

# Data Warehouse und Business Intelligence

Grundlagen und  
sd&m-Projekterfahrungen

Siegen, 19.01.2007  
André Speulmanns

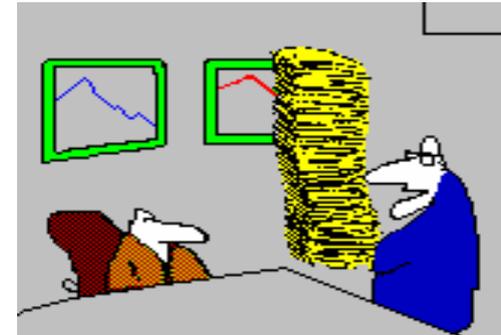
**sd&m**  
A Company of  Capgemini

UNIVERSITÄT SIEGEN



# Die Aufgabe

„Ich brauche mal schnell folgende Berichte, die das Wichtigste auf einer Seite zusammenfassen!“



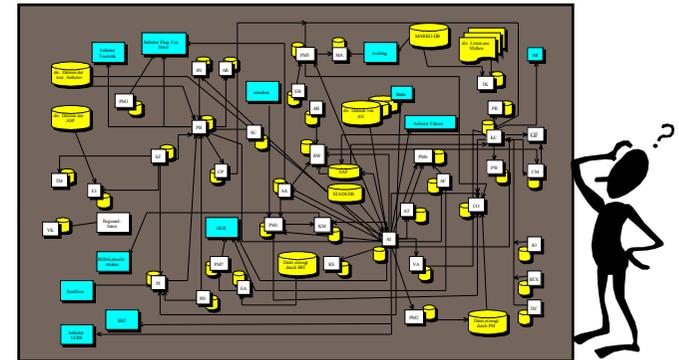
## ■ Typische Fragestellungen

- Kaufverhalten                      Wer ist seit > 1,5 Jahren Kunde und hat keines unserer Top-Produkte?
- Markttrends                        Wie kommt der neue Tarif auf dem Markt an?
- Qualitätssituation                Wie lange muss ein Anrufer in der Warteschleife warten?
- Geschäftsperformance            Wie lange dauert eine Adressänderung?

## ■ Ausgangsbasis: operative Daten des Tagesgeschäfts

# Das Problem

- Daten sind auf viele Systeme verstreut (Heterogenität).
  - Komplexität
  - Know How für die Datenextraktion liegt bei den jeweiligen Systemexperten
  - Daten sind z. T. nicht vergleichbar, z. B. aufgrund uneinheitlicher Begriffsbildung (Inkompatibilität)
- Daten sind teilweise schlecht gepflegt (mangelnde Datenqualität).
- Es gibt keinen zentralen Ort zur Speicherung externer Zahlen (z. B. Marktforschungen oder Daten über Konkurrenten).
- Das vorhandene Standardberichtswesen ist unflexibel, die Umsetzung von Änderungen dauert zu lange.
- Speicherung auf Belegpositionsebene ist für hochaggregierte Berichte ungeeignet (falscher Aggregationsgrad).
- Komplexe Berichte belasten operativen Systeme (Performanz).



# Die Lösungsidee

- Daten aus den verschiedensten operativen Systemen plus Daten aus externen Quellen sind zentral für Berichtszwecke verfügbar
- Daten können dort eingesammelt werden wie Produkte in einem Warenhaus:
  - Selbstbedienung (Abfragen ohne DV-Kenntnisse)
  - große Auswahl
  - alles sofort und schnell lieferbar
  - komplettes Sortiment, viele Fachgeschäfte in einem
  - günstige Preise
- Operative Systeme werden durch Abfragen nicht belastet
- Zielgruppe: Manager / Entscheidungsträger



 Ein Datenwarenhaus!



