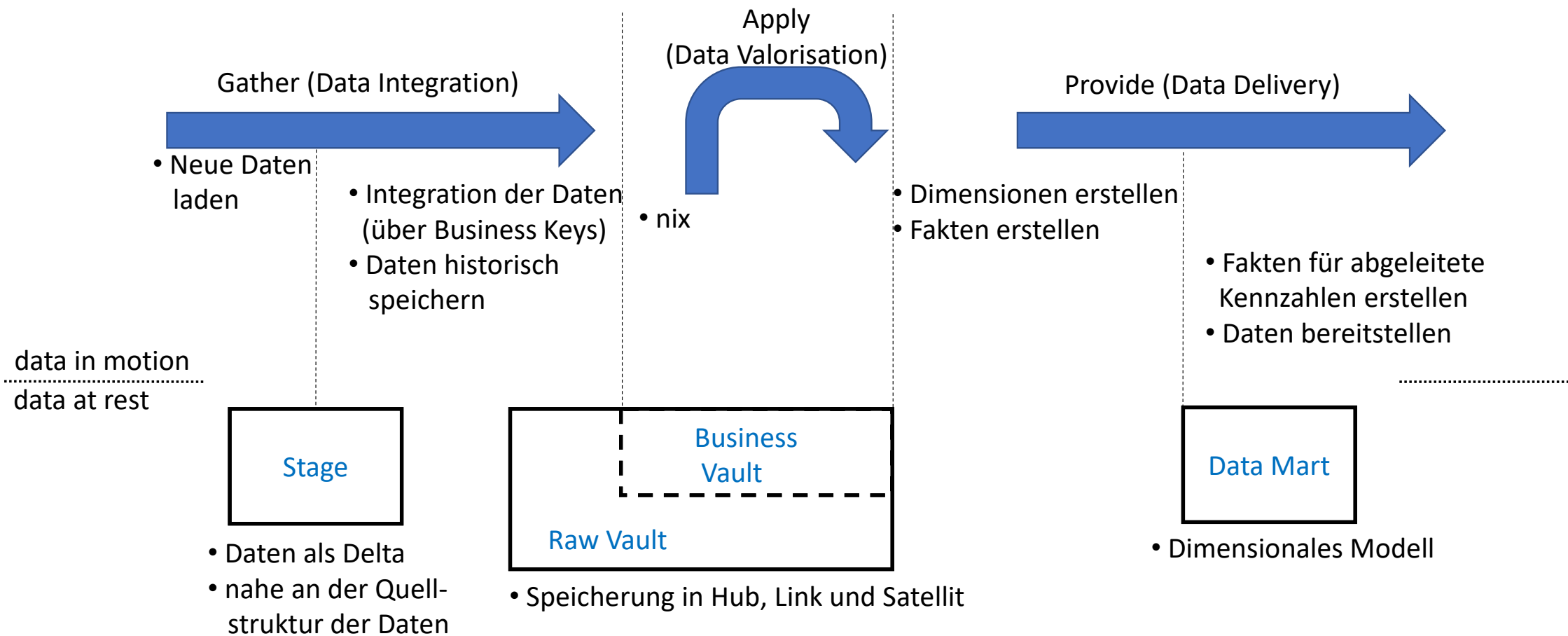


Datenmodellierung, Datenmanagement und das Data Warehouse

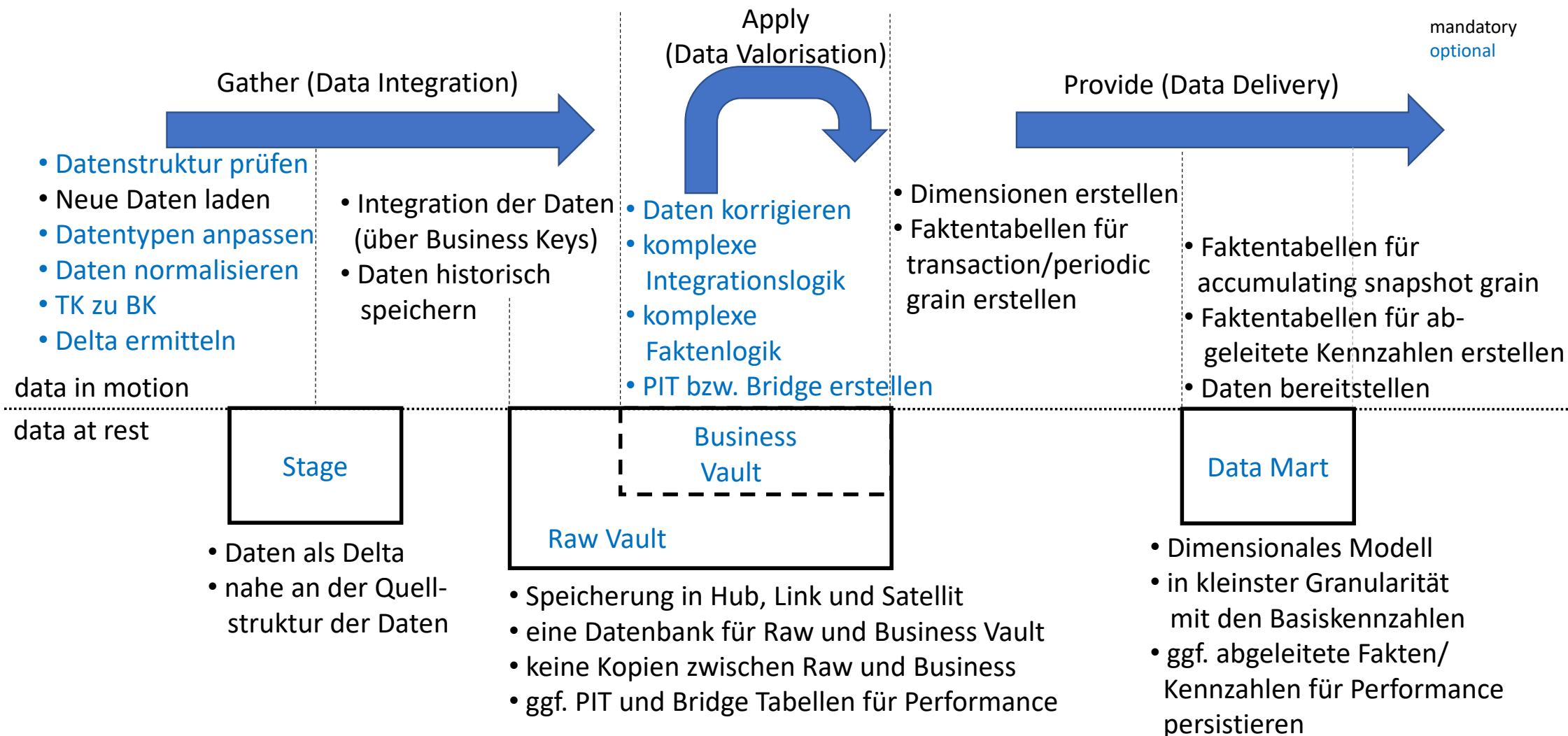
DDVUG Online am 27. Januar 2022

Architektur Datenintegration

