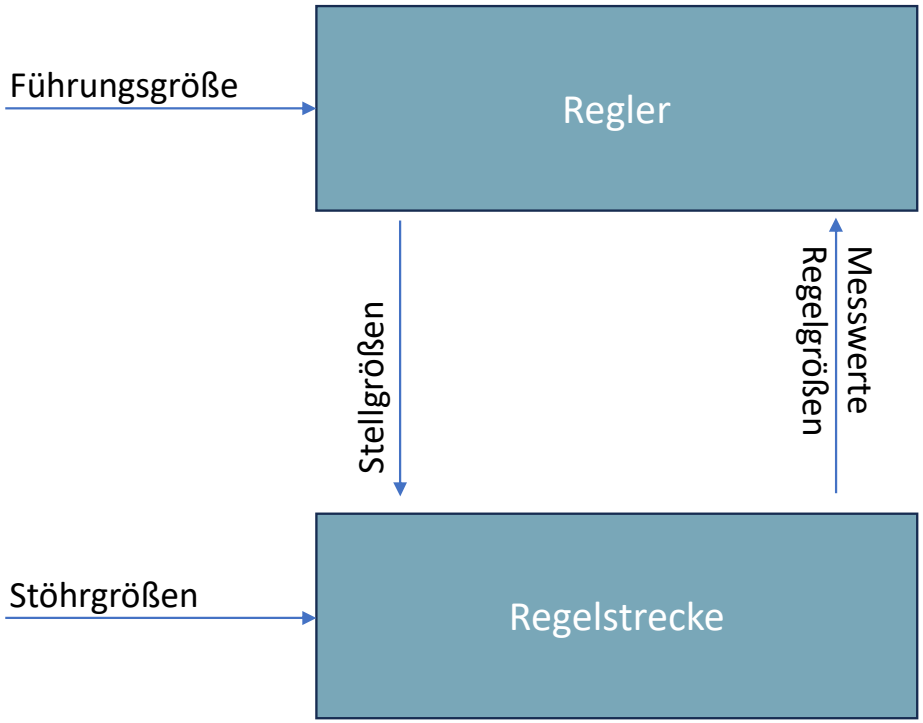
m[ethod] 2 data

Das gemeinsame Verständnis über die eigenen Daten

19. Tagung der DDVUG am 1. Oktober 2026 in Köln

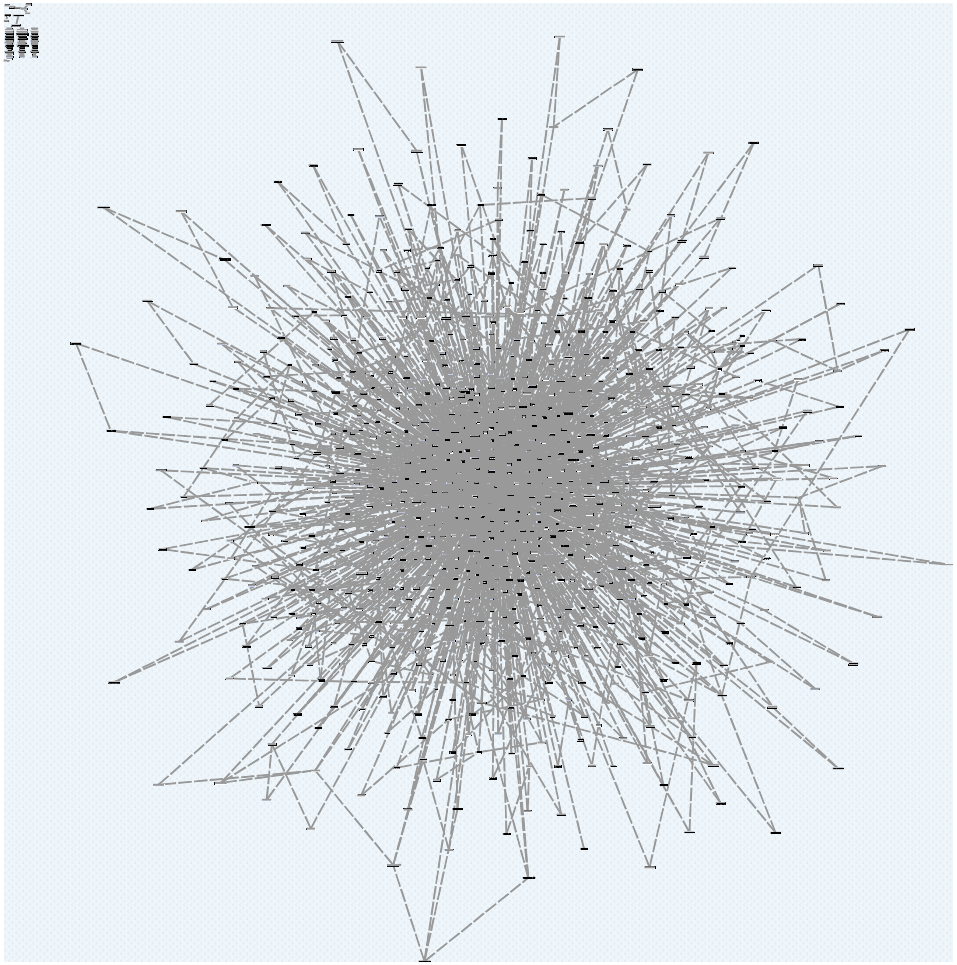
Warum modellieren wir Daten?

- Messen
- Steuern
- Regeln



- Beschreibt nicht den Einfluss des Beobachters
- Je komplexer es wird, desto mehr indirekte Einflussfaktoren
- Und dennoch hilft es, wenn man die Wertschöpfung klar benennen kann.

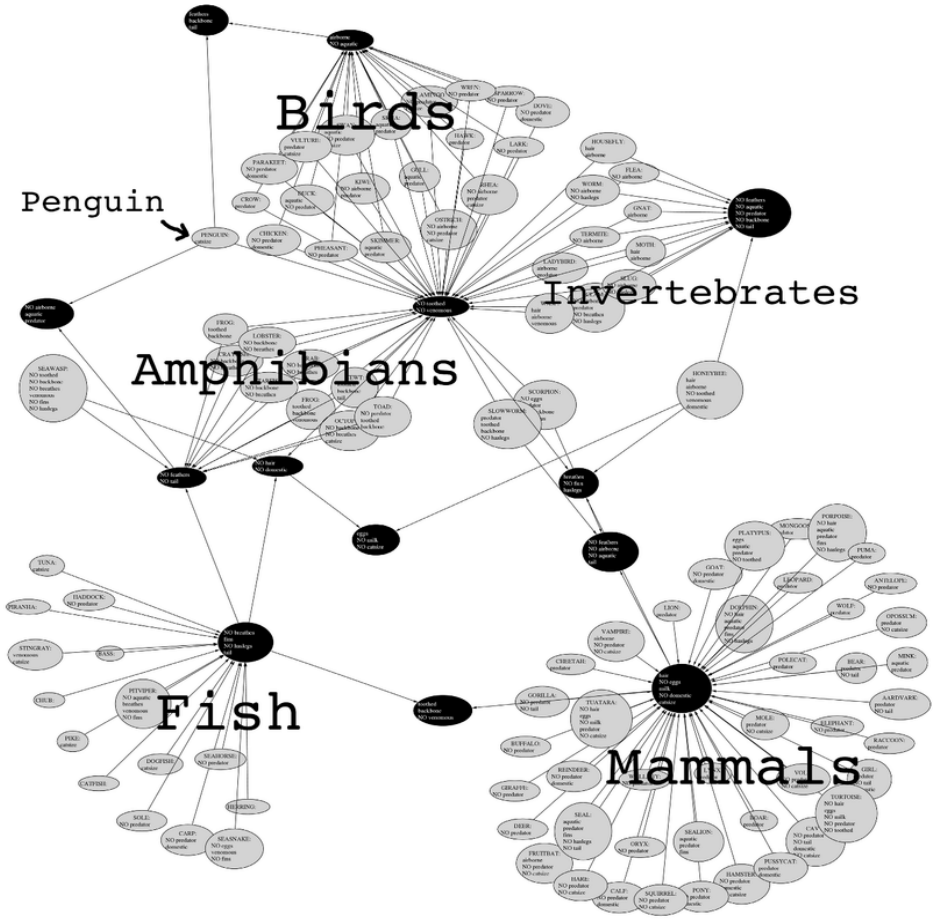
Warum haben Datenmodelle so einen schlechten Ruf?



- Datenmodell, das war doch dieses Riesendiagramm an der Wand in der IT, dass außer Klaus Dagmüller nie jemand verstanden hat.
- Drei Jahre an einem Datenmodell diskutiert und nie verwendet.
- Das Datenmodell lies/lässt sich nicht ändern.
- Datenmodellierer verbreiten im Namen des ‚universe of discourse' Angst und Schrecken

Datenmodelle wie auch Ontologien ermöglichen den Diskurs durch gemeinsame Begriffe

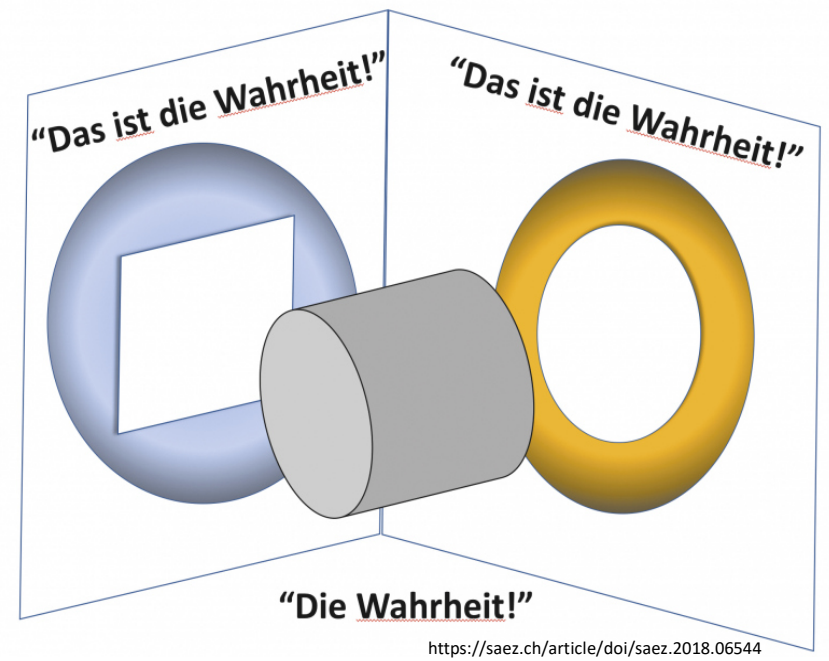
- Einigung auf gemeinsame Begriffe, um zusammenarbeiten zu können
- Ontologien als wesentlicher Baustein moderner Wissenschaft
- Von Aristoteles über Baruch Spinoza bis hin zu Ludwig Wittgenstein
- Deswegen:
  - Univers of Discourse
  - ein Begriff, der alle glücklich macht



