



## Data-Vault-Automation aus dem Datenmodellierungstool

1. Tagung der DDVUG am 24. Juni 2014



# AGENDA

- 1** MID & Innovator
- 2** Modell & Methode
- 3** Architektur & Automatisierung
- 4** Nutzen & Veränderung

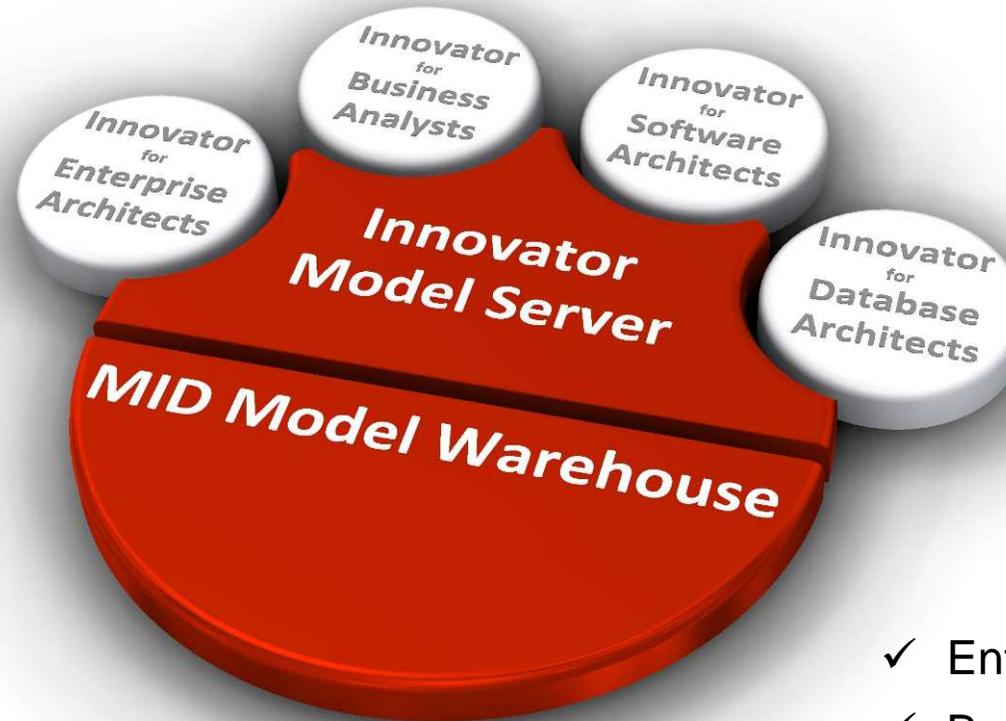
# MID ist seit über 20 Jahren mit Innovator im Markt für Business- und IT-Modellierung erfolgreich



- 600+ Organisationen, Unternehmen und Behörden nutzen heute Innovator
- Kunden sind Unternehmen aus nahezu allen Branchen, darunter BMW, Daimler, Deutsche Telekom, Siemens, AXA, Migros, Bundesagentur für Arbeit, und viele weitere
- 130 MID Mitarbeiter arbeiten in den Büros in Nürnberg, Köln, Stuttgart, München und Zürich
- Seit 2012 vertreibt MID Produkte und Dienstleistungen über eine eigene Landesgesellschaft in der Schweiz mit Büro in Zürich
- Eine eigene Akademie sorgt dafür, dass auch große Organisationen in kurzer Zeit Ihre Mitarbeiter qualifizieren können
- MID ist 2011 zu einem der top 100 innovativsten mittelständischen Unternehmen gewählt worden



**Rollenbasierte Editionen für alle am IT  
Innovationsprozess beteiligten Disziplinen bieten jeweils  
eine perfekt optimierte Arbeitsumgebung.**



- ✓ Enterprise Architektur
- ✓ Business Architektur
- ✓ Software Architektur
- ✓ Datenbank Architektur

**Komplexität beherrschen,  
Transparenz schaffen,  
Freiräume gewinnen**

**Die Lücke zwischen Fachbereich und IT schließen**

→ Anforderungen in umsetzbare Konzepte überführen

**Ein Analysemodell schaffen**

→ von Big Picture ins Detail für Fachbereich und IT

→ eindeutige Kennzahlen ohne langwierige Alignment-Prozesse

**Vom Modell profitieren**

→ Re-Design vermeiden

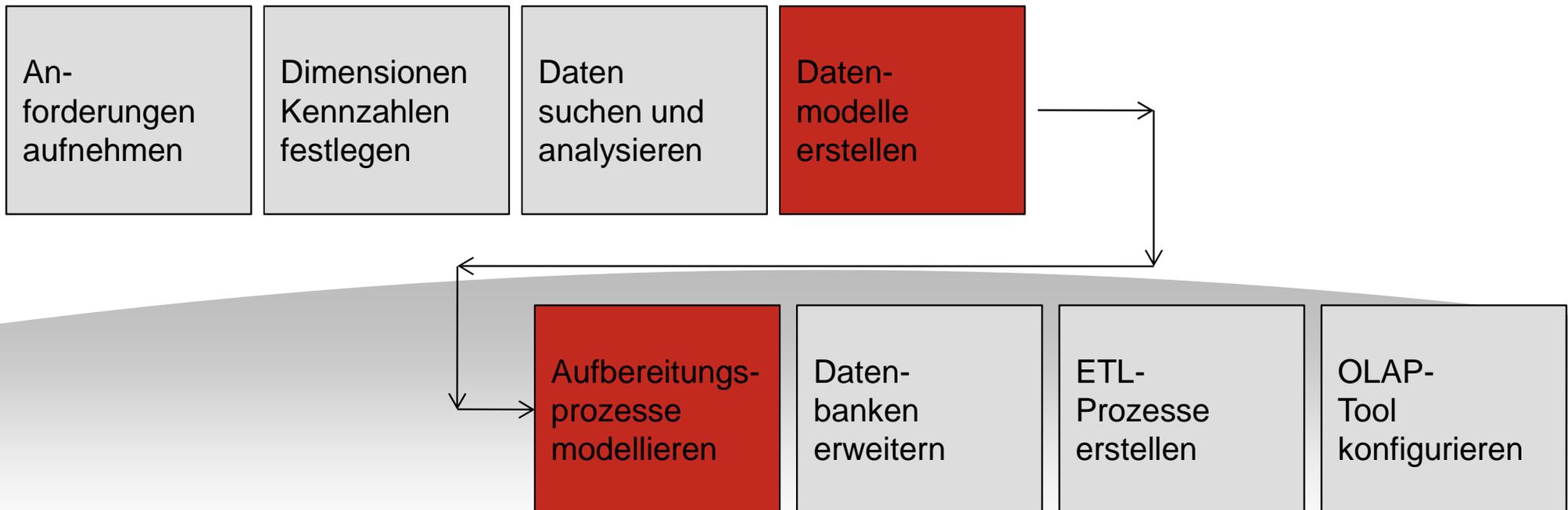
→ Zielsicherer und dadurch schneller arbeiten

→ andere Themen möglich machen

## Die methodische Basis unseres Vorgehens liegt in einfachen Grundprinzipien

- **Komplette Beschreibung von In- und Output...**
- **... auf oberster Ebene ...**
  - Schnittstellen
  - Fakten, Dimensionen und Kennzahlen
- **... bis hin zur Beschreibung eines Prozessschritts**
  - Integration von Daten und Prozessen für das komplette DWH von der Quelle zum Ziel
  - Vollständige Darstellung
- **Trennung von Aspekten**
  - Jedes Element hat seine eigene klare Aufgabe
  - Auflösen von Komplexität

**Ohne Modellierung wird auf die Erhebung von wesentlichen Metadaten und damit auf entscheidenden Nutzen verzichtet.**



Modellieren heißt nur explizit arbeiten. Dafür liefern Modelle einen erheblichen Nutzen:

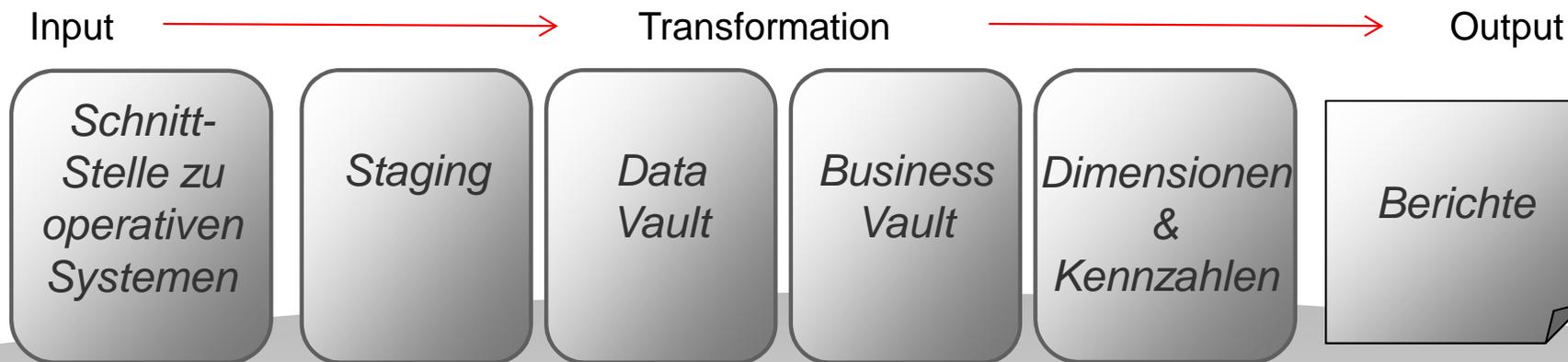
- Dokumentation als integrierter Teil des Entwicklungsprozesses
- Bewahren der Analyseergebnisse statt beschreiben der Lösung

## Was ist so toll an Data Vault?

### Data Vault löst das 80/20 Problem des Data Warehouse

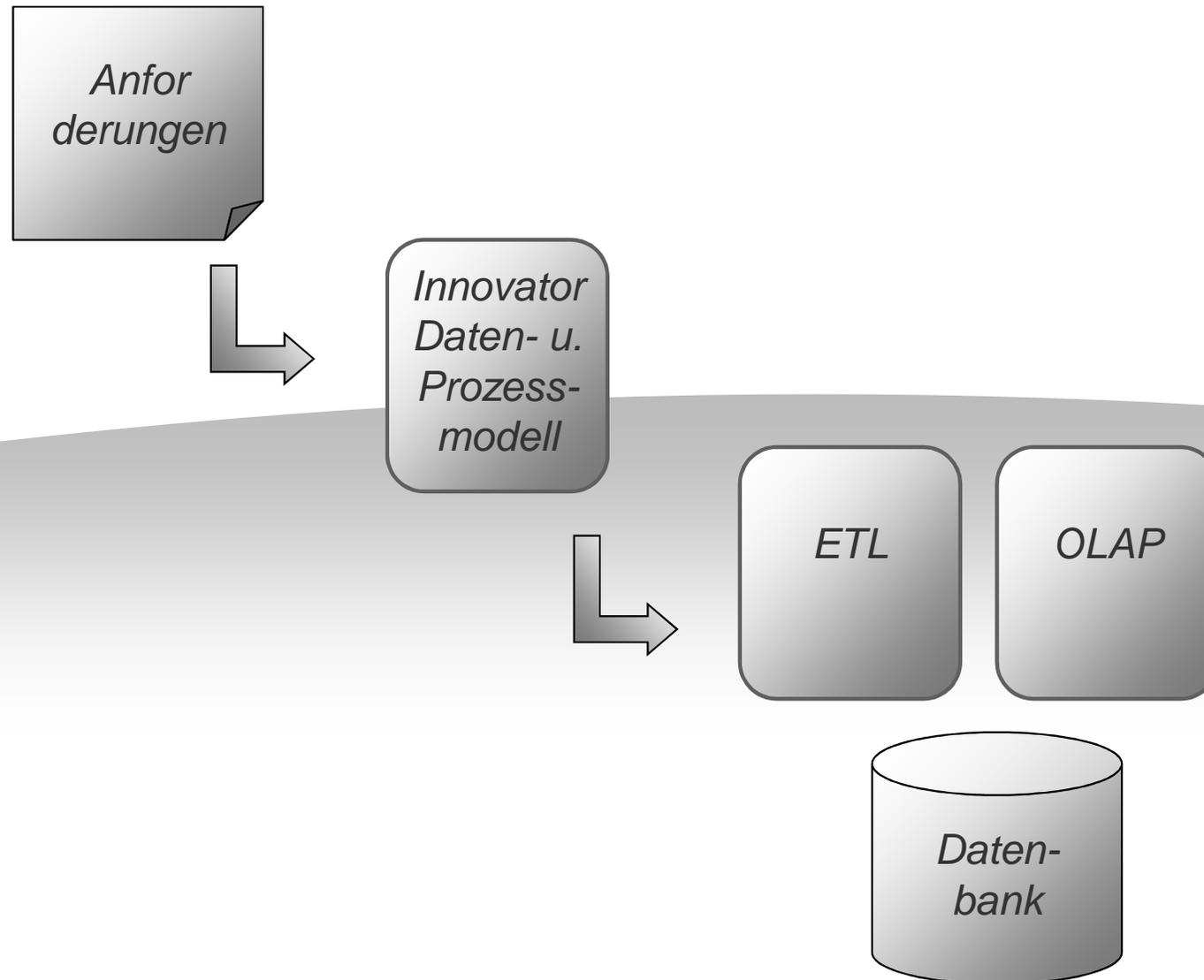
- Data Vault ist ein standardisiertes Vorgehen mit dem Datenmodelle für BI schnell und erweiterbar gestaltet werden können.
- Änderungen sind weniger umfangreich und sind sehr stark begrenzt auf unmittelbar betroffene Daten.
- Data Warehouses mit Data Vault sind auf massiv-parallelen Systemen sehr schnell, es gibt nur inserts im Ladeprozess
- Für komplexe und jeweils für ein Data Warehouse spezifische Lösungen gibt es einen Bereich in dem die Geschäftsregeln hinterlegt sind (Business Vault)
- Die Standardisierung erlaubt einen hohen Grad an Generierung -> schneller werden
- Die Erweiterbarkeit macht Änderungen schneller und leichter -> agil werden
- Es ist einfach möglich CMMI-Level 5 zu erreichen -> besser werden

**Eine Architektur liefert einen festen Rahmen, in den die Anforderungen übertragen und in kleinere Einheiten gespalten werden.**



- Durch die Aufteilung der Arbeitsschritte in Schichten werden die Anforderungen in kleinere Pakete geteilt
- In jeder Schicht werden die Prozesse soweit zerlegt, dass für jedes Attribut einer Schicht genau die Quellattribute festgelegt werden
- Klare Entscheidungen in der Schicht reduzieren die Komplexität
- In dem resultierenden Modell kann nun für jede Dimension und jede Kennzahl nachvollzogen werden, aus welchen Elementen der Schnittstelle sie gebildet wird.

**Der Innovator bildet eine logische Schicht / Sicht über alle BI-Werkzeuge hinweg.**



# NOTATION DATENMODELL

## ER - Elemente

### Entität

«Attribute»

«Fremdschlüssel»

Eindeutig zu bestimmendes Objekt, über das Informationen gespeichert oder verarbeitet werden sollen



### Sicht

«View-Attribute»

«From-Klauseln»

Virtuelle Entität, die bestimmte Aspekte aus einer oder mehreren Entitäten zusammenfasst

### Dimension



«Attribute»

«Fremdschlüssel»

Filter zur Bestimmung einer Kennzahl, Analysegegenstand z.B. Kunde

### Fakt



«Attribute»

«Fremdschlüssel»

Enthält Kennzahl(en), die mit Hilfe der Dimensionen ausgewertet werden sollen



### Würfel

«View-Attribute»

«From-Klauseln»

Ausschnitt der Faktenentität mit den relevanten Hierarchien, Kennzahlen und berechneten Kennzahlen

### Hub



«Attribute»

«Schlüssel»

Fokussiert auf die *stabilen* Elemente innerhalb eines Unternehmens, die von den täglichen Geschäftsvorfällen betroffen sind

### Satellit



«Attribute»

«Schlüssel»

Speichert die *zeitabhängige* Information zu den einzelnen BusinessKeys eines HUB auf granularer Ebene

### Link



«Attribute»

«Schlüssel»

Ermöglicht die Abbildung von Geschäftsprozessen/Transaktionen zwischen einzelnen HUBS (bzw. BusinessKeys).

## DB - Elemente

### Tabelle

«Tabellenspalten»

Relationale Darstellung der Entität auf physischer Ebene



### Sicht

«View-Spalten»

«From-Klauseln»

Relationale Darstellung der konzeptionellen Sicht auf physischer Ebene

### Kennzahlen

«Kennzahlen»

Darstellung der Kennzahl als Anforderung

### Hierarchien

«Projekt»

▷ «Projekttyp»

▷ «Projekt»

Die Hierarchie fasst mehrere Ebenen (Entitäten) zusammen. Den Entitäten sind Tabellen und Views zugeordnet.

### Attribute

«Attribut»

Eigenschaft einer Entität

«Sichtattribut»

Eigenschaft einer Sicht

«Basiskennzahl»

Kennzahl, die auf Basis der DWH-Schicht ermittelt wird. Bildet Basis für alle weiteren Kennzahlen.

«Abgeleitete Kennzahl»

Geht mit Hilfe von Rechenoperationen aus einer Basiskennzahl hervor.

«Berechnete Kennzahl»

Spezialform der abgeleiteten Kennzahl. Sie wird im OLAP- Würfel erst zur Laufzeit ermittelt.

## Mit Modellen haben wir beim ersten Erstellen des BI-Systems folgende Vorteile

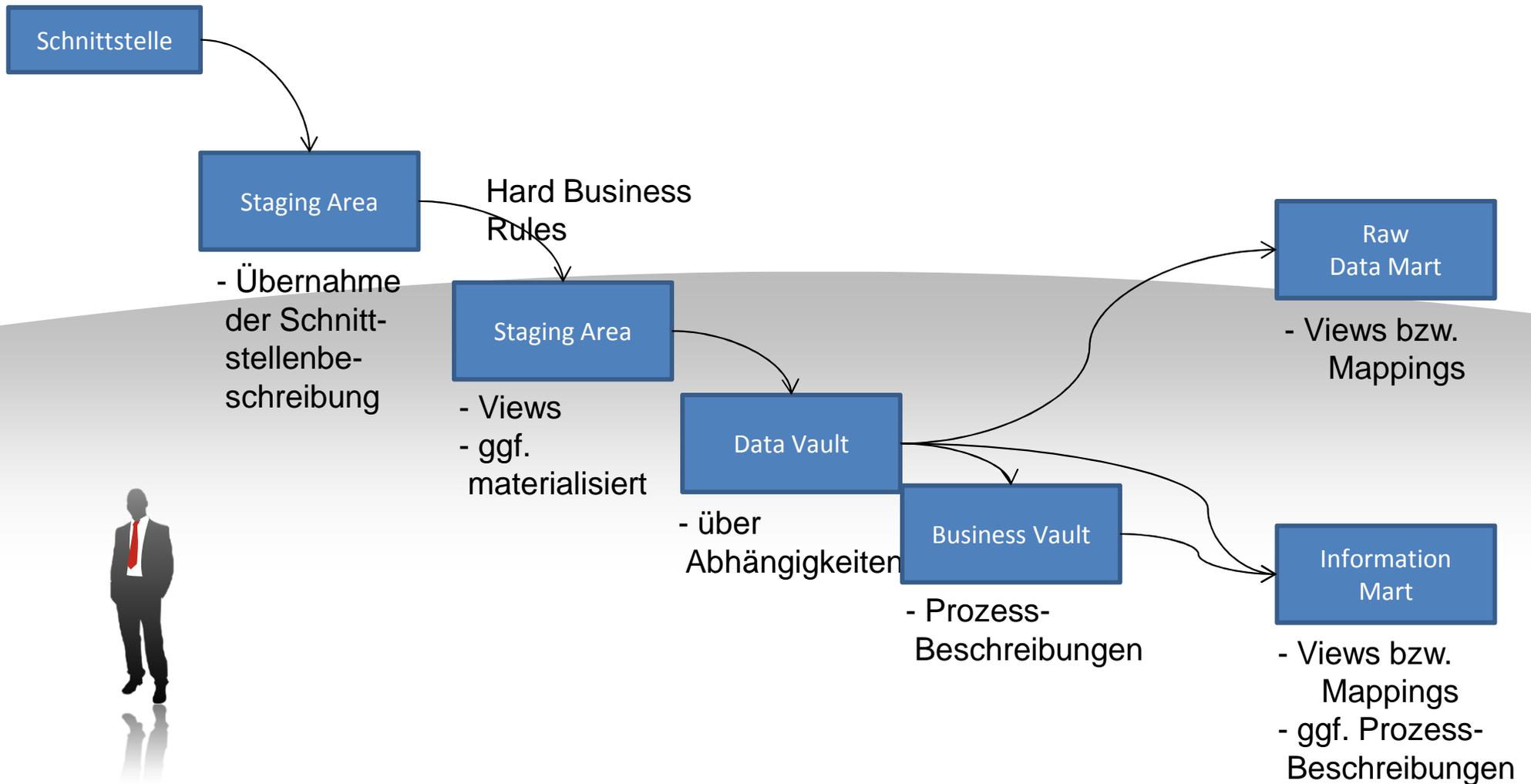
- Senken der Fehlerquote während der Analyse
- Aufteilen komplexer Prozesse in einfachere Schritte, dadurch weniger komplexe ETL-Schritte
- Klare Vorgaben für die Umsetzung
- Auch bei BI-Systemen, die über mehrere verschiedene Zielsysteme gehen, kann der Gesamtblick hergestellt werden

Damit sinkt die Abhängigkeit von einzelnen Entwicklern oder besonders komplexen Prozessschritten.

## Übergang zu Innovator

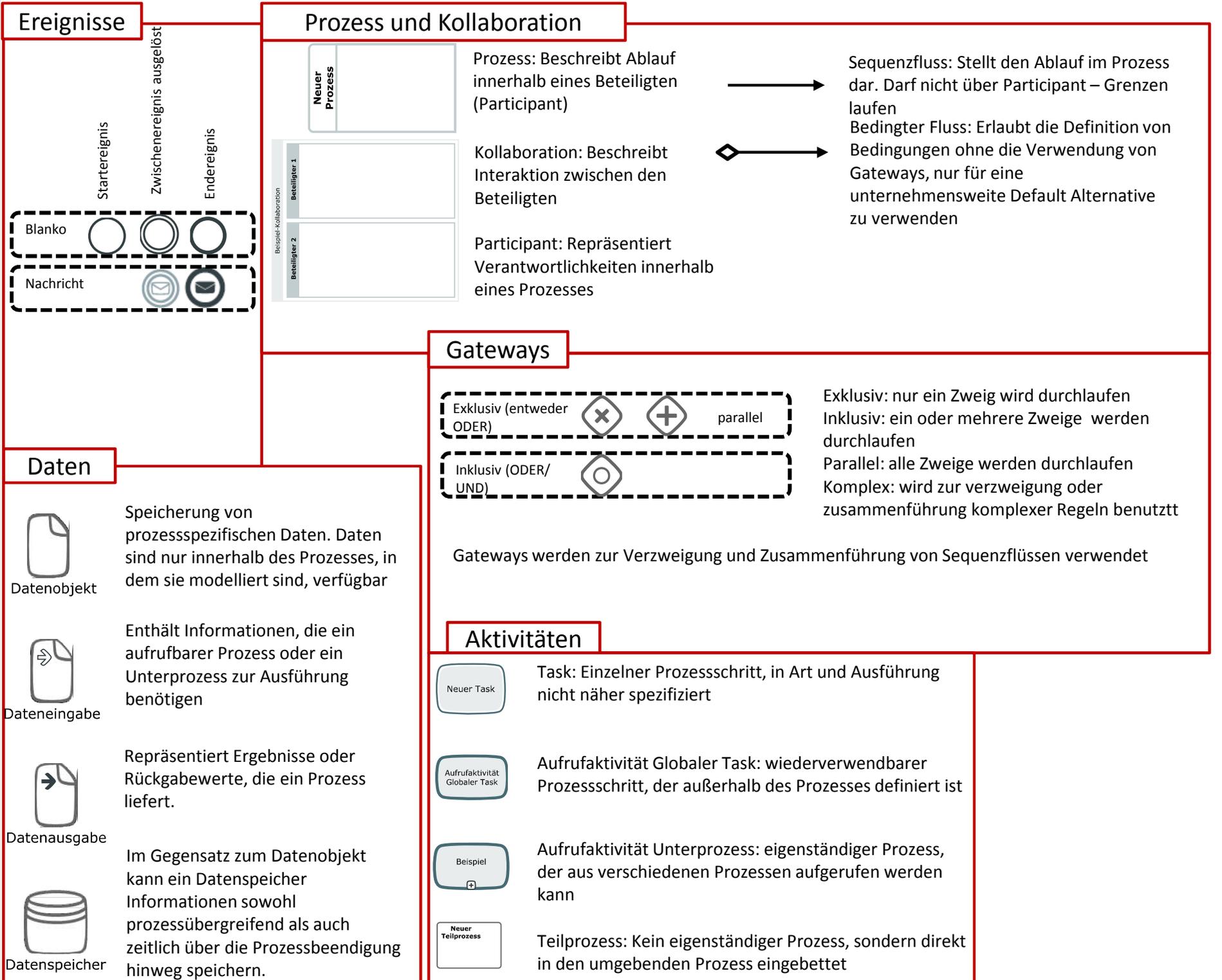
- Zeigen der Beispieldaten
- Erstellen eines Hubs, Sateliten mit dem Beamer
- Generieren von SQL
- Generieren einer SSIS Strecke

## Nur mit standardisierten Architekturen kann auch eine Automation erfolgen





# NOTATION PROZESSMODELL



---

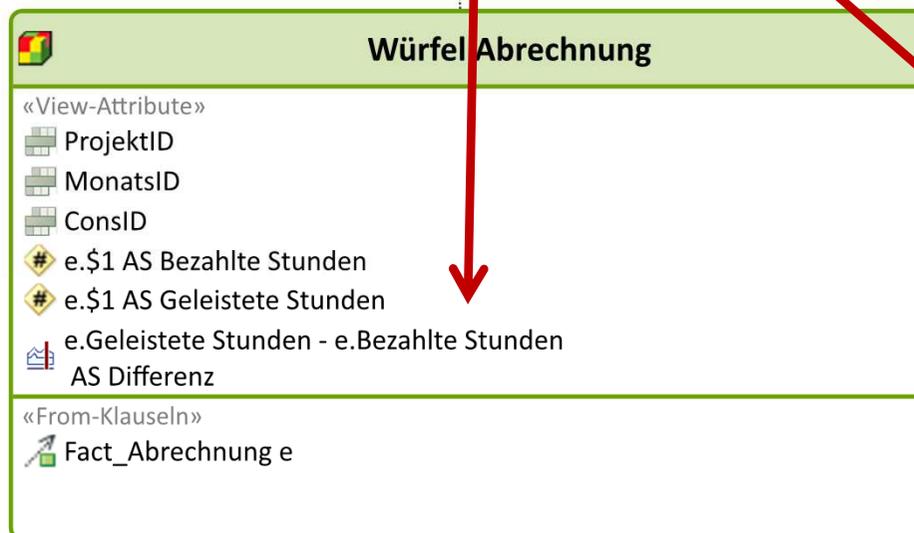
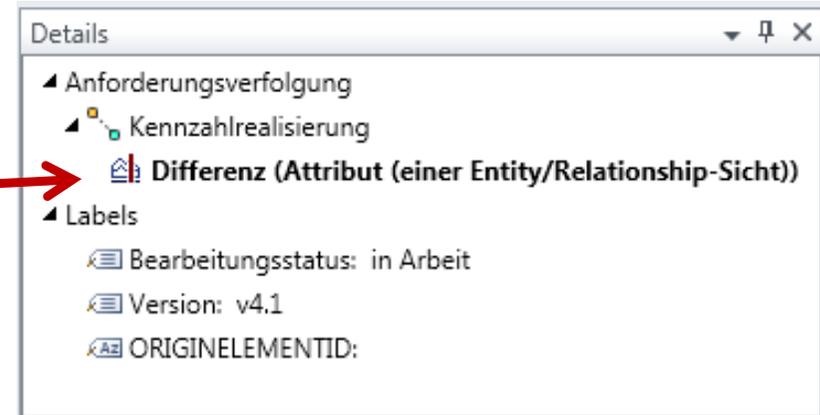
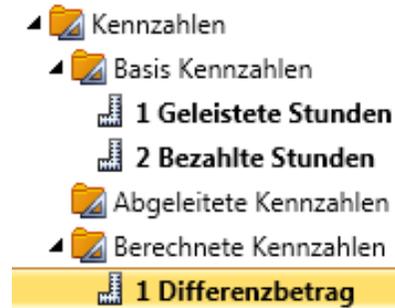
## Übergang zu Innovator

- Prozessbeschreibung für Business Vault
- Beschreibung auf Attributebene
- Impact Analyse

# KENNZAHLEN

Kennzahlen werden in einem eigenen Modellelement textuell beschrieben

Im Detailfenster ist ersichtlich, an welcher Stelle die Anforderung realisiert wird



Differenzbetrag

Würfel kann Basis-, berechnete und abgeleitete Kennzahlen enthalten

## Folie 18

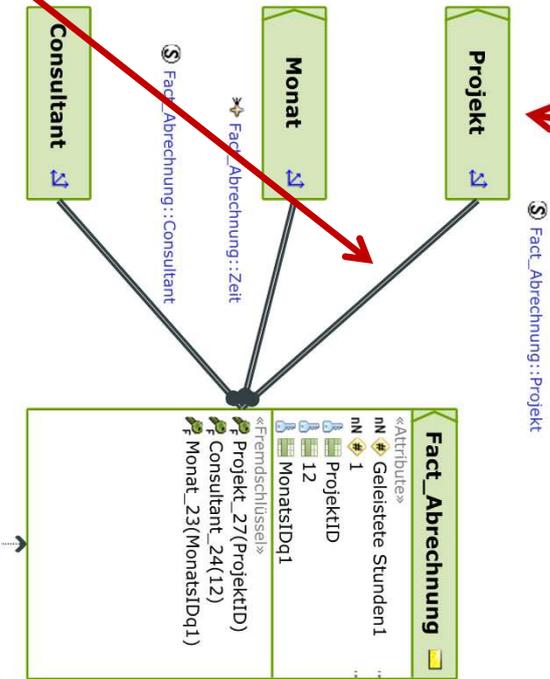
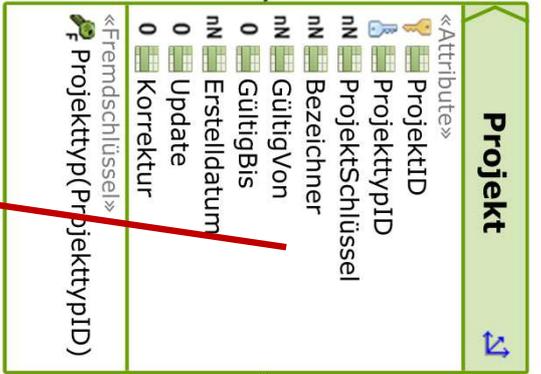
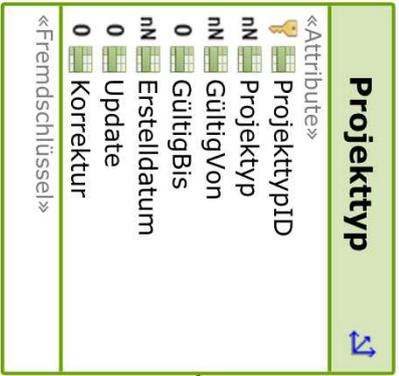
---

**BNB4**

da ist zumindest das ICON vom ersten ATtribut falsch  
besser?

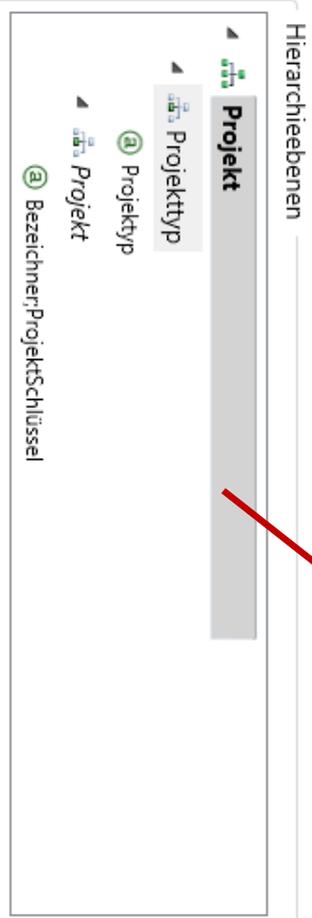
Beate Navarro Bullock; 16.01.2013

# HIERARCHIEN



Geleistete Stunden

Bezahlte Stunden

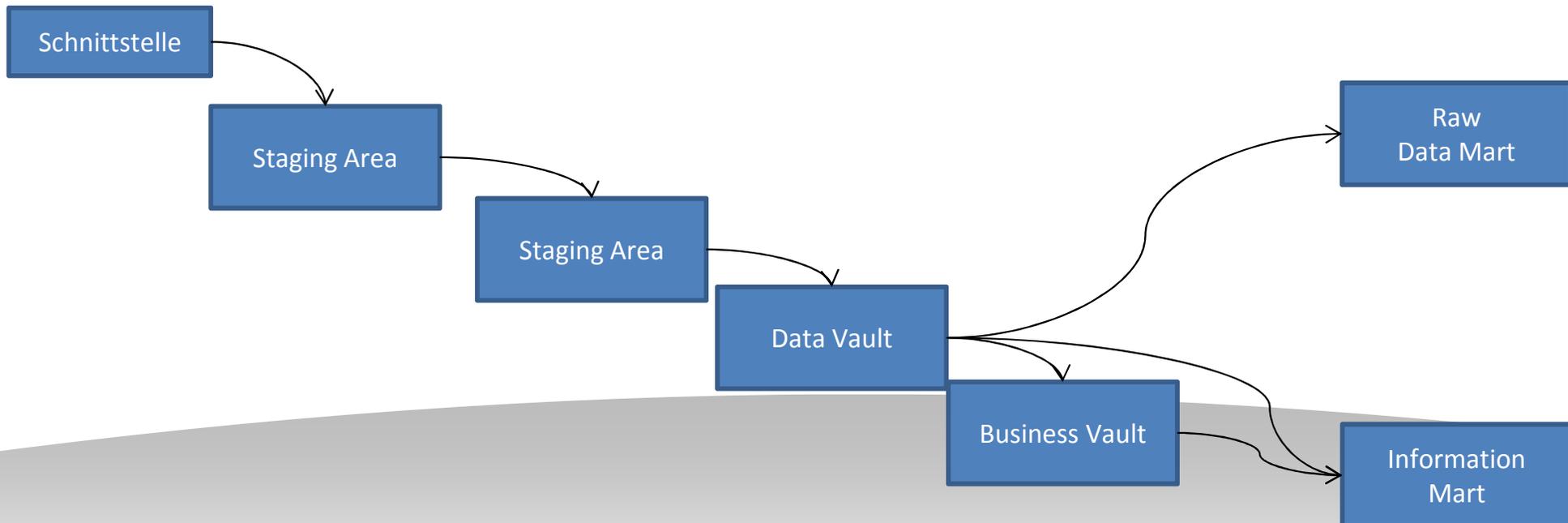


---

## Übergang zu Innovator

- Zeigen eines Data Marts
- Kennzahlenbeschreibung
- Hierarchieeditor
- Generieren der OLAP-Beschreibung

## Mit Modell und Methode einen zentralen Ort schaffen, an dem alle Informationen zum Data Warehouse hinterlegt sind.



- Die Modelle sind verständlicher als SQL, da sie den Analyseteil enthalten und so Abhängigkeiten aufzeigen
- Änderungen sind einfacher zu planen und zu steuern
- Ursprüngliche Architekturideen sind einfacher durchzuhalten
- Das BI-System wird stabiler, ein Re-Design seltener nötig

Bei einer kompletten Neu-Implementierung kann auf den bisherigen Ergebnissen weiter gearbeitet werden, da die Modelle unabhängig von den verwendeten Werkzeugen sind



## Die Zusammenarbeit mit der Fachabteilung wird erleichtert

- Anhand der Modelle kann man sich einfacher verständigen
- Mit dem Modell lässt sich aufzeigen, warum eine Dimension oder Kennzahl wenige Tage zur Erstellung braucht, während andere mehrere Wochen brauchen. Das erhöht das Vertrauen in die IT
- Die Konsolidierung der Dimensionen und Kennzahlen zwischen Endanwendern kann nach und nach erfolgen – die Unterschiede sind klar ersichtlich und können jederzeit für die Klärung der Differenzen in den Zahlen und auch in den Ansichten/Werten/Positionen dienen
- Senken der Recherchekosten für die Erschließung neuer Auswertungen

**Alles in allem können wir damit die Komplexität in der Datenkonsolidierung und in der Fachlichkeit der Auswertung (Dimensionen und Kennzahlen) aufnehmen, begreifbar machen und damit beherrschbar. Modelle bieten volle Kontrolle über die Zusammenhänge im kompletten Unternehmen. Je mehr Quellsysteme zu integrieren sind und je mehr Zielsysteme zu bedienen sind, desto mehr lohnen sich Modelle.**

# Vielen Dank !

**Michael Müller**

**MID GmbH**

Kressengartenstrasse 10  
90402 Nürnberg

*Telefon:* +49 911 / 968 36 0  
*Fax:* +49 911 / 968 36 100  
*E-Mail:* [m.mueller@mid.de](mailto:m.mueller@mid.de)  
*Internet:* [www.mid.de](http://www.mid.de)

**MID**  
the modeling company



*Internet:* [www.mid.de](http://www.mid.de)  
*E-Mail:* [m.mueller@mid.de](mailto:m.mueller@mid.de)  
*Fax:* +49 911 \ 968 36 100  
*Telefon:* +49 911 \ 968 36 0