



Die richtige Grafik als Zauberstab zur Mart Modellierung

Referat vom 22.11.2016

Peter Brunner, NTTData

Dr. Andrea Kennel, InfoPunkt Kennel GmbH

NTT DATA

- 1 Semantische Modellierung
- 2 Business Intelligence – Mehrdimensionale Modellierung in Projekten die BI Analytics betreffen
- 3 Modellierungs-Methodologie : ADAPT
(Application Design for Analytical Processing Technologies [Bulos 1996])
- 4 Beispiel: Finanzindustrie – Zahlungsverkehr (ISO20022)
- 5 Generator ADAPT-Modelle in PowerDesigner





Methoden-Position

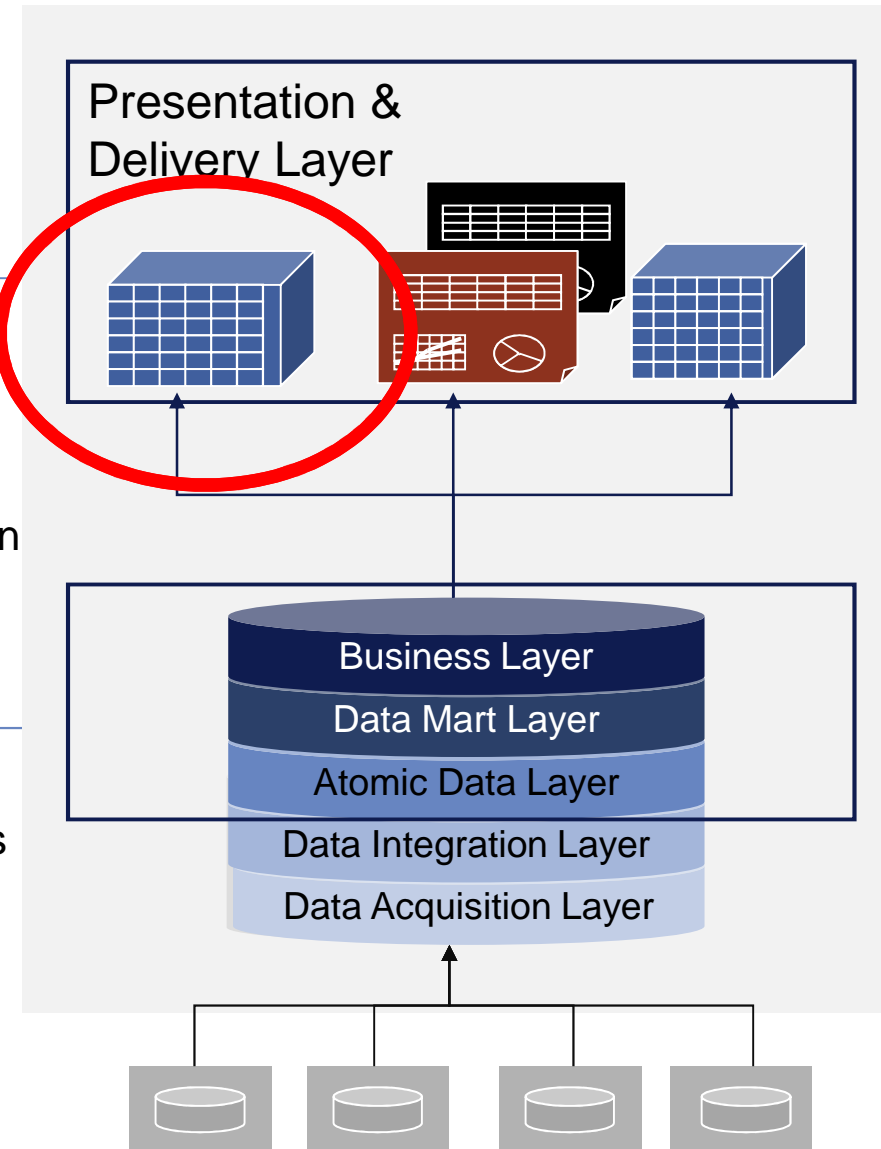
Das **Reporting Repository** als konzeptionelles Datenmodell zur Abbildung der verfügbaren Reporting-objekte inkl. Auswertungsstrukturen und Kennzahlen, auf die die Fachbereiche Zugriff haben.



Anforderungen



Das **Reporting Repository** als konzeptionelles Datenmodell zur Abbildung der verfügbaren Reportingobjekte inkl. Auswertungsstrukturen und Kennzahlen, auf die die Fachbereiche Zugriff haben.



Modellierung einheitlicher Fakten und Dimensionen (nach Kimball) ist ein zusätzlicher Grund die Fachlichkeit verständlich aufzunehmen und abzustimmen

- Nutzen der einheitlichen Fakten und Dimensionen
- Alle sprechen dieselbe Sprache und nutzen dieselben Dimensionen und Faktenbasis
- Daten unterschiedlicher Auswertungen sind vergleichbar
- Festlegen der einheitlichen Fakten zwingend zusammen mit allen betroffenen Bereichen des Business
- Festlegen der einheitlichen Dimensionen zwingend zusammen mit allen betroffenen Bereichen des Business.
- Suchen von Konsenslösungen, und nicht Abbilden von allen Spezialfällen



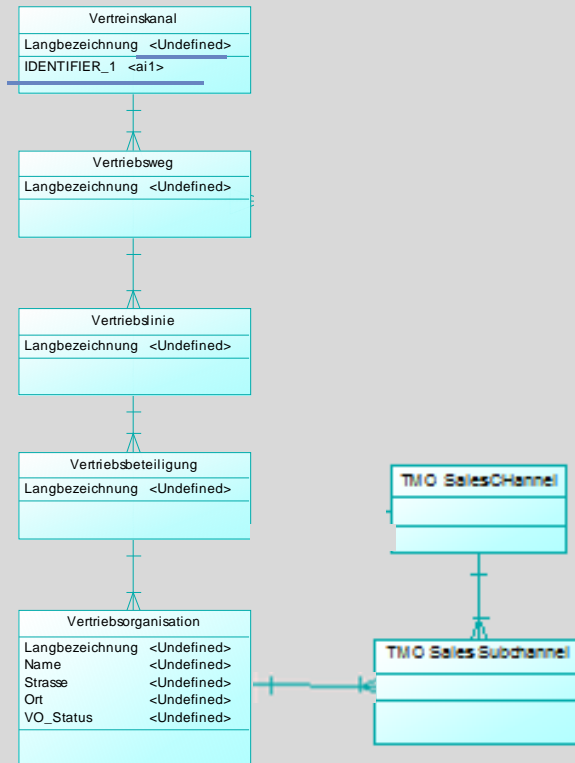
ADAPT Modellierungs-Methodologie

Literatur-Referenz

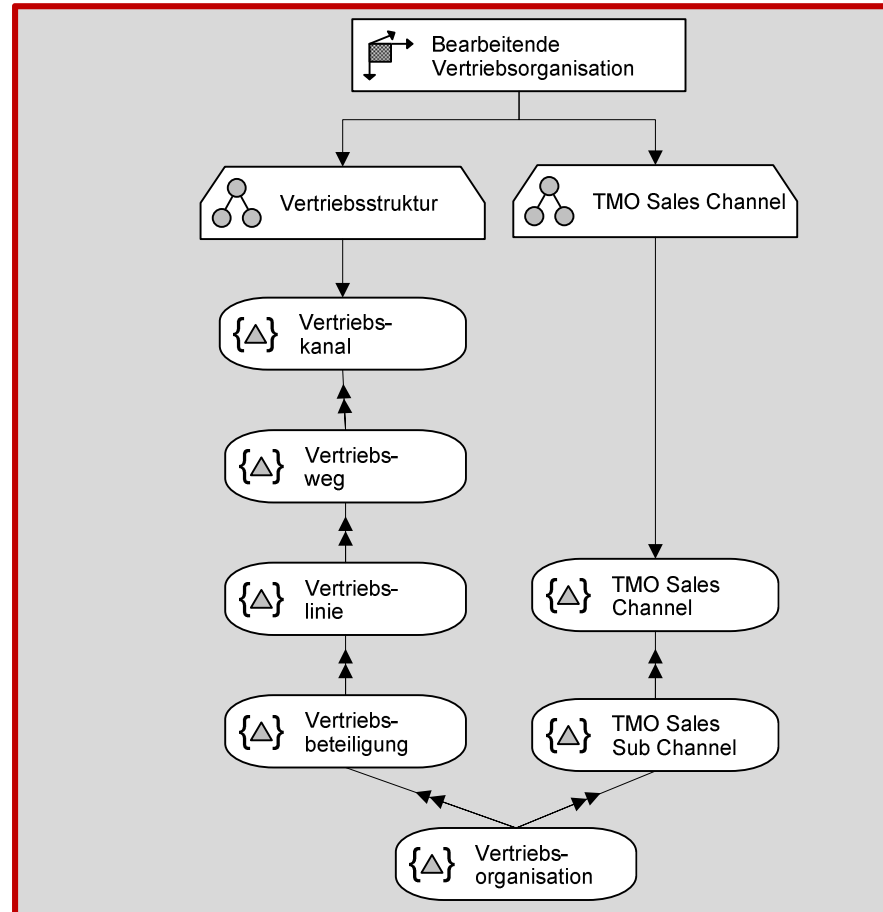
Michael Hahne;
u.a. Modellierung von Business-Intelligence- Systemen
dpunkt.verlag, 2014

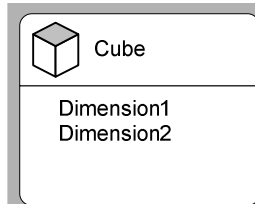
Ralph Kimball;
The Data Warehouse Toolkit: The Definitive Guide to Dimensional
Modeling, 3rd edition, 2013

Konzeptionelles / semantisches Datenmodell in Krähennuss-Notation (PowerDesigner)

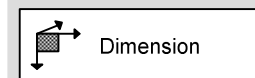