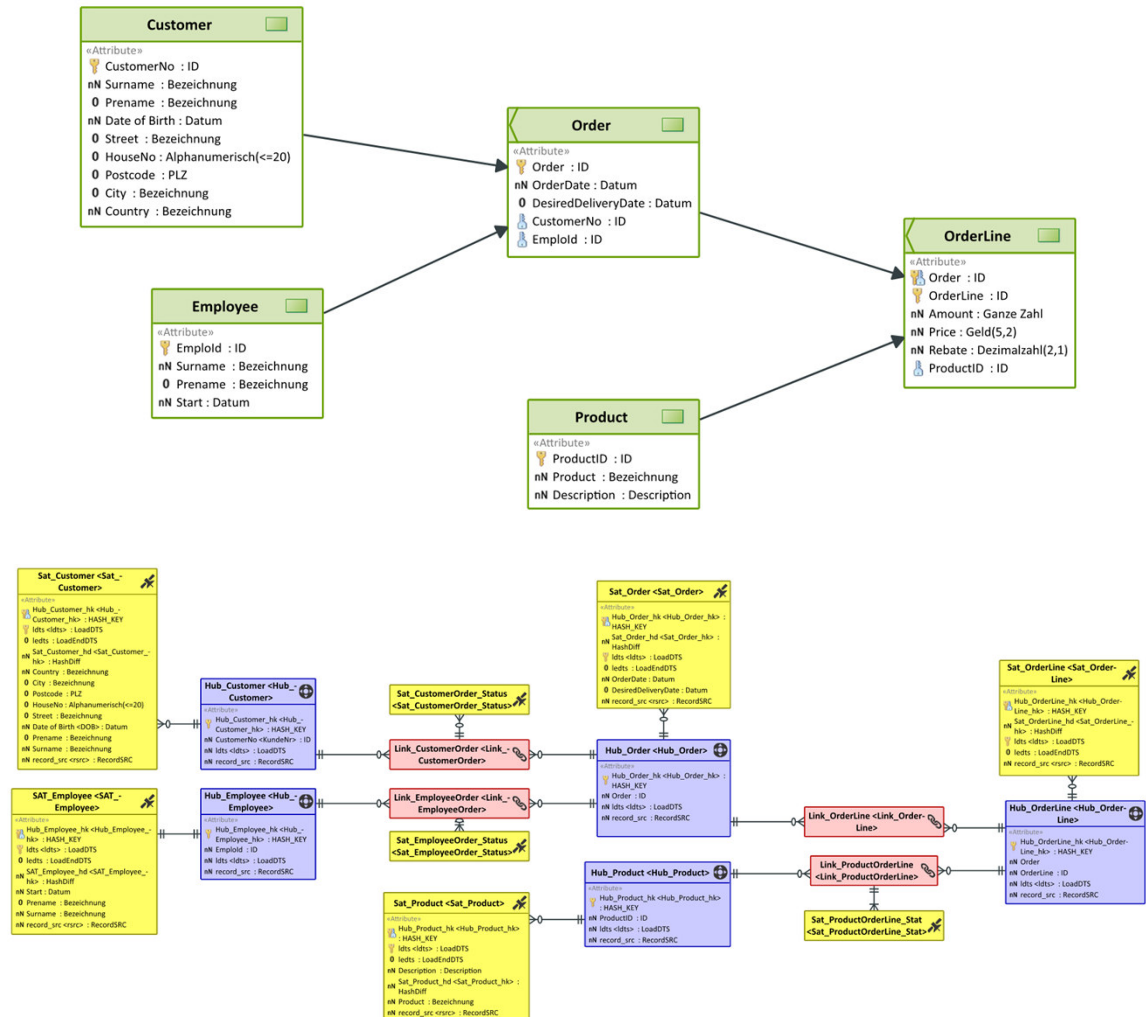
– eigentlich ganz einfach



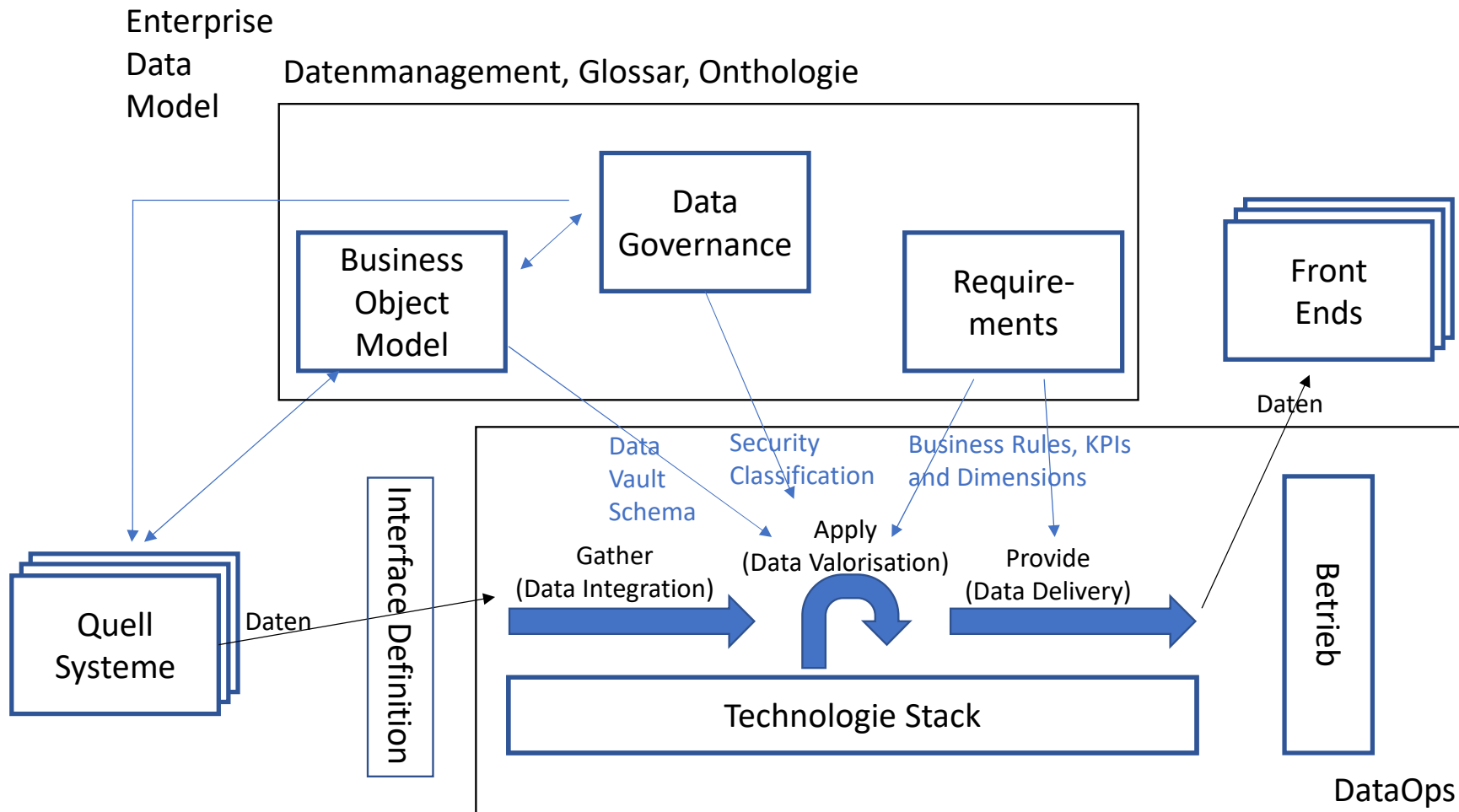
Die Realität der Datenintegration



Ein
Geschäftsobjekt
modell ist ideal
für die
Kommunikation
über die Daten im
Core Warehouse