## Agenda

---

### → **Vorstellung**

- Begriffsklärung: DWH und BI
- Technische Grundlagen
- sd&m-Erfahrungen im Umgang mit BI-Projekten
- Aktuelle Themen und Ausblick

# Vorstellung

- Studium der Informatik mit Nebenfach BWL an der RWTH Aachen, vorher 6 Semester Physik
- seit 11 Jahren bei sd&m
- alle Projektphasen und Rollen vom Entwickler bis zum Projektleiter

- Schwerpunkte im Bereich Data Warehouse
  - Konzeption
  - Technologie (Theorie, sd&m-Ladeschicht, SAP BW, Frontend-Tools)
  - Projektleitung (Vorgehensweise in einem Data Warehouse-Projekt)

André Speulmanns  
Managing Consultant

sd&m AG  
software design & management  
Wanheimer Str. 68  
40468 Düsseldorf, Germany  
Direct +49 211 56623-208, Fax -50  
andre.speulmanns@sdm.de  
www.sdm.com

**sd&m**

A Company of

 **Capgemini**  
CONSULTING. TECHNOLOGY. OUTSOURCING

# sd&m AG – software design & management

## Geschäftsfelder

- Entwicklung und Integration maßgeschneiderter Informationssysteme für unternehmenskritische Prozesse
- IT-Beratung mit Umsetzungskompetenz

## Kunden

- Namhafte Unternehmen und Organisationen, die durch Einsatz individueller Lösungen Wettbewerbsvorteile erlangen

## Kernkompetenz

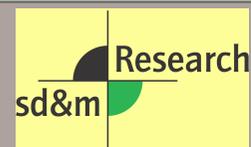
- Software-Engineering und Projektmanagement