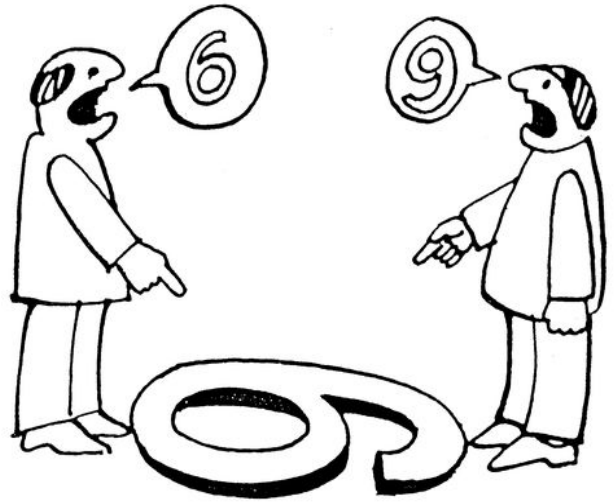
The Cruncher: Automatic Concept Formation Using Minimum Description Length - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/The-Automatically-Created-Zoo-Ontology-Arrows-represent-inheritance-and-the-attributes_fig3_262405011 [accessed 26 Jan, 2022]

Leider gibt es da noch Platos Höhle

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



https://saez.ch/article/doi/saez.2018.06544



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.“

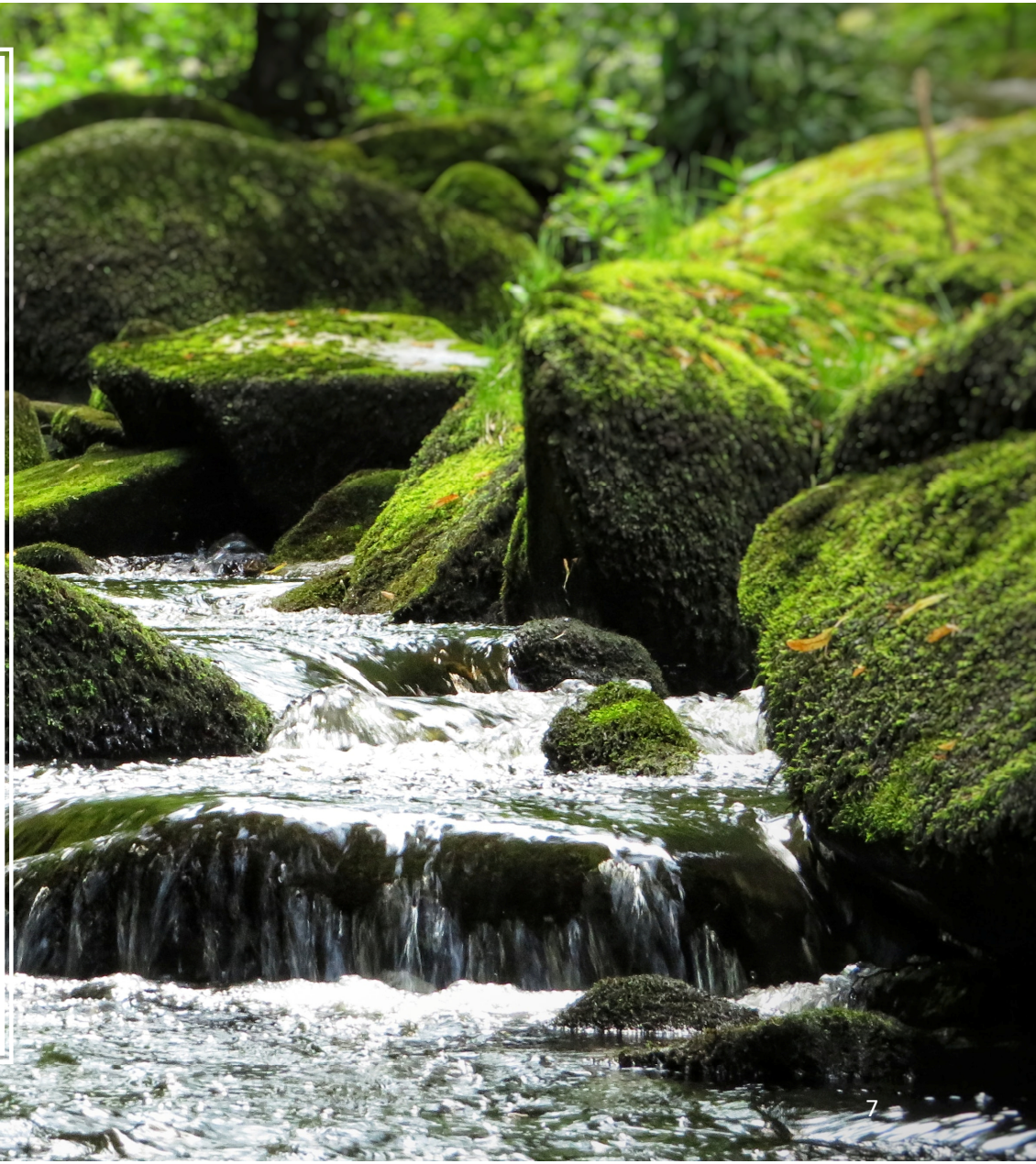
– Niklas Luhmann: Die Gesellschaft der Gesellschaft Seite 44/45

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

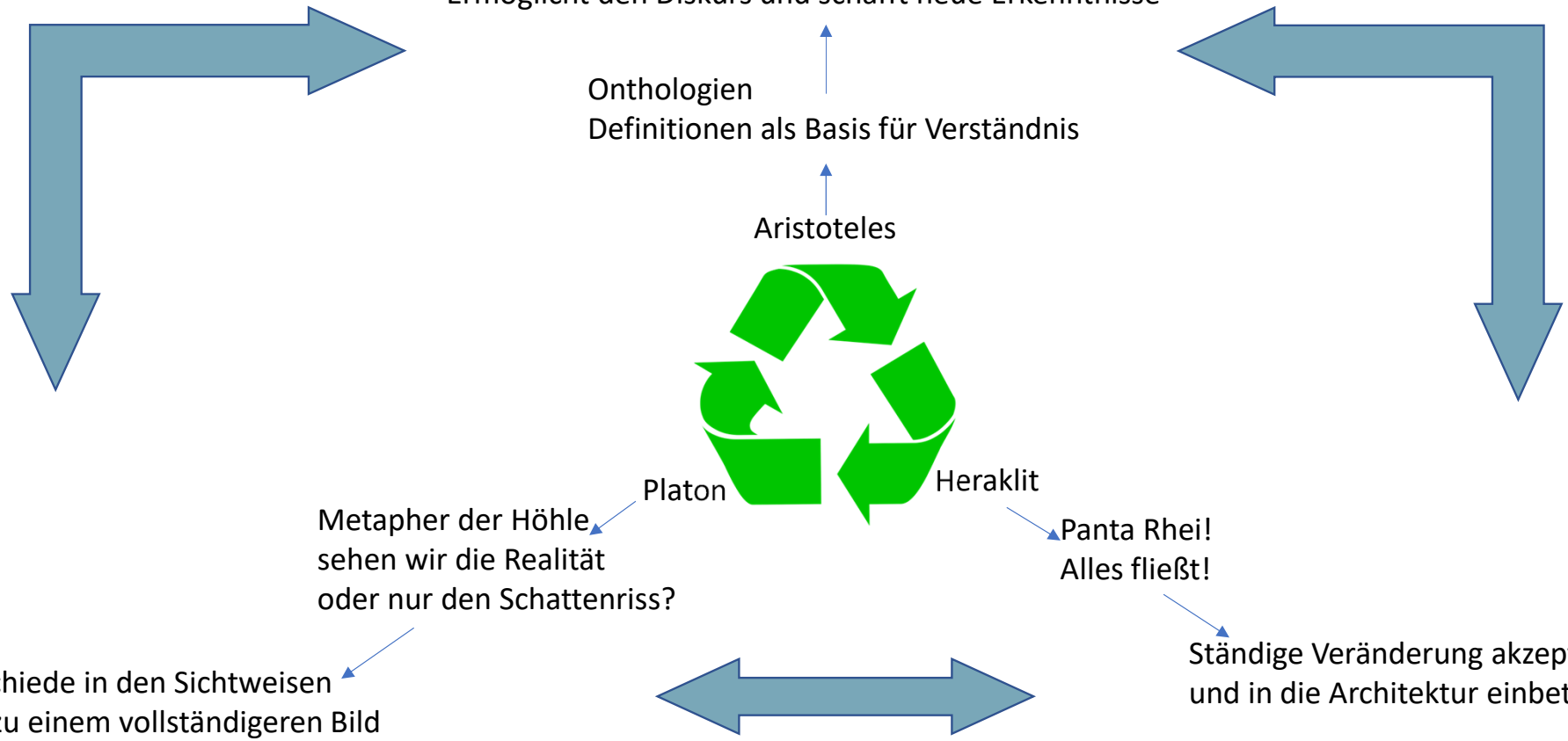
Panta Rhei!
Alles fliesst!

Über die Zeit ändern sich
Ansichten und Konsent

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



Spannungsfeld der Datenmodellierung



Datenmodellierung: Methodik und Kompetenz in Merksätzen

- Datenmodellierung ist Abstimmung: *„Eine gemeinsame Sprache finden“*
- Datenmodellierung ist die Basis zur Kommunikation: *„Weniger Mißverständnisse“*
- Datenmodellierung ist eine Vereinbarung zur Zusammenarbeit

- Unterschiede würdigen und modellieren

- ‚Gut genug‘ als Ziel für das Datenmodell

- Konsent finden: „Damit kann ich leben.“

- Änderungen als Teil der Architektur



Das soll die KI machen, die Daten sind ja schon da

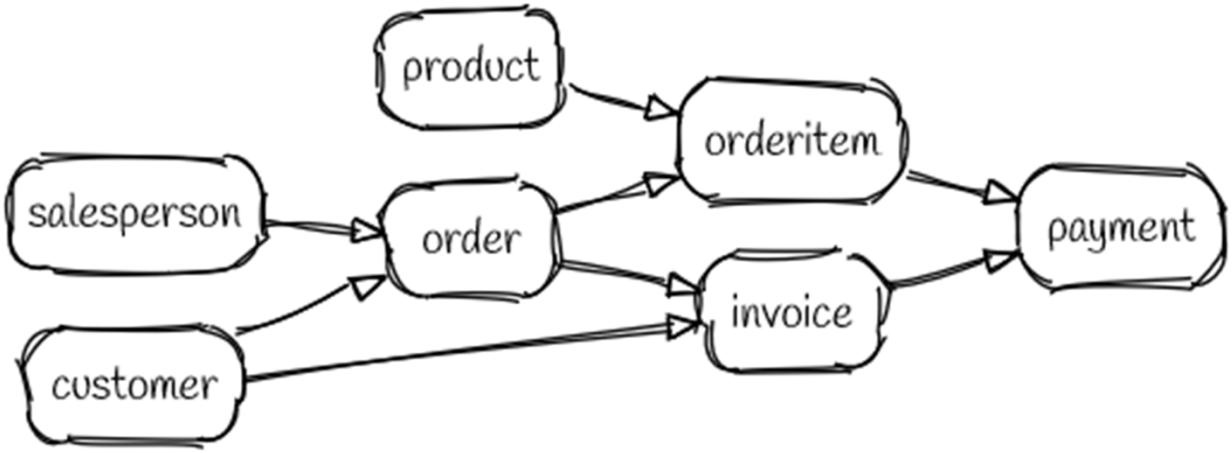
- Die aktuellen Datenmodelle beschrieben den Ist-Zustand
- Die Lücken sind meist nicht dokumentiert:
  - Masterdaten mit unterschiedlichen ID
  - Manueller Übertrag von Daten von einem System ins Andere, gerne auch mit Veränderung der Daten
  - All die kleinen Prozesse, die IT Probleme lösen
- Wie wäre es mal mit einer geschickten Kombination von neu und alt?
- Gutes bestehen lassen, kaputtes gezielt ersetzen