Über Metadaten das DWH fachlich Steuern



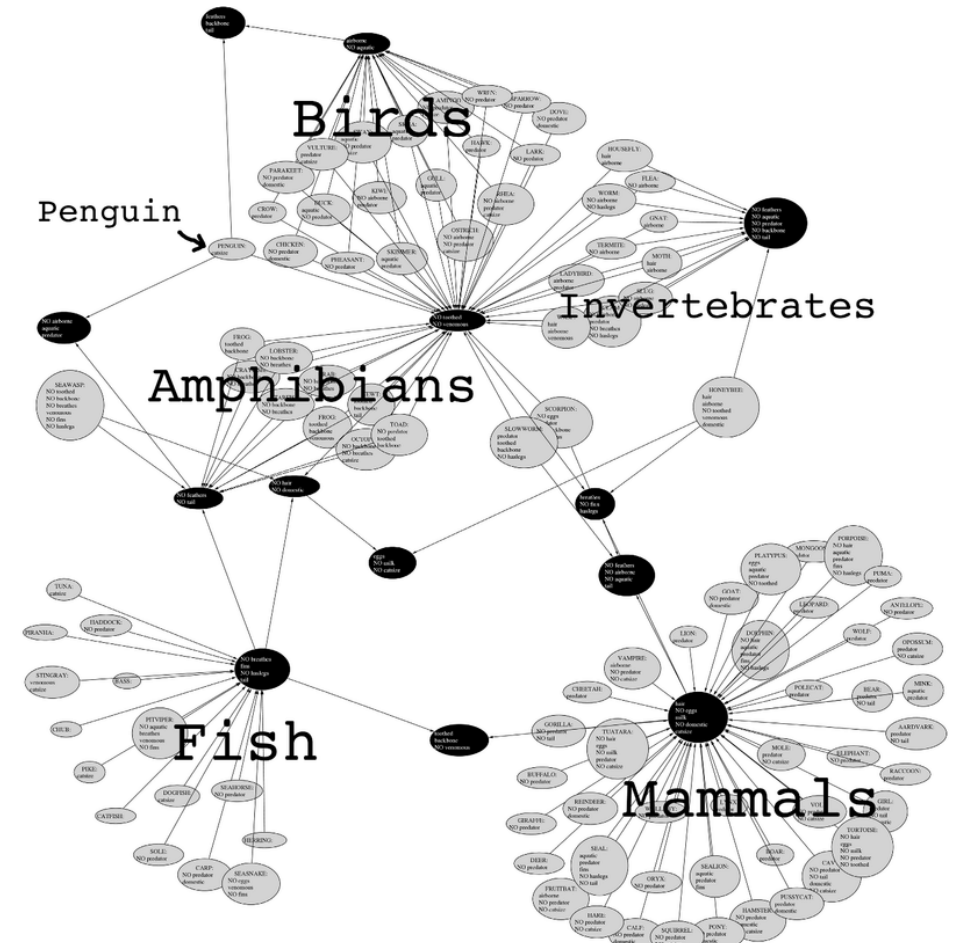
Warum macht es dann niemand?

- Ein rein fachlicher Ansatz ist selten
- Es ist mehr Methode als Tool
- Und ist oft mit einzelnen Personen verknüpft



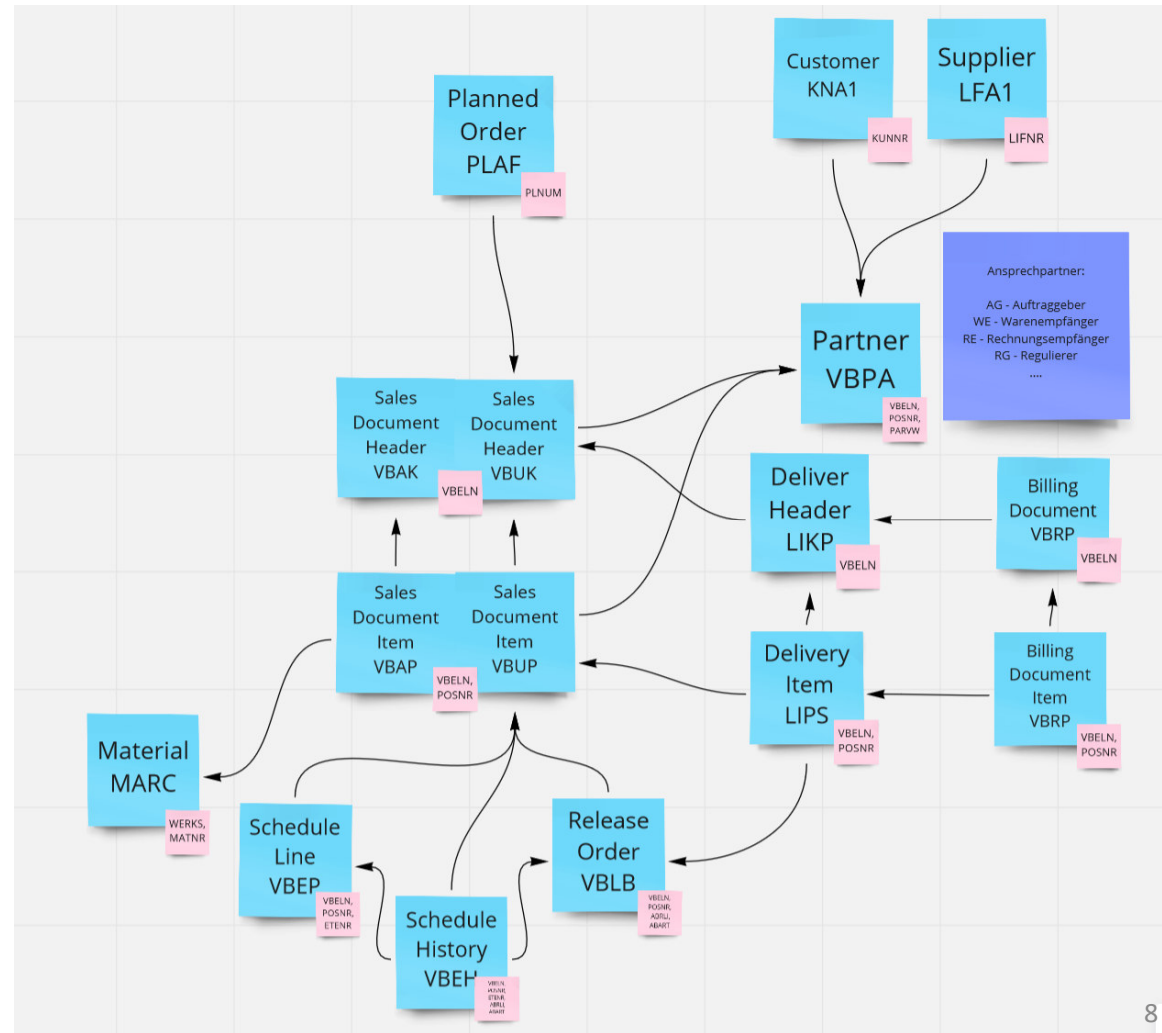
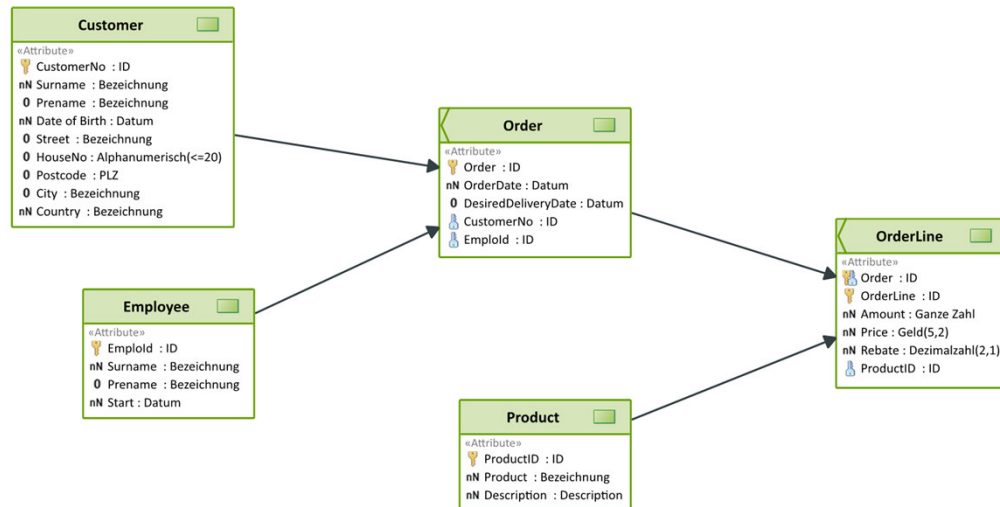
Datenmodelle beschreiben die Welt

- Einigung auf gemeinsame Begriffe, um zusammenarbeiten zu können
- Ontologien als wesentlicher Baustein moderner Wissenschaft
- Von Aristoteles über Baruch Spinoza bis hin zu Ludwig Wittgenstein
- Davon für die Datenmodelle übernommen:
 - Univers of Discourse
 - ein Begriff, der alle glücklich macht