## Eckdaten 2006

- Mitarbeiter: ca. 1.200
- Umsatz: 184 Mio. €

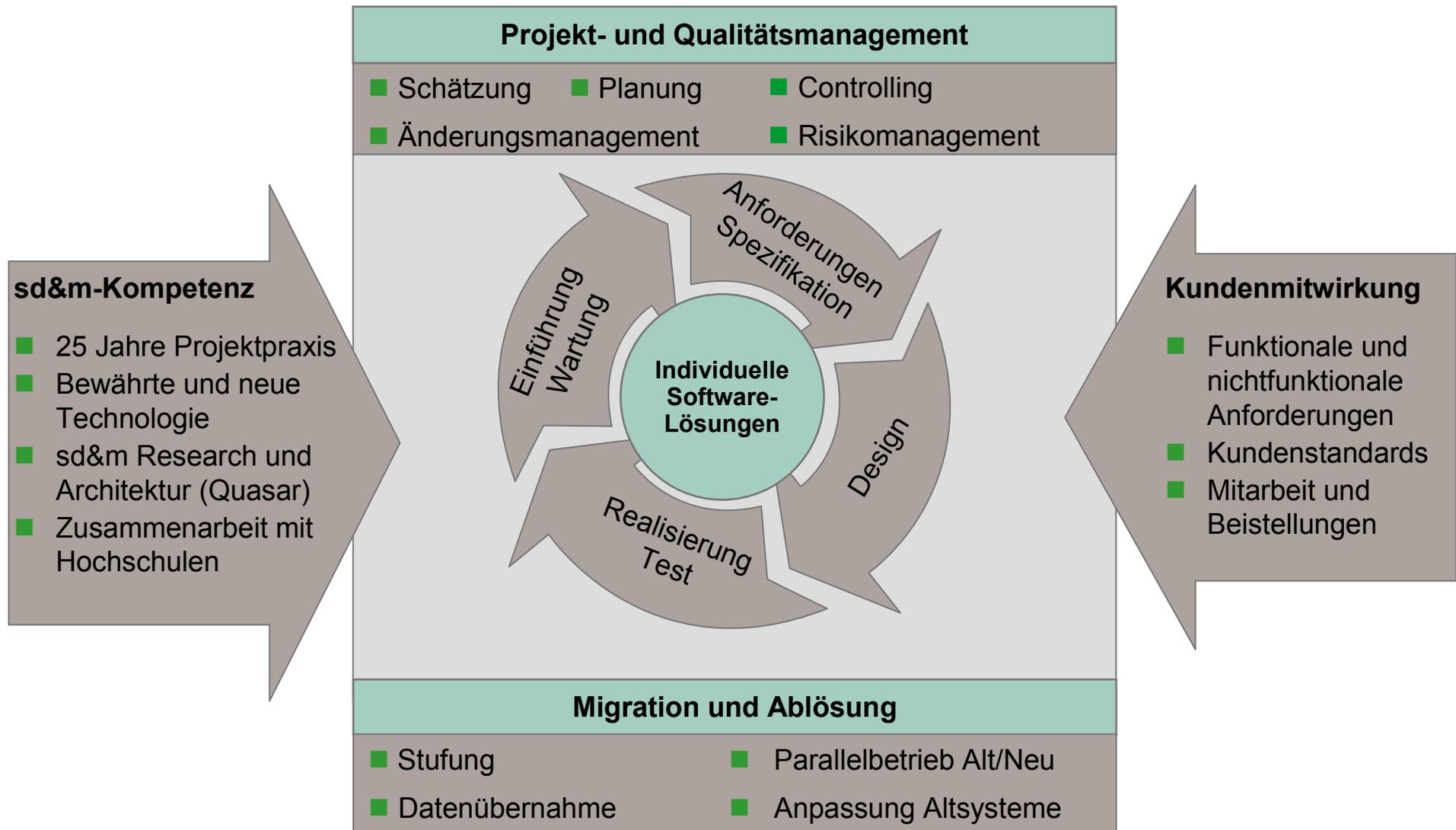
## Forschung



## Aktionär

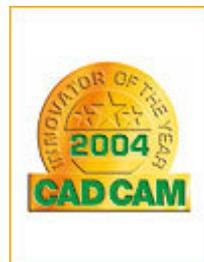


# Die Kernkompetenz von sd&m liegt in der Fähigkeit, (schwierige) IT-Projekte gut zu machen



## Aspekte der sd&m-Unternehmenskultur

- Streben nach Spitzenleistung, Ehrlichkeit, Wirtschaftlichkeit
- Attraktiv für Kunden und Mitarbeiter
- Alle Führungskräfte kommen aus der Projektarbeit
- Alle Führungskräfte akquirieren
- Es gibt keinen „Personalchef“
- Offene Türen und duzen bis zum Vorstand
- Kaffeeküchenkultur





## Agenda

---

- Vorstellung
- **Begriffsklärung: DWH und BI**