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

m[ethod] 2 data

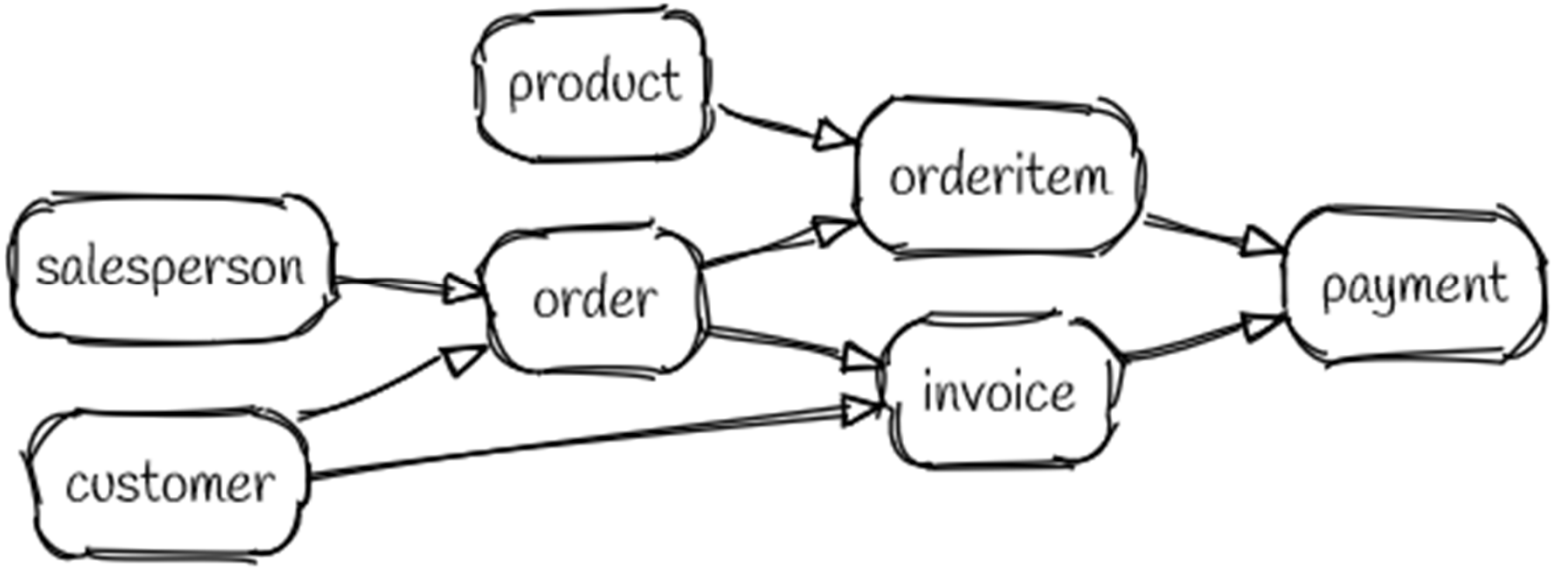
- Ein Unternehmensdatenmodell ist viel zu groß, um es in einem Zug zu verstehen
- Abstraktion ist der Schlüssel
  - ein konzeptionelles Modell verwenden
  - bestehend aus zentralen Geschäftskonzepten und deren Beziehungen
  - nur einen relevanten Teil des Modells betrachten
  - Geschäftsbegriffe verwenden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



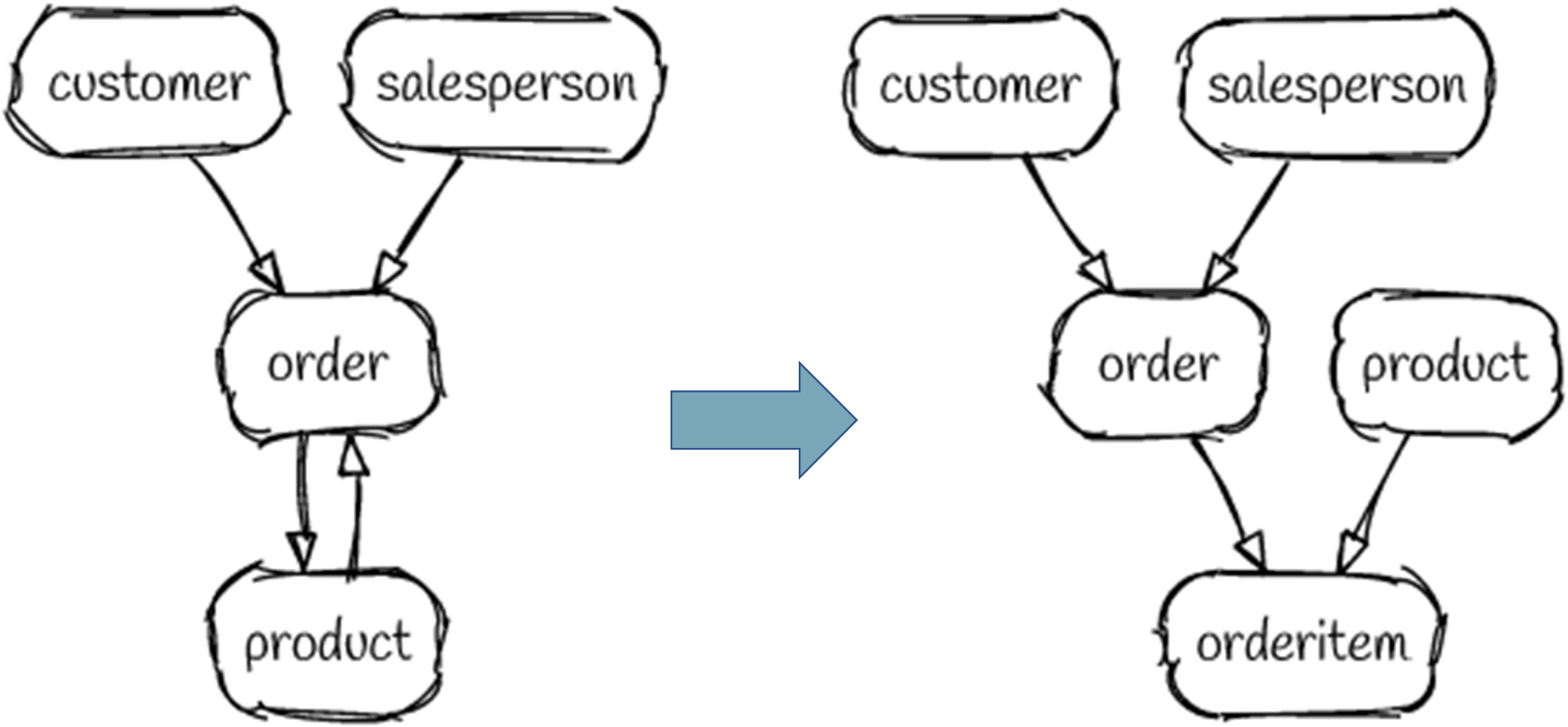
Verständliche Diagramme & Darstellung

- Basiert auf structured entity relationship modelling von Prof. Sinz

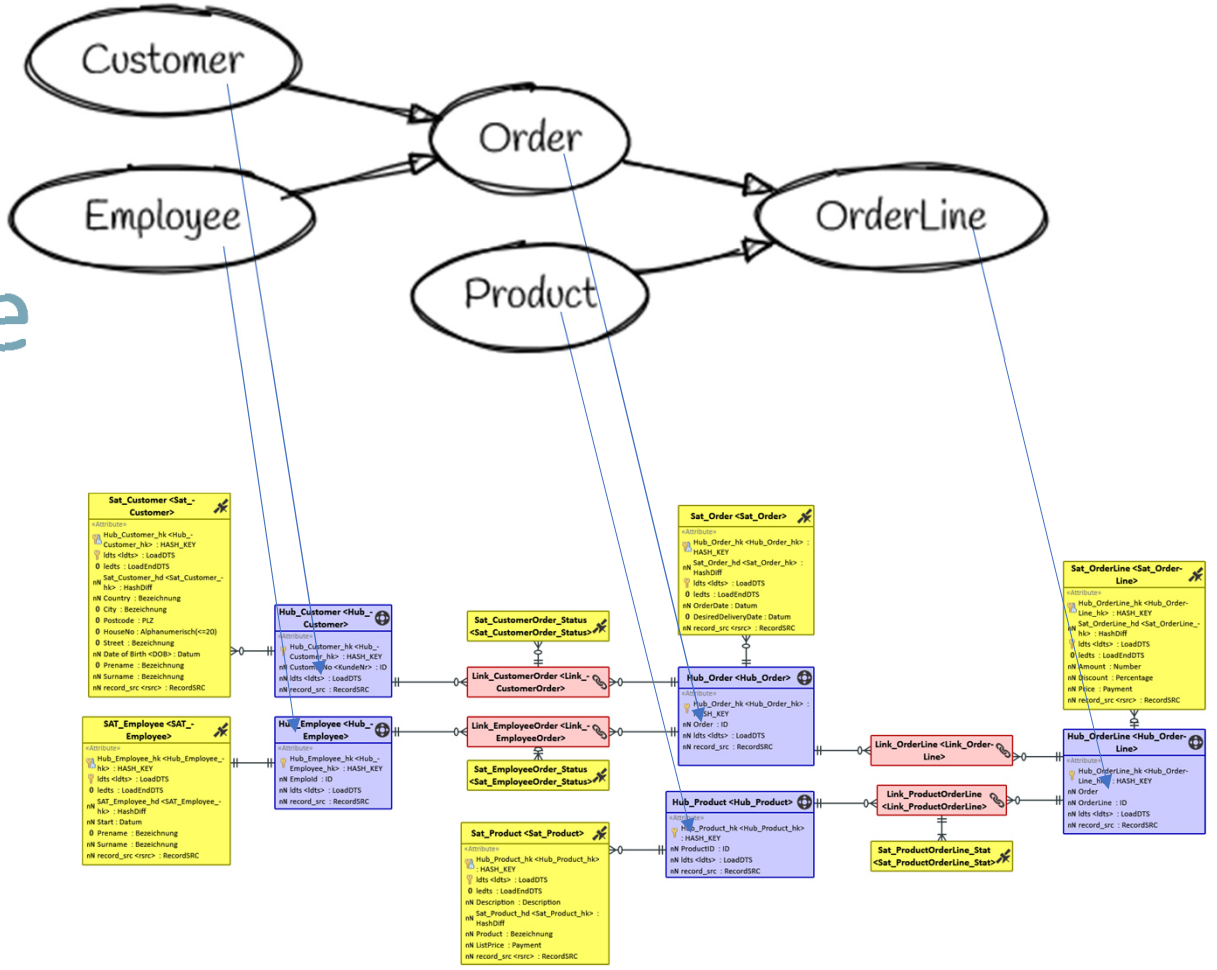
- Entitäten mit einem Fremdschlüssel, sind rechts von der Entität mit dem passenden Primärschlüssel zu platzieren
- Es gibt keine M:N Beziehungen, diese werden in einer Entität dargestellt



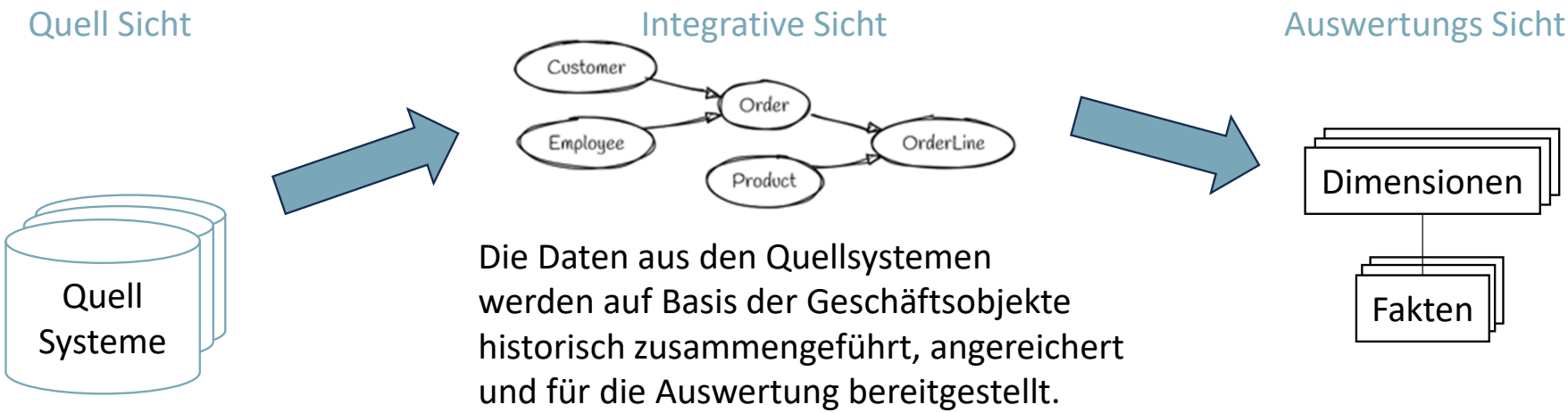
Was bedeutet keine m:n Beziehung?



Die Hubs bilden dieselbe Sicht wie die Geschäfts-objekte



Wirklich aktuell sind Datenmodelle nur, wenn man das Modell als Metadaten nutzt

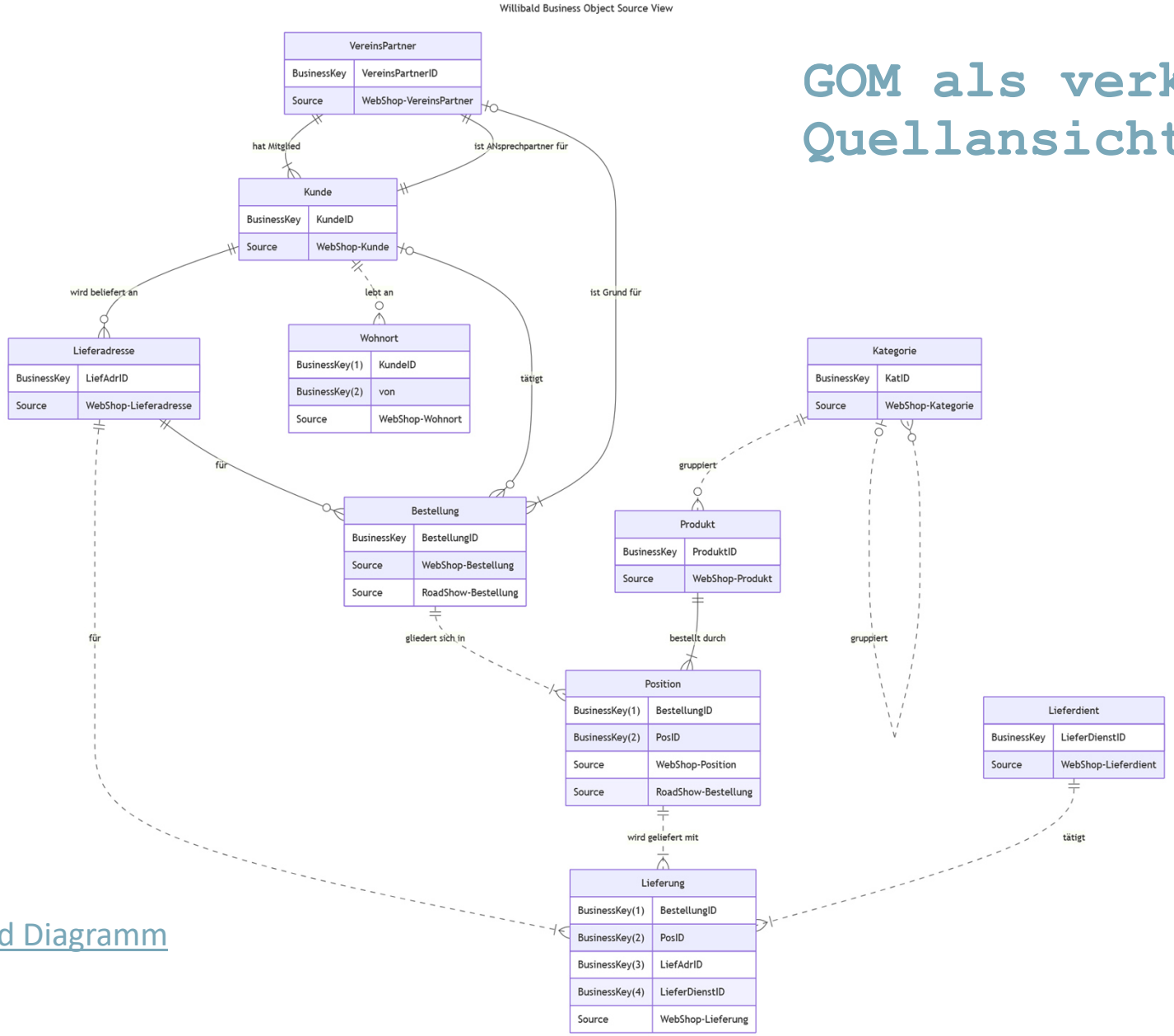




Datenmodell als Code

```yaml
Kunde:
  Domain: Willibald
  desc: Repräsentiert eine kaufende Person mit Stamm-, Kontakt-, Geburts- und Zahlungsdaten.
  business_keys:
  - KundeID
  sources:
  - Kunde
  schema.org:
  attributes:
  - name: KundeID
    type: char(13)
    nullable: false
    primary_key: true
  - name: VereinsPartnerID
    type: varchar(30)
    nullable: true
    foreign_key: true
    references: VereinsPartner.VereinsPartnerID
  - name: Vorname
    type: varchar(128)
    nullable: false
  - name: Name
    type: varchar(128)
    nullable: false
  - name: Geschlecht
    type: char
    nullable: true
  - name: Geburtsdatum
    type: date
    nullable: false
  - name: Telefon
    type: varchar(20)
    nullable: true
```

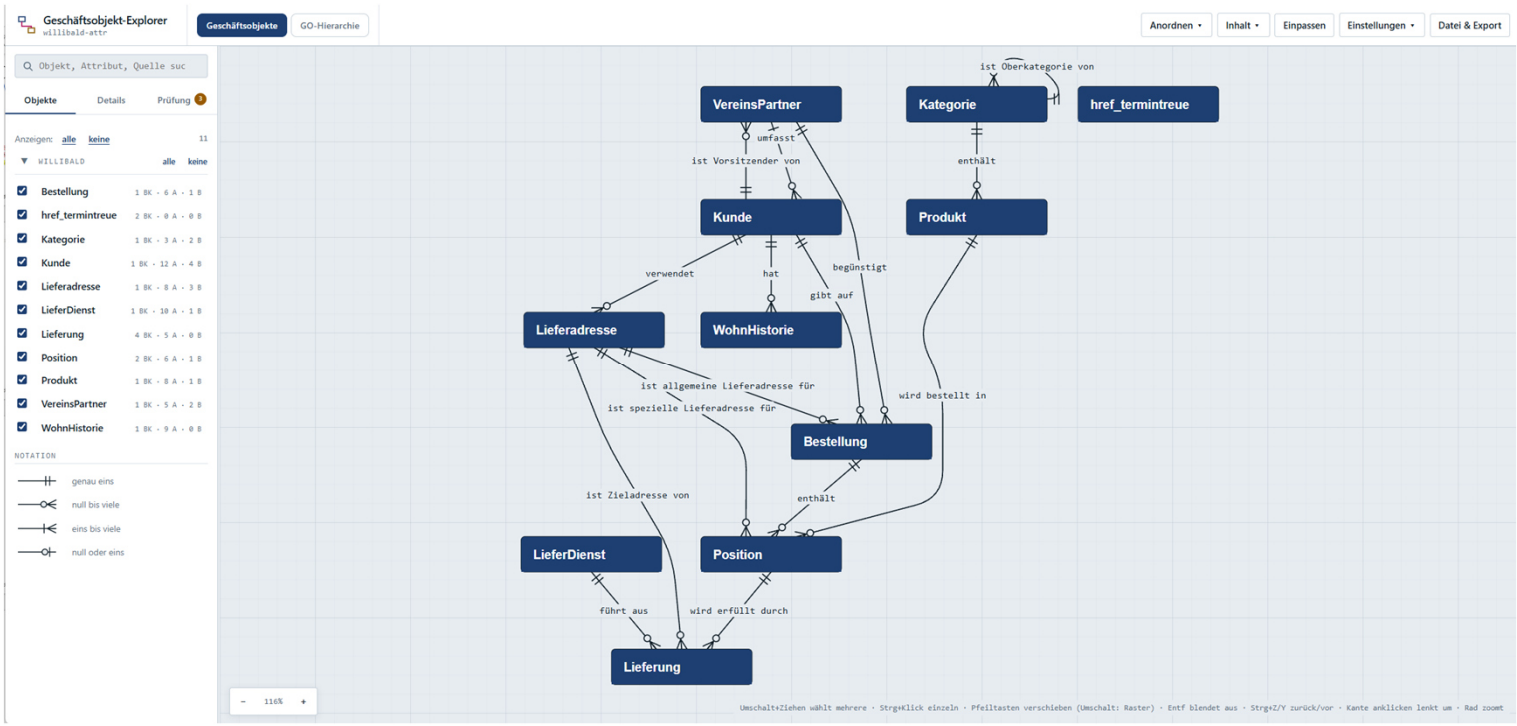
GOM als verknüpfte Quellansicht



[Mermaid Diagramm]

Je mehr Geschäftsobjekte man hat, desto schwerer wird die Pflege der Konsistenz in mermaid

- für das Verständnis sind redaktionell gestaltete Diagramme von Teilaspekten wichtig
- dennoch will man die Daten nur einmal anlegen
- bei mermaid wird es ab 30 Objekten sehr scher pflegbar
- leider gibt es nicht viele Werkzeuge am Markt, die diese Funktionalität bieten und das Datenmodell als code ablegen
- zur Anschauung und zum Test gerne mal dies hier probieren: https://github.com/m2data/rumi.git



Ziel und Grenzen bei der Geschäftsobjektmodellierung

Ziel

- mit Geschäftsobjektmodell und Glossar ein gemeinsames Datenverständnis schaffen
- das Glossar als Basis für Requirements
- individuelle, abweichende und notwendige Sichtweisen bzw. Blickwinkel abbilden
- Klärung der Frage:
  'Welche Daten haben wir?'
- Es muss lediglich 'Gut genug' sein

Out of Scope

- Enterprise Data Model
- Universe of discourse
- endlose Abstimmungen über den richtigen Begriff
- alles was nicht direkt zu einem Data Warehouse führt

Danke für Ihre Aufmerksamkeit!