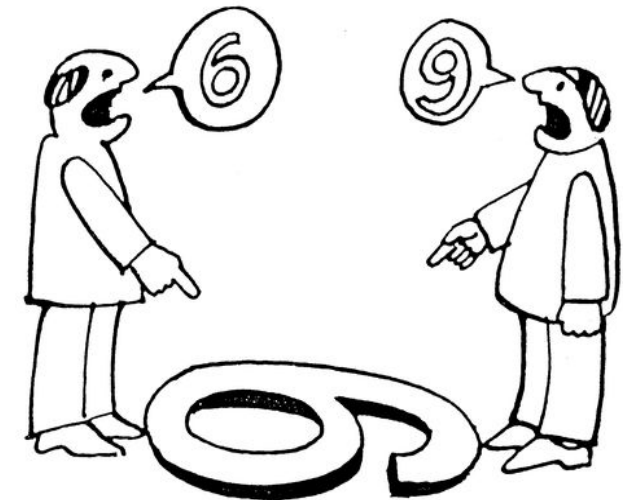
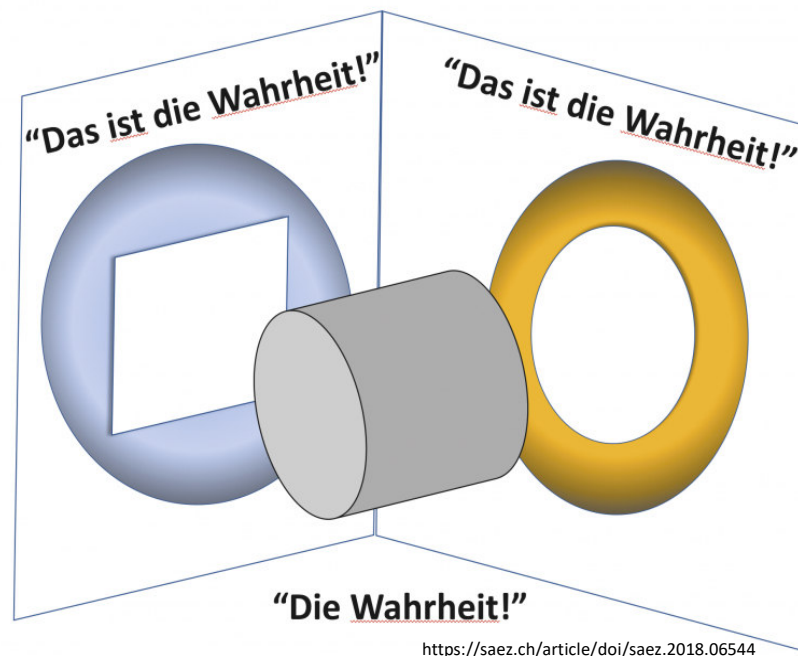
The Cruncher: Automatic Concept Formation Using Minimum Description Length - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-Automatically-Created-Zoo-Ontology-Arrows-represent-inheritance-and-the-attributes_fig3_262405011 [accessed 26 Jan, 2022]

Das Problem der Granularität



Und dann noch: Platos Höhle

- wir können nur beschreiben was wir sehen
- aber sehen wir auch das ganze Bild?



<https://leveragepoints.org/2017/10/31/alles-eine-frage-der-perspektive/>

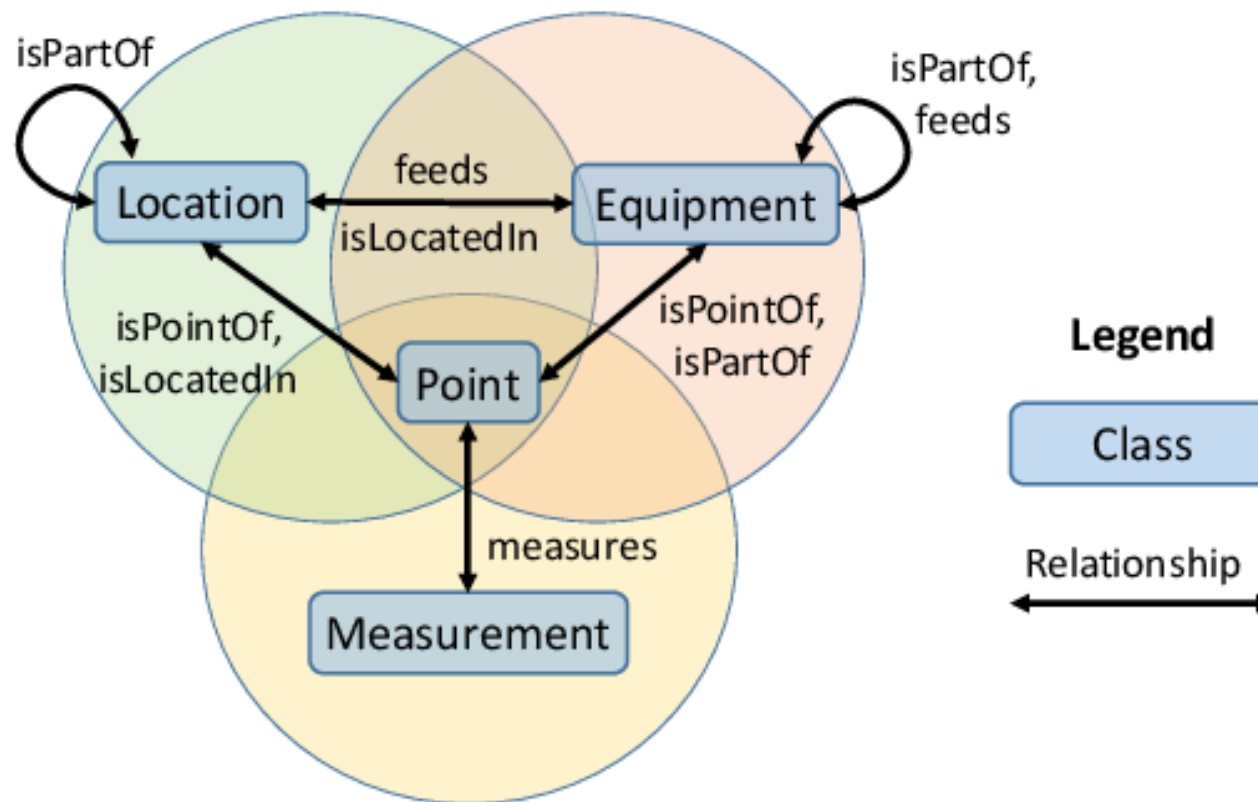
„Es ist eine Selbstillusionierung sinnkonstituierender Systeme, wenn sie meinen, zeitüberdauernde Identitäten habe es immer schon gegeben und werde es weiterhin geben, und man könne sich daher auf sie wie auf Vorhandenes beziehen. Alle Orientierung ist Konstruktion, ist von Moment zu Moment reaktualisierte Unterscheidung.“

– Niklas Luhmann: Die Gesellschaft der Gesellschaft Seite 44/45

Warum dann die eine Gemeinsamkeit suchen?
Was wenn wir auch die Unterschiede modellieren?