- Technische Grundlagen
- sd&m-Erfahrungen im Umgang mit BI-Projekten
- Aktuelle Themen und Ausblick

## Definition

Nicht jede Datensammlung ist ein Data Warehouse!

**"Ein Data Warehouse ist eine**

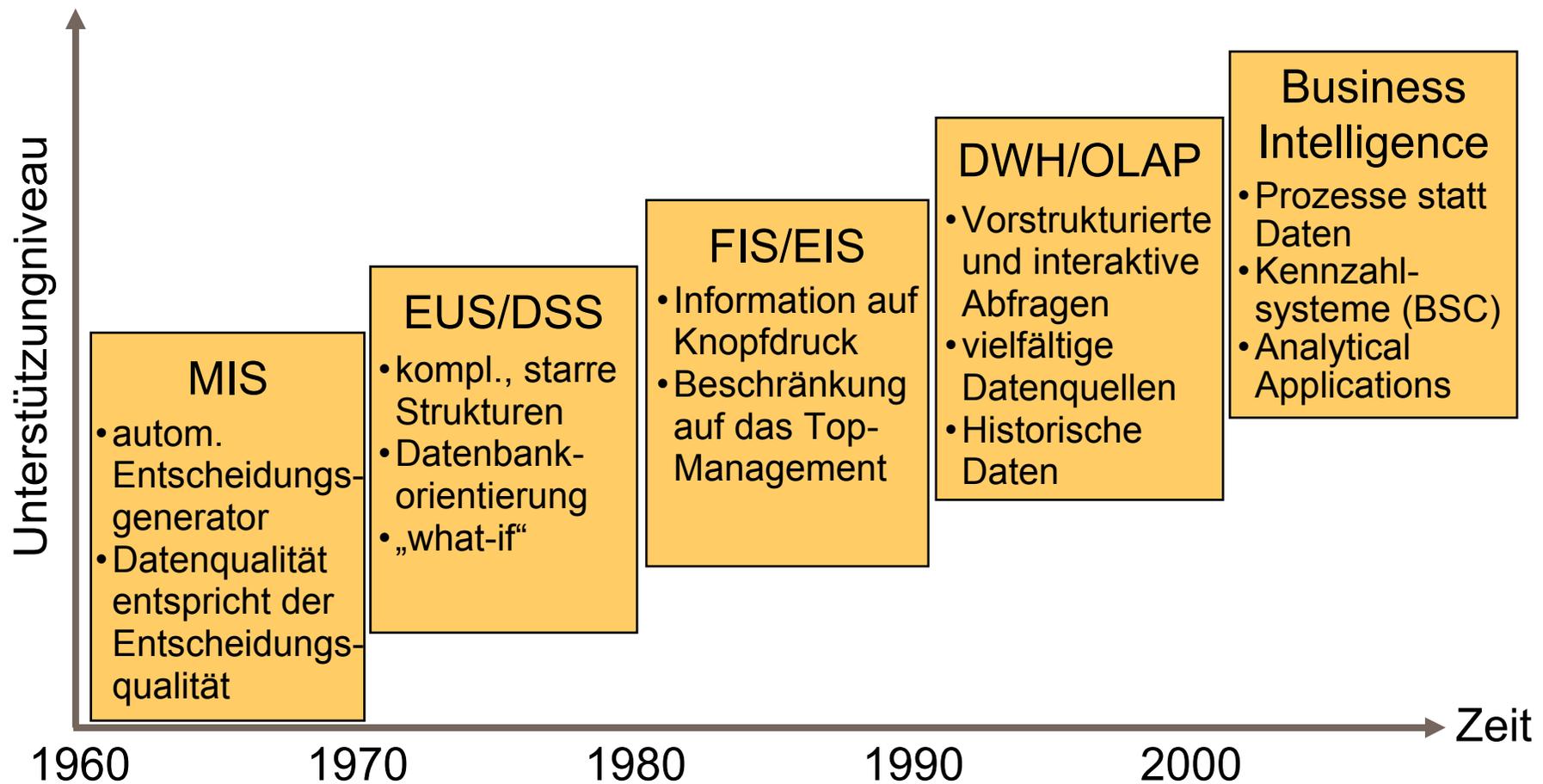
- **themenorientierte** *Alle ins Data Warehouse übernommenen Daten lassen sich bestimmten Themen zuordnen (z. B. Umsatz pro Kunde, Vertriebsseinheit und Produkt)*
- **zeitorientierte** *Zeitreihenanalysen sind möglich (Dimension „Zeit“)*
- **integrierte** *Verteilte Daten werden konsolidiert*
- **unveränderliche** *Gleiche Abfragen liefern zu jedem Zeitpunkt gleiche Ergebnisse (Dauerhafte Speicherung, keine Änderung an bestehenden Daten)*