Konzeptionelles / semantisches Datenmodell zur Darstellung multidimensionaler Dimensionsstrukturen (ADAPT)

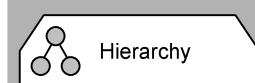




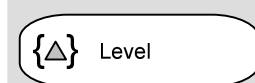
Der Datenwürfel (**cube**) repräsentiert den multidimensionalen Datenraum. Jeder Datenwürfel besitzt eine Bezeichnung und besteht aus mehreren Dimensionen, die auf Kennzahlen zugreifen. Eine OLAP-Anwendung kann aus einem oder mehreren Datenwürfeln bestehen.



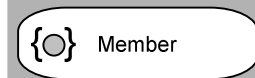
Die Dimension (**dimension**) bildet ein diskretes Kriterium zur Auswahl, Zusammenfassung und Navigation innerhalb der Datenstrukturen ab. Dimensionen spannen letztlich den multidimensionalen Datenraum auf.



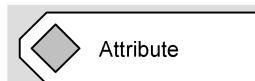
Jede Dimension kann einen oder mehrere Konsolidierungspfade ausweisen, die als Hierarchie (**hierarchies**) bezeichnet werden. Eine Hierarchie ist in mehrere Ebenen (Hierarchiestufen) unterteilt.



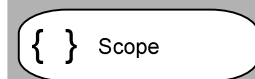
Eine Hierarchiestufe (**level**) definiert eine Ordnungsstruktur innerhalb einer Hierarchie. Mitglieder einer Hierarchiestufe teilen eine gleiche implizite Charakteristik. Durch die stufenweise Zusammenfassung von Dimensionselementen entstehen innerhalb der Dimensionen Elementhierarchien, über die eine Navigation im Datenbestand durch den Endbenutzer (Drill-Down-/Roll-Up-Operationen) erfolgen kann.



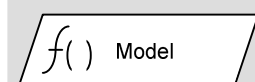
Eine einzelne atomare Einheit (**member**) innerhalb einer Dimension wird als Dimensionselement bezeichnet. Grundsätzlich werden Dimensionselemente in Hierarchiestufen zusammengefasst.



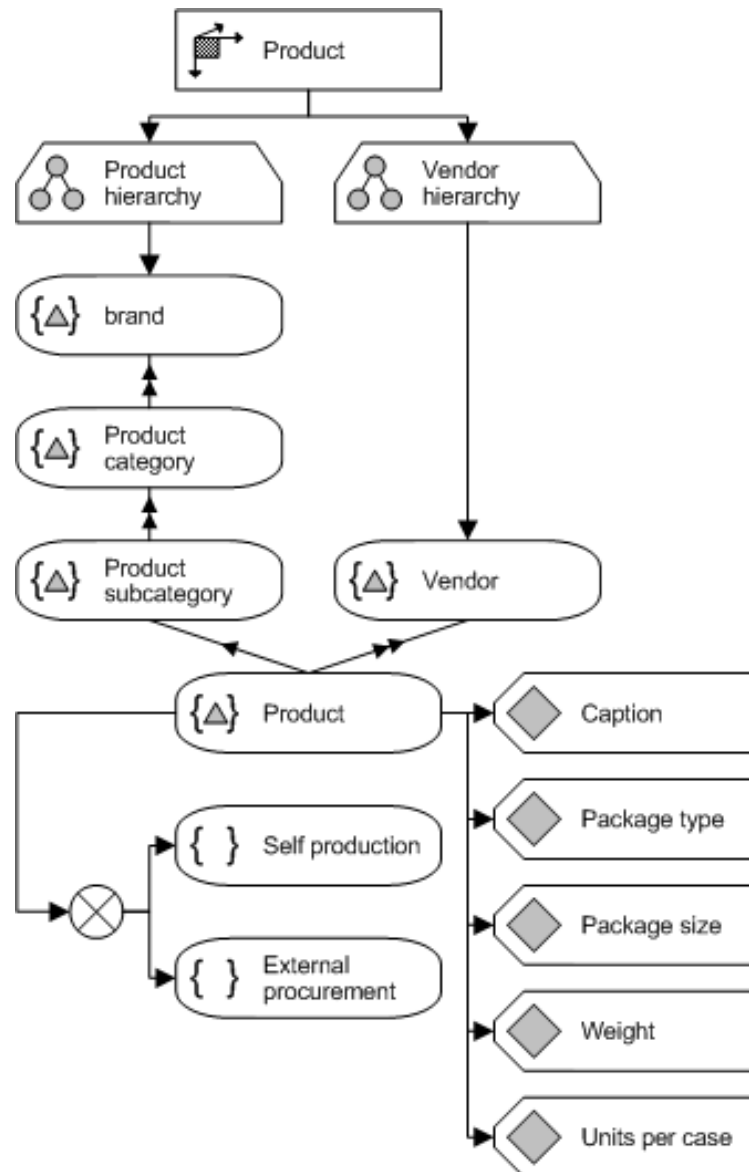
Dimensionsattribute (**attributes**) geben zusätzliche Informationen zu einzelnen atomaren Einheiten innerhalb einer Dimension.



Für bestimmte Sachverhalte werden lediglich Dimensionsausschnitte (**scope**) als Mengen gültiger Dimensions-elemente von speziellem Interesse benötigt.



Die hinter abgeleiteten Modellen (**model**) stehenden Berechnungsformeln, z.B. zur Berechnung von abgeleiteten Kennzahlen, können algebraisch spezifiziert werden.



- Die **Konsolidierung** in der Produkt-Hierarchie erfolgt durch Marken, Produktkategorien und Produktunterkategorien.
- Eine Aggregation zu den verschiedenen Herstellern der Produkte wird darüber hinaus betrachtet.
- Produkttitel, Größe und Art der Verpackung sowie Gewicht sind weitere relevante Informationen neben der Produkt-Nr. Sie sind die **Attribute der Produkthierarchieebene**.
- Selbstproduktion und externe Beschaffung der Produkte sind differenziert, wobei das Objekt der **Dimensionssicht** benutzt wird .



Papierübung anhand eines konkreten Beispiels

- Kunden haben eine Adresse
- Wir wollen Kundenkennzahlen nach geografischen Gesichtspunkten auswerden.
- Dabei sollen die Daten nach Länder ausgewertet werden.
- Für die Schweiz soll dann ein Drill-Down über die Region bis zum Kanton gemacht werden können
- Zeichnen Sie die dazu nötige Hierarchie in ADAPT



Konkretes Beispiel: Data Mart Dimensionen im Zahlungsverkehrs-Teilbereich

Harmonisierung des Zahlungsverkehrs: Die Zukunft des Schweizer Zahlungsverkehrs heisst ISO 20022

Überweisungen

Harmonisierung der Überweisungsverfahren von Banken und PostFinance.

Verfahren werden weitestgehend an die Regeln des europäischen SEPA-Überweisungsverfahrens (ISO 20022) angeglichen.

Lastschriftverfahren

Konsolidierung der Schweizer Lastschriftverfahren.

E-Rechnung wird forciert und mit einer neuen Lastschriftlösung kombiniert.

Einzahlungsschein mit Datencode (QR-Code)

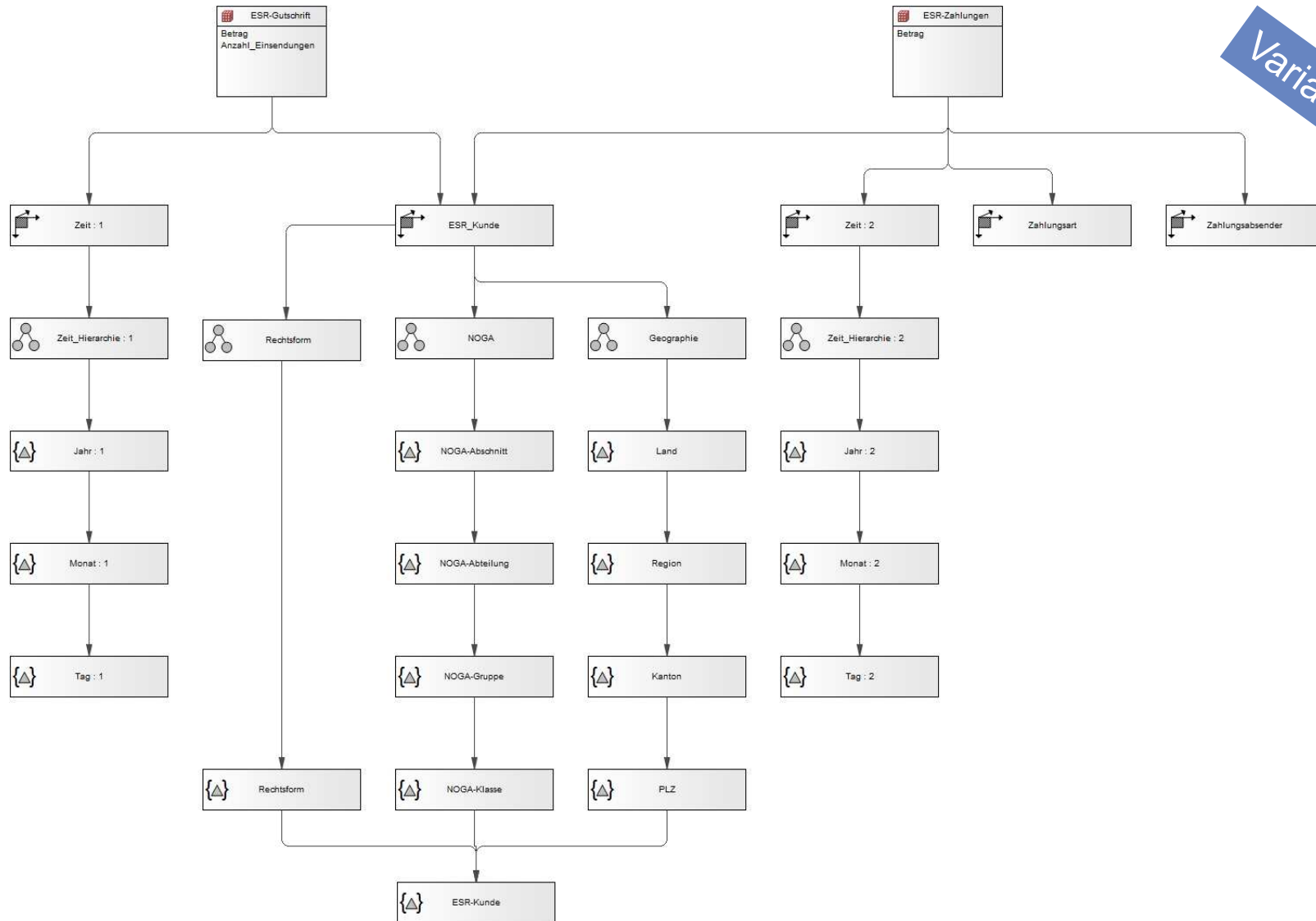
Heutige Belege (rote und orange Einzahlungsscheine) werden durch einen einheitlichen neuen Einzahlungsschein ersetzt.

Der neue Einzahlungsschein mit Datencode erleichtert für Kunden die Zahlungsabwicklung.

Die Harmonisierung betrifft alle zentralen Bereiche und Verfahren des Zahlungsverkehrs:

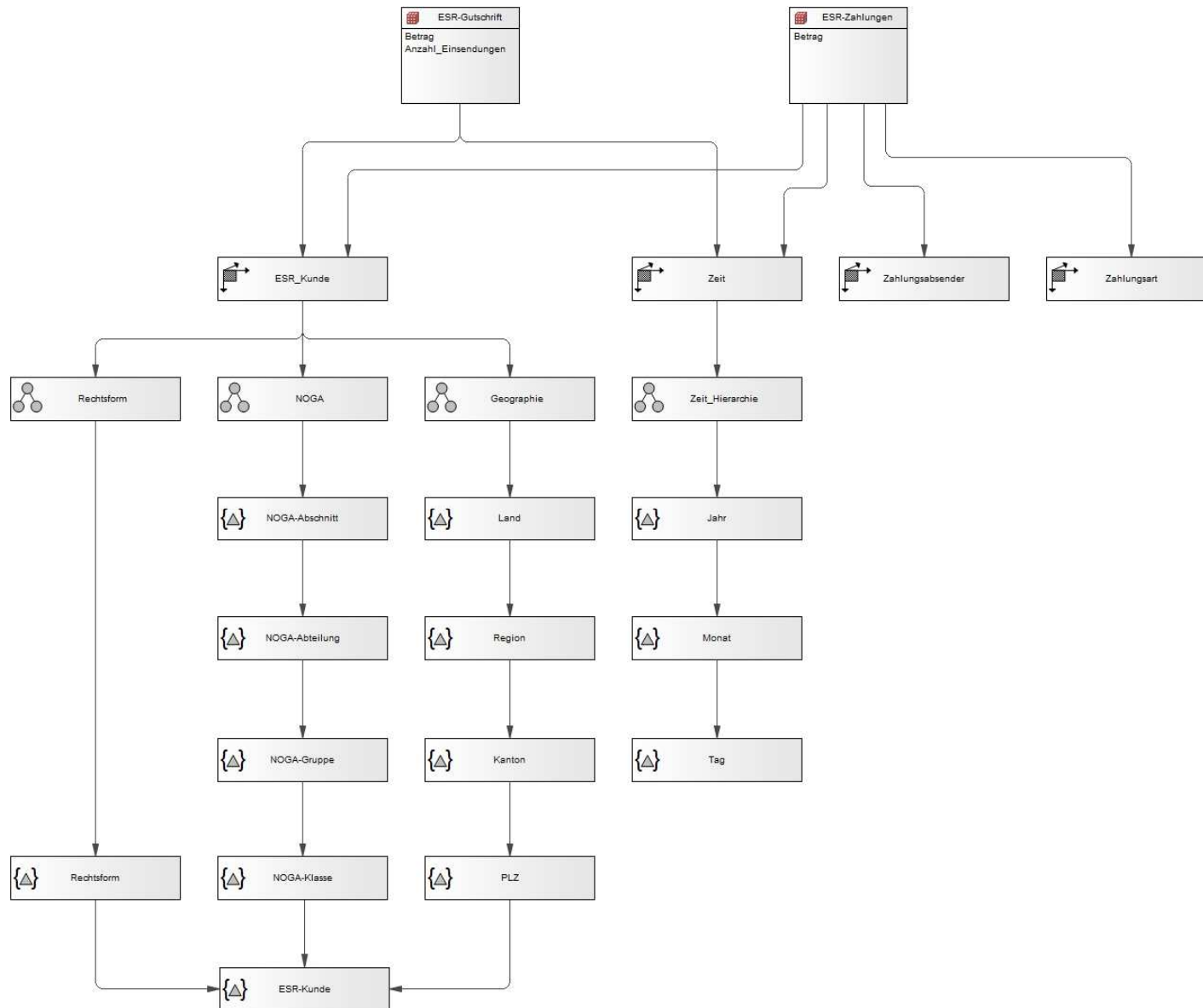
- Überweisungen
- Lastschriften (mit Integration in E-Rechnungsprozesse)
- Avisierung & Reporting
- Einzahlungsscheine

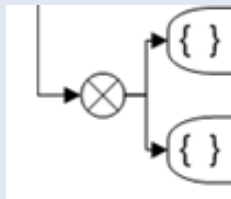
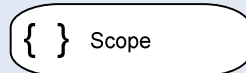
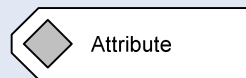
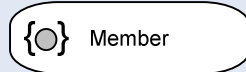
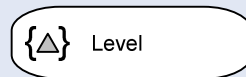
Cube – Modellierung (Ausschnitt ZV)



Cube – Modellierung (Ausschnitt ZV)

Variante 2





- ADAPT dient zur Darstellung eines betriebswirtschaftlichen Kontextes
- Hierarchien – ob ein-/mehrstufig – bilden die Konsolidierungsstufen in Dimensionen der über die Fakten zu führenden Kennzahlen
- Die verschiedenartigen hierarchischen Dimensionsstrukturen lassen sich modellieren
- Dimensionen können die fachliche Relevanz der Verdichtungsstufen umfassend modelliert werden
 - Ebenen
 - Wertelisten
 - Scope (Unterteilung in Teilmengen)
 - Attributierung (Sortierung/Filterung)
 - Beziehungstypen für Dimensionsausschnitte

und ist somit auch ein zentraler Input für fachliche Metadaten, die auf der Datenzugriffsebene für die Anwender verfügbar sind



Reporting Repository



Generierung von ADAPT (+) Modellen in PowerDesigner

MDD – Cube Dokumentation in ADAPT



- **Konzeptionelles Modell**
 - Relational: Hierarchien die in einer Entität durch verschiedene Attribute abgebildet sind, sind nicht ersichtlich
 - Dimensional: Alle Hierarchien und auswerterelevante Attribute werden sichtbar gemacht

- **Logisches/Physisches Modell**
 - Dimensional (Star-Schema) wird eine Dimension mit allen Hierarchien in eine physische Tabelle gespeichert.
 - Hierarchien werden also bewusst redundant gespeichert
 - Das bedeutet, dass in der Regel die Daten aus mehreren Core-Tabellen zusammen geführt werden
 - Vorteil: Einfacher und schneller Lesezugriff
 - Nachteil: Hierarchien nicht offensichtlich

Modellierung ist das Zeichnen
basierend auf einem exakten Dialog