brickschema.org

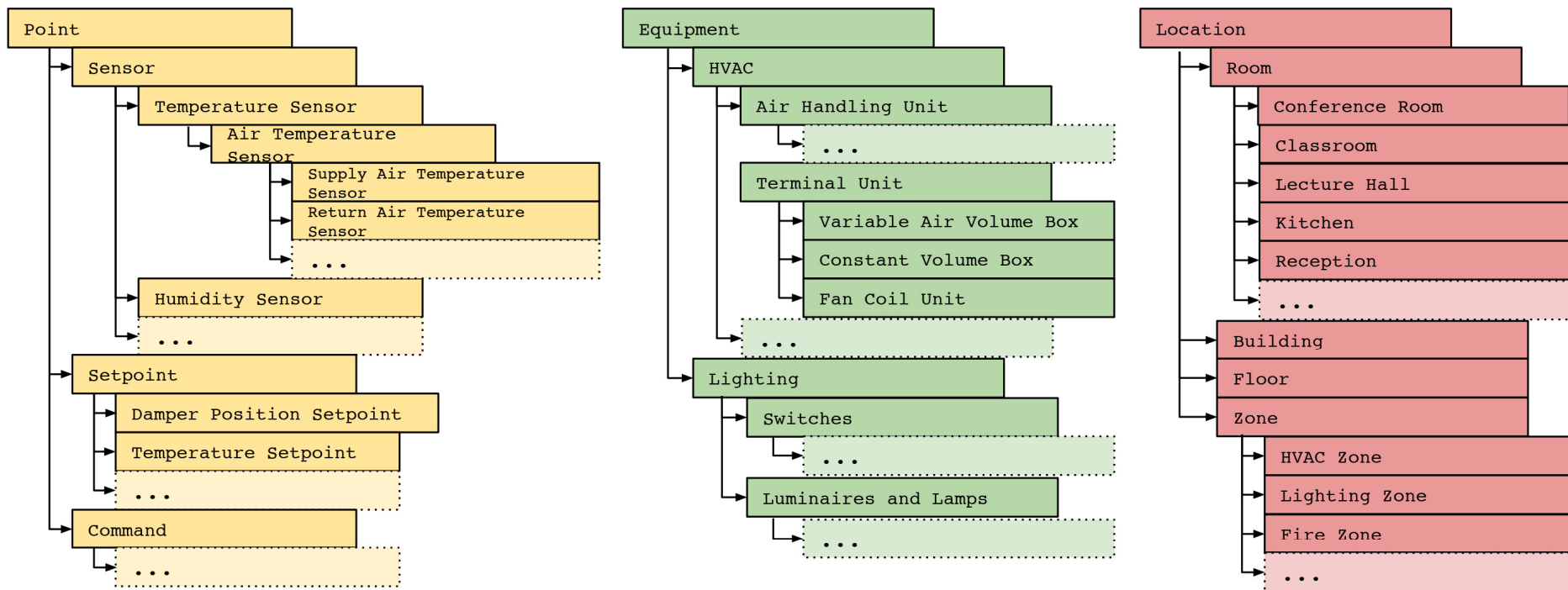
Eine Ontologie für Gebäudemanagement



brickschema.org

An Ontology for Building Management

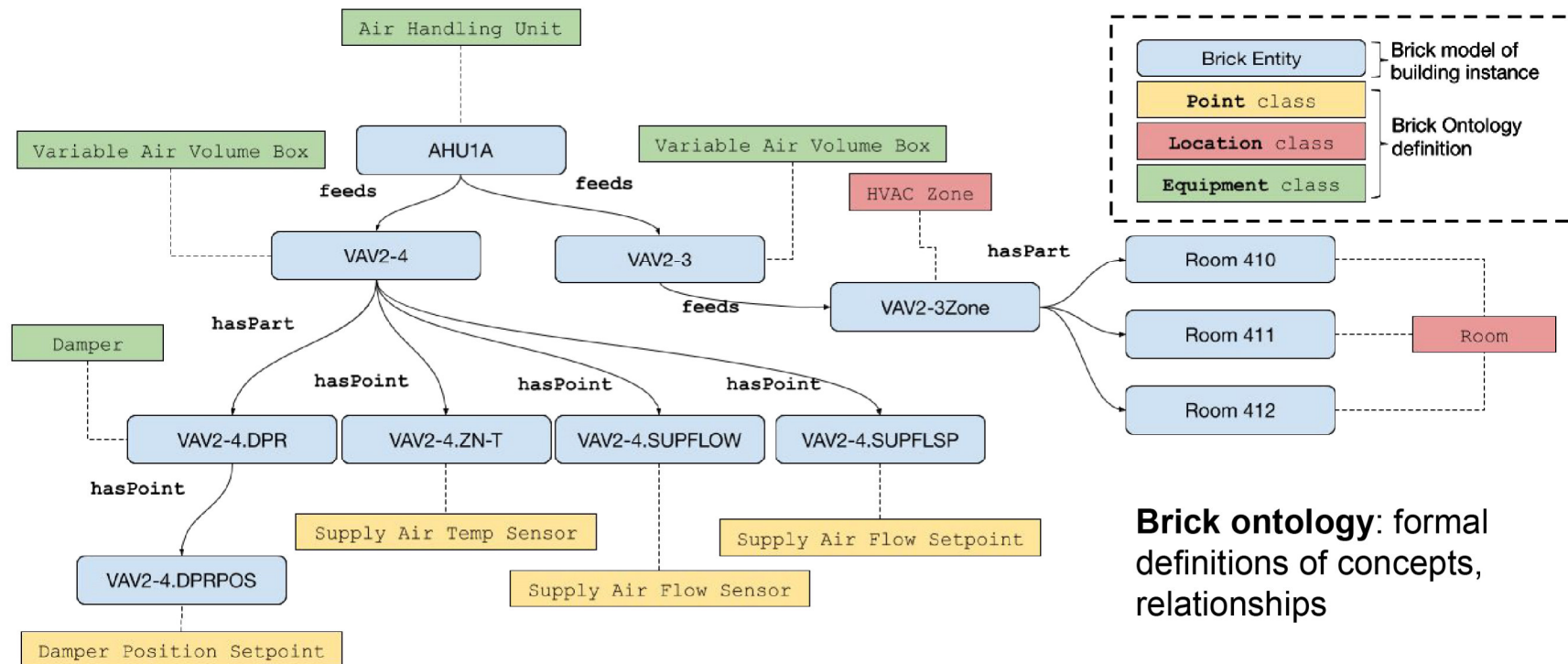
Brick Classes



Brick

<https://brickschema.org>

A Brick Model in a Graph Database



Fachliche
Unterschiede und
Gemeinsamkeiten
modellieren

Klare Geschäftsobjekte

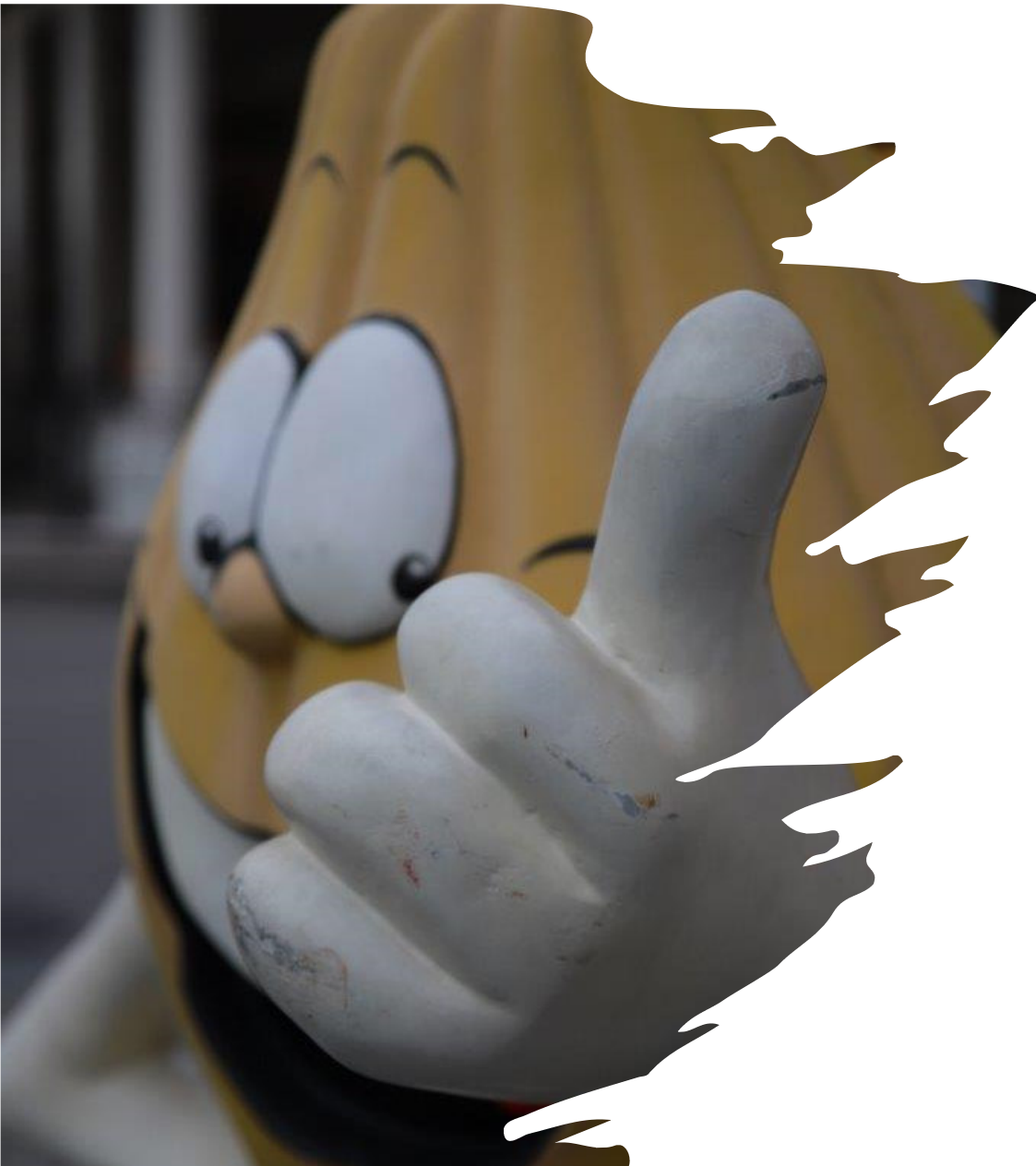
Eindeutige und durchgängige
Schlüssel

Hierarchien für sich einander
ausschließende Ausprägungen

Beziehungen für kumulative Rollen

Technische Unterschiede in der Sichtweise





Einigung einfach: Konsent finden

- der Geist dahinter ist ‚Damit kann ich leben‘
- Irrelevant ob der Unterschied fachlich oder organisatorisch ist

Panta Rhei! Alles fließt!

Über die Zeit ändern sich
Ansichten und Konsent

Heraklit:
„Du kannst niemals zweimal in
denselben Fluss steigen.“





Lineage

Das Geheimnis agiler Datenmodelle

Datenschnittstellen festlegen mit Inhalt, Quelle und Ziel

Beobachtung:
Datenmodelle
sind sehr oft
veraltet

entstehen für Verständnis

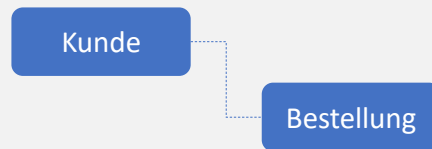
oder zur Dokumentation

reicht nicht als Nutzen

werden nicht aktuell gehalten

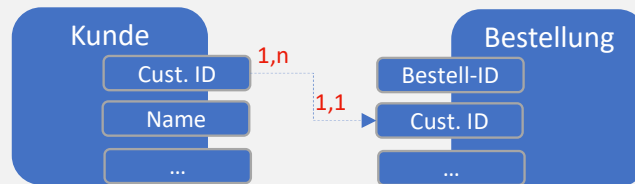
Datenmodelle werden schnell umfangreich

Konzeptionell Datenmodell



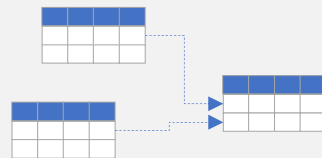
Business-Objekte und ihre Verknüpfungen
Nur wenn hilfreich: IDs und Attribute

Logisches Datenmodell



Detaillierte Beschreibung der Entitäten,
ihrer IDs und Attribute sowie der
Beziehungen zwischen Entitäten

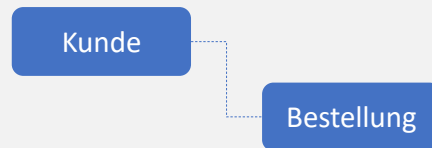
Physikalisches Datenmodell



Physikalische Repräsentation von Daten
(Felder, Datentypen, Schlüssel ...) in Bezug
auf einen bestimmten Datenbanktyp (z. B.
relational, Graph, Dokument ...)

Datenmodelle: die vereinfachte Sicht

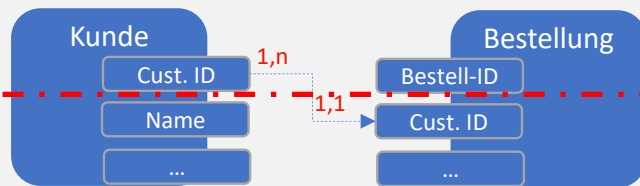
Konzeptionell Datenmodell



Umfangreiche Beschreibung
Business-Objekte und ihre Verknüpfungen
Nur wenn hilfreich: IDs und Attribute

Einmalig im Unternehmen
Mit Gemeinsamkeiten *und*
Unterschieden

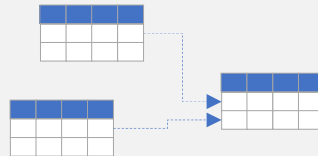
Logisches Datenmodell



Semantik

System

Physikalisches Datenmodell

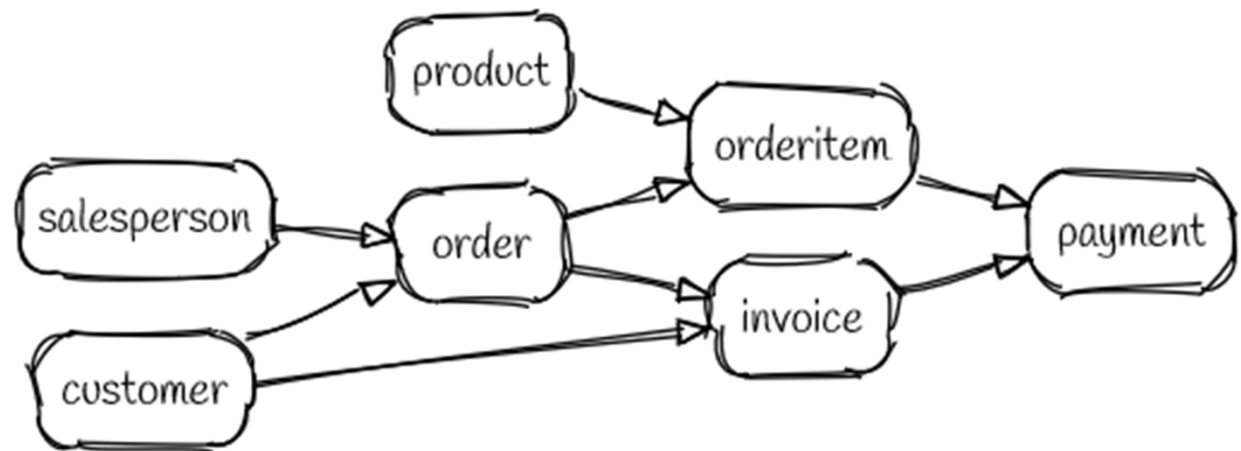


Scannen der Daten
-> Data Catalog

Pro System

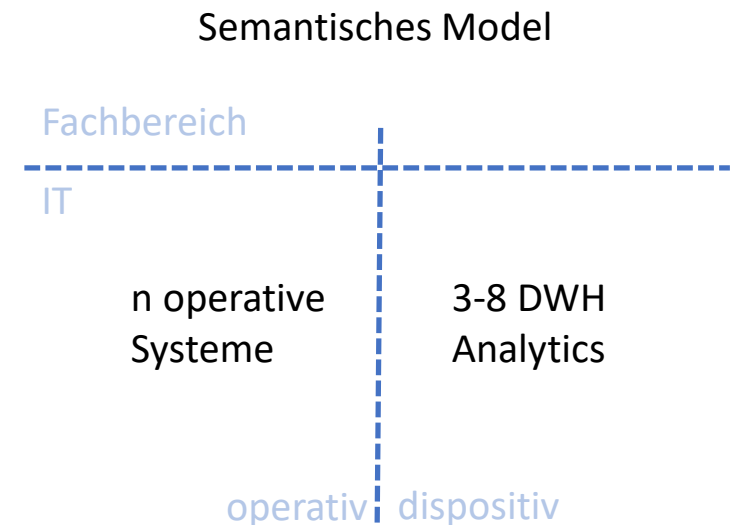
Um Daten
verwalten zu
können,
müssen alle
Beteiligten
wissen,
welche Daten
verwaltet
werden

- Ein Unternehmensdatenmodell ist viel zu groß, um es in einem Zug zu verstehen
- Abstraktion ist der Schlüssel
 - ein konzeptionelles Modell verwenden
 - bestehend aus zentralen Geschäftskonzepten und deren Beziehungen
 - nur einen relevanten Teil des Modells betrachten
 - Geschäftsbegriffe verwenden
 - das Modell in einer lesbaren Form zu formatieren