**Datensammlung, deren Daten sich für Managemententscheidungen auswerten lassen."**

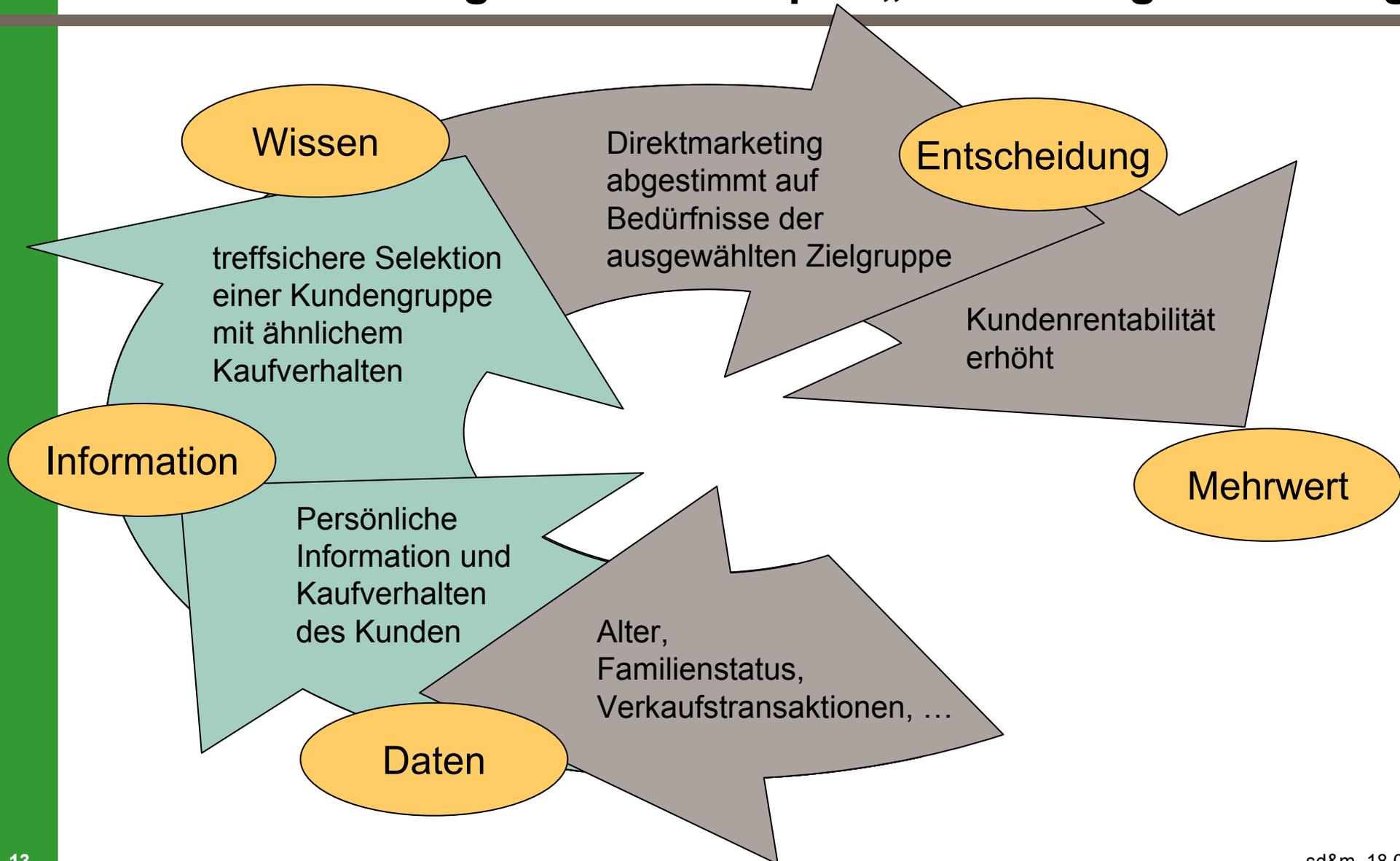
(Inmon, "Building the Data Warehouse")