Viele Rollen und Systeme erzeugen ein Abgleichsproblem

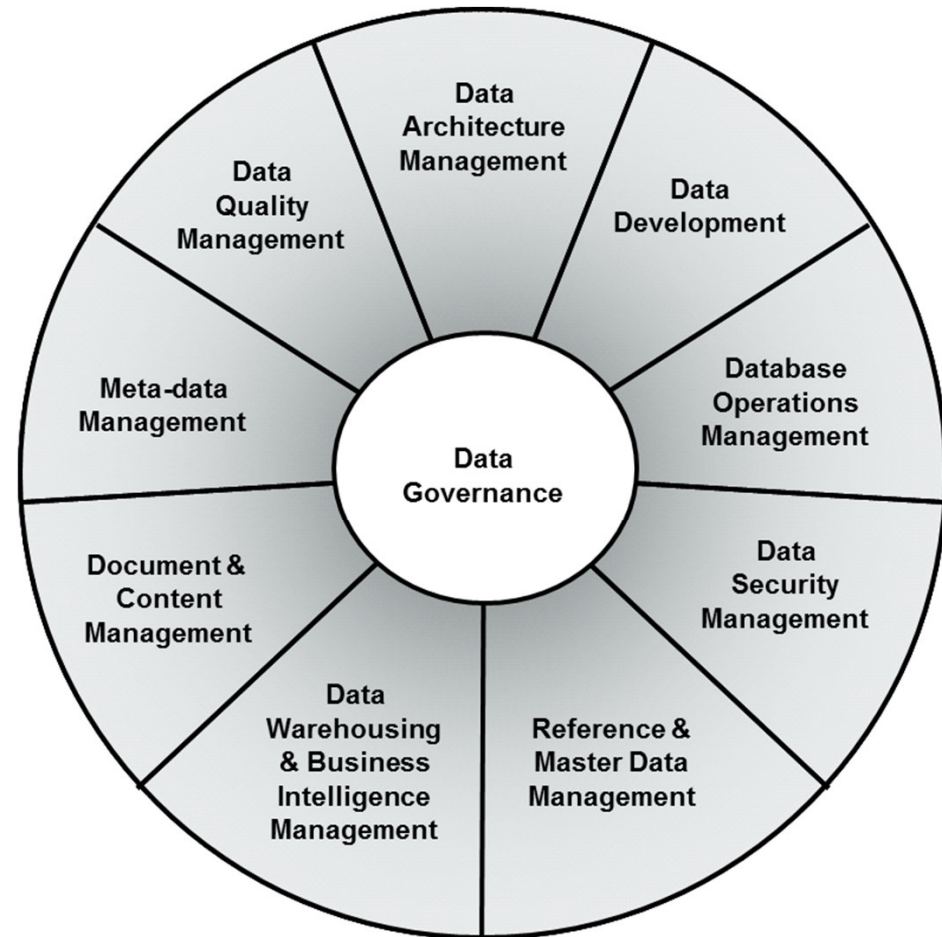
- Wasserfall nicht mehr denkbar
- Automatisch bzw. maschinell gestützt
- Ein beständiger Abgleich anhand von Metadaten



DAMA hat Datenmanagement definiert

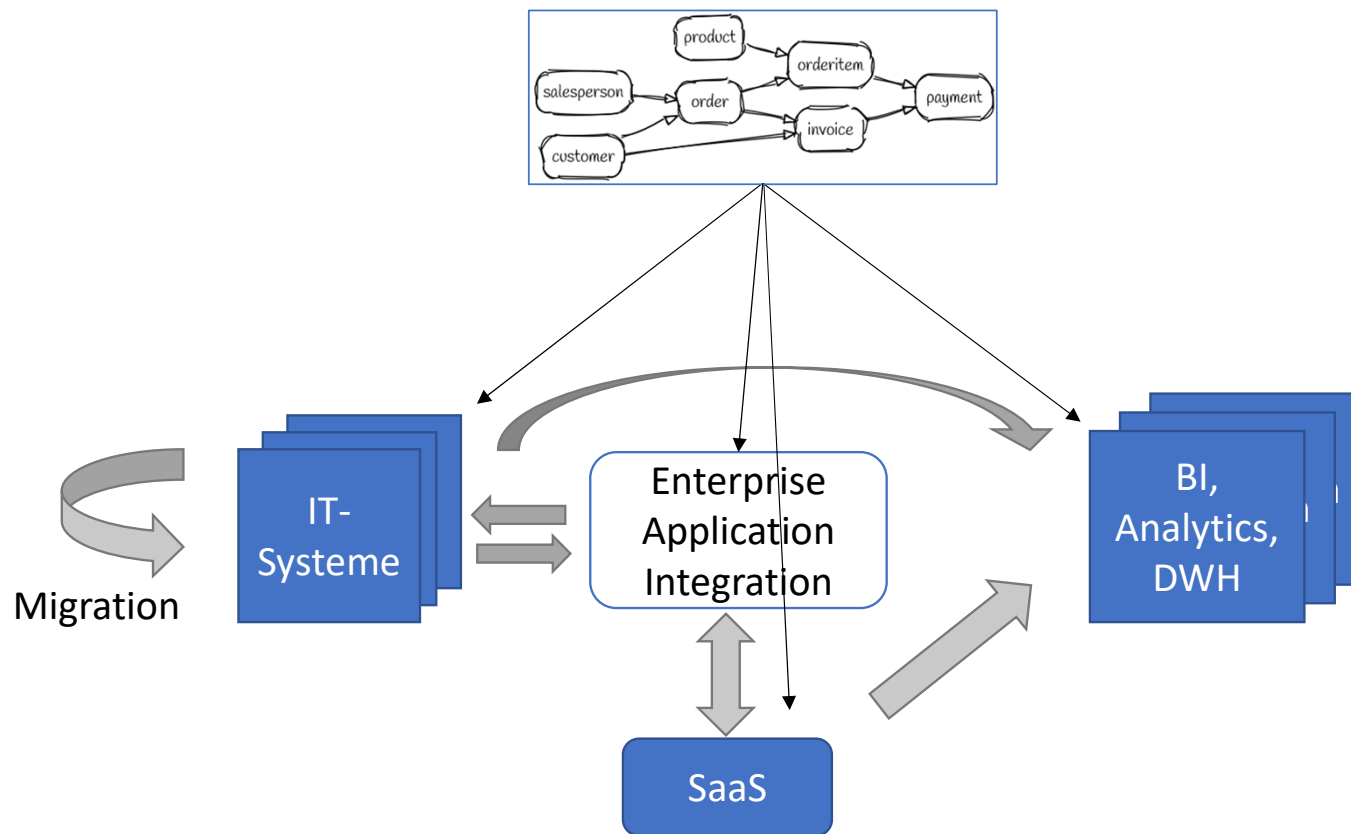
Leider nur als Liste
von Best Practices

Haben keinen Ablauf
beschrieben,
sondern lediglich das
Bedürfnis formuliert.



Copyright © by DAMA International

Das konzeptionelle Modell als Treiber



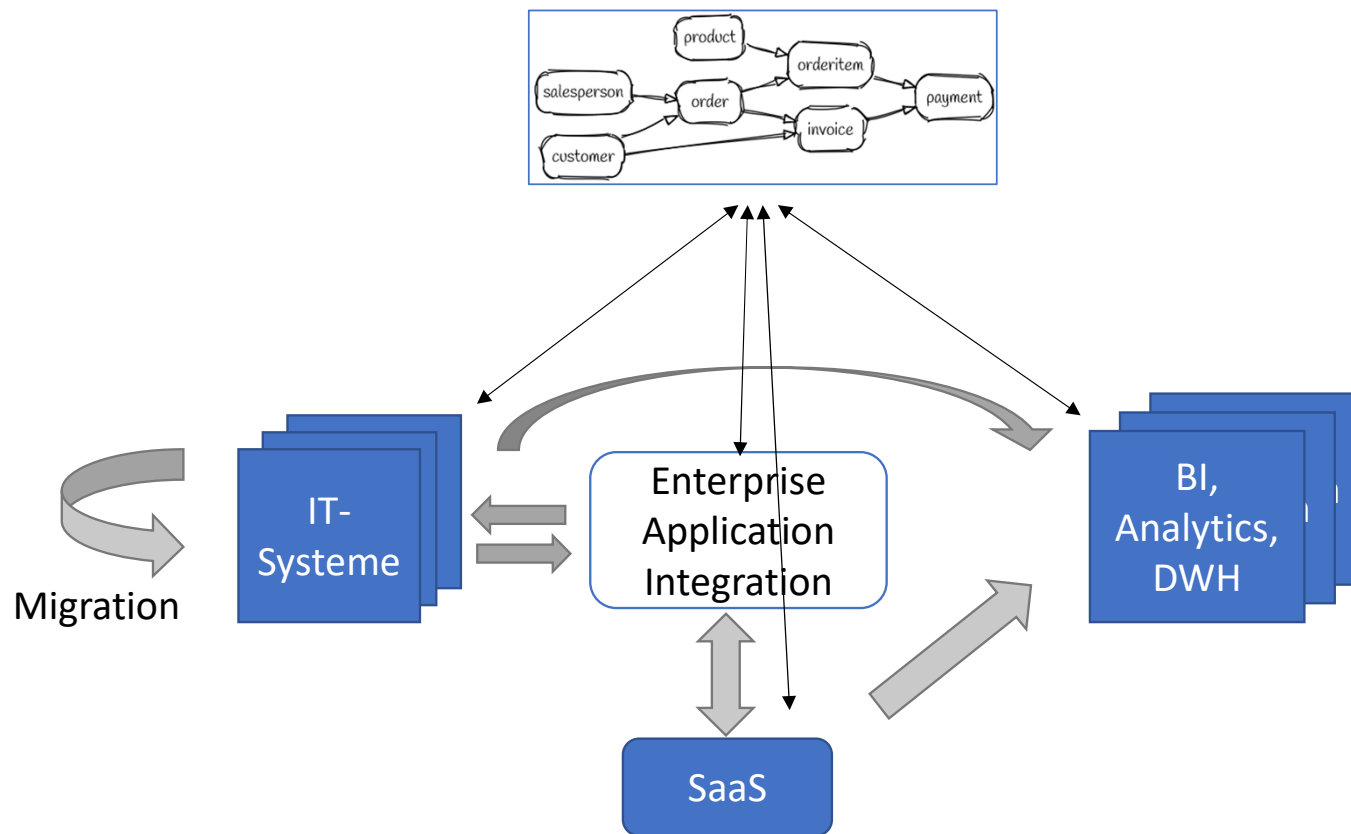
Diese Metadaten (Daten über Daten) beantworten Sie die folgenden Fragen:

- Wo finde ich bestimmte Daten?
- Wie werden die Daten systemübergreifend kopiert? (Abstammung der Daten)
- Wo liegt das führende System?
- Angaben zu Datenschutz und -sicherheit

Die Daten zwischen Systemen

- überschneiden sich, trennen sich oder sind identisch?
- haben gleiche oder unterschiedliche Schlüssel?
- mit dem gleichen oder einem unterschiedlichen Satz an Attributen?

Die Verknüpfung zum konzeptionellen Modell ist keine Einbahnstraße

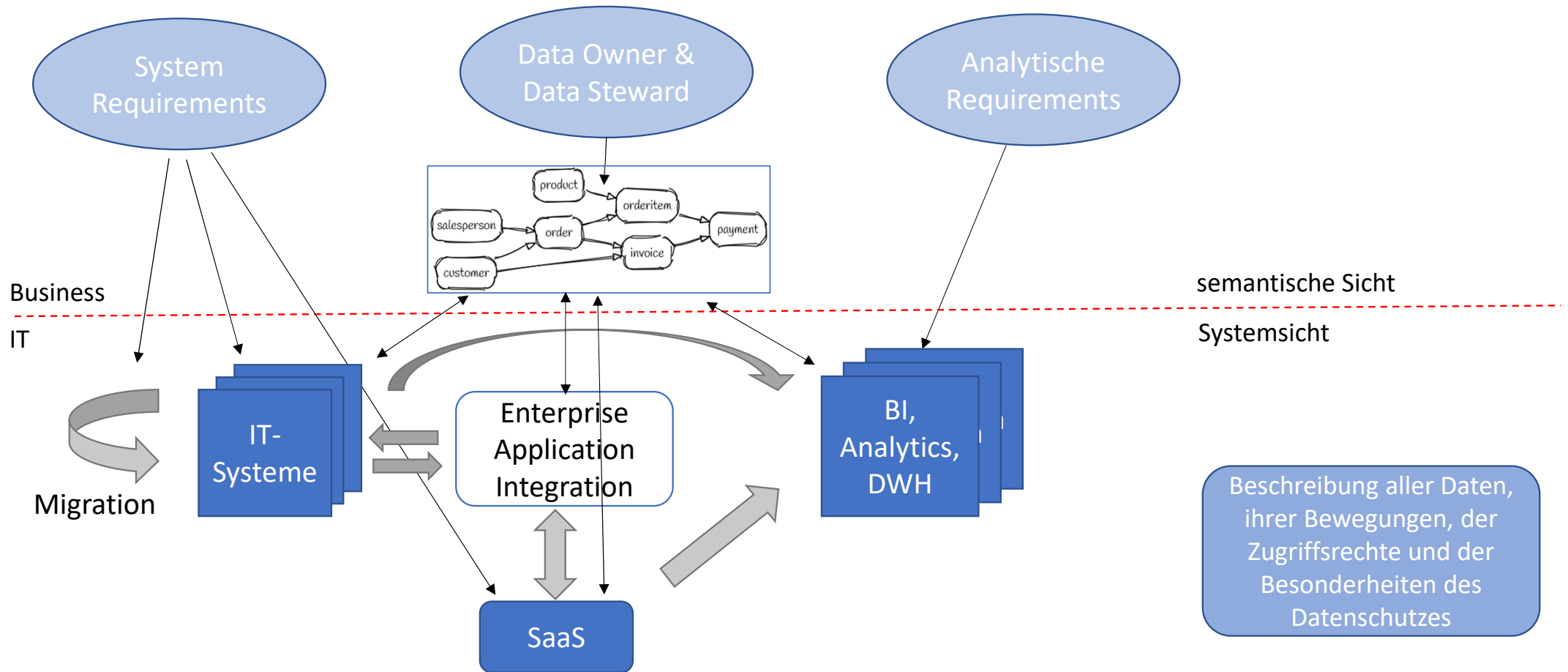


Alle Entitäten in all diesen Datenmodellen gehören immer zu einem Core Business Concept.

Damit lassen sich gleiche Daten über Systeme hinweg identifizieren.

Somit lassen sich parallele Entwicklungen über die Metadaten verfolgen und abgleichen.

Ein sich abgleichendes System



Ein solche System ist keine Utopie

Viele neue Werkzeuge für Datenmanagement, Datensicherheit, Data Governance, etc.

- Colibra, Erwin, Informatica, Alation, atacama, data.world
- mit BI Fokus: dataspot, synabi

Die Lösungen haben eine hohe Varianz in der Funktionalität.

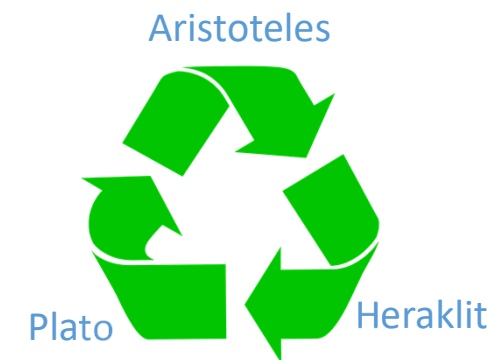
Und es entstehen immer noch weitere.

Eine vollständige Unterstützung eine fachlich getriebenen BI fehlt noch.



Es braucht zudem Methodik und Kompetenz

- Datenmodellierung ist Abstimmung
- Die Basis zur Kommunikation
- Als Vereinbarung zur Zusammenarbeit
- Unterschiede würdigen und modellieren
- Andere Fachlichkeit respektieren
- Konsent finden



Danke für Ihre Aufmerksamkeit.