→ Konsequenz: Es müssen zum Teil ganz andere Prinzipien beachtet werden als für transaktionsorientierte Systeme (z. B. beim Datenmodell)

# Die Werkzeuge zur Managementunterstützung werden kontinuierlich besser und ändern regelmäßig den Namen



## Business Intelligence am Beispiel „Kundensegmentierung“



## „Definition“ von Business Intelligence (BI)

*Business Intelligence ist der Prozess, der Daten in Informationen und weiter in Wissen umwandelt (Gartner Group)*

*Unternehmensentscheidungen und Prognosen stützen sich auf dieses Wissen.*

*Business Intelligence umfasst ein breites Spektrum an Anwendungen und Technologien und ist bei sd&m der Oberbegriff für Data Warehousing, Data Mining, Online Analytical Processing und Analytical Applications.*

- Während Business Intelligence die Unterstützung des Geschäfts ausdrückt, wird Data Warehousing mitunter als rein technische Bereitstellung von Daten gesehen („Datenfriedhof“).
- Aus unserer Sicht müssen alle Komponenten über fachliche und technische Aspekte optimal zusammenspielen.



## Agenda

---

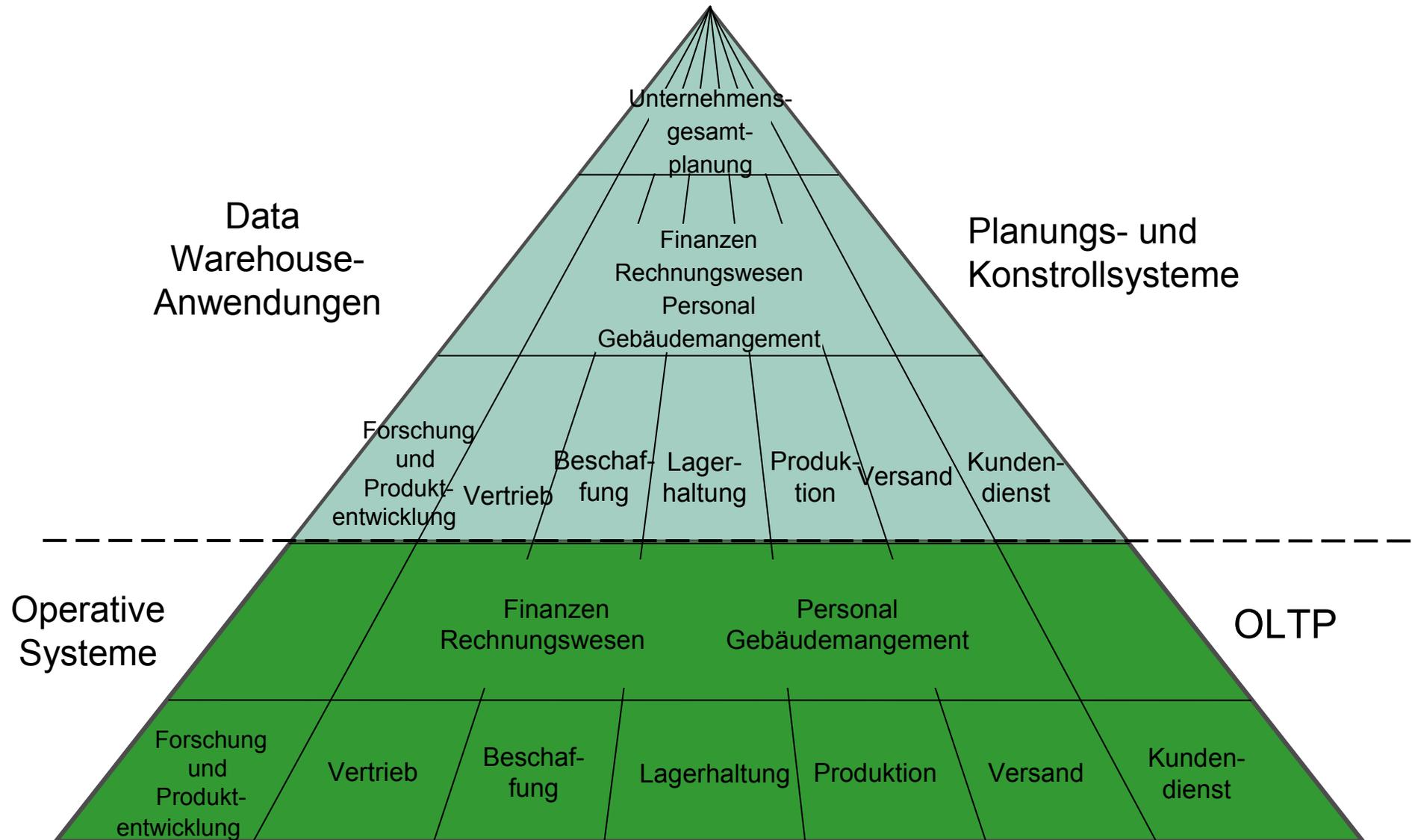
- Vorstellung
- Begriffsklärung: DWH und BI
- **Technische Grundlagen**
- sd&m-Erfahrungen im Umgang mit BI-Projekten
- Aktuelle Themen und Ausblick

# Unterschiedliche Anforderungen an operative Systeme und Data Warehouses

	Operative Systeme	Data Warehouse Systeme
<b>Datenbestand:</b>	operative Daten: <ul style="list-style-type: none"> <li>■ in der Regel aktuelle Daten</li> <li>■ oft nur für abgegrenzten Unternehmens- und Funktionsbereich</li> <li>■ Unterst. des op. Geschäftsprozesses</li> </ul>	Analysedaten: <ul style="list-style-type: none"> <li>■ historisch, verdichtet, aufbereitet</li> <li>■ Kombination von Daten mehrerer Systeme</li> </ul>
<b>Datenformate:</b>	<ul style="list-style-type: none"> <li>■ Datenbestände oft unkoordiniert gewachsen</li> <li>■ unterschiedliche Formate</li> <li>■ kein einheitliches Datenmodell</li> </ul>	<ul style="list-style-type: none"> <li>■ einheitliches Datenmodell</li> <li>■ einheitliches Format</li> <li>■ semantisch angepasst</li> </ul>
<b>Datenzugriff:</b>	<ul style="list-style-type: none"> <li>■ satzorientiert, „vorgegeben“</li> <li>■ Kurze Transaktionszeiten</li> </ul>	<ul style="list-style-type: none"> <li>■ benutzerdef., komplexe Analysen</li> <li>■ multidimensional, Drill-down, Roll-up</li> </ul>

- Die Trennung operativer Daten von Analysedaten ist ein fundamentales Konzept eines Data Warehouses

# DWH-Anwendungen bauen in der IT-Landschaft auf den OLTP-Anwendungen auf



# Kernkomponenten eines Data Warehouse

- ETL
  - Extraktion aus den Quellsystemen
  - Transformieren und Bereinigen der Daten gemäß der Metadaten
  - Laden wahlweise komplett oder inkrementell durch Bestimmen der Differenz (“Delta”)
- Datenbank
  - Nutzdaten : relational („ROLAP“) als Sternschema oder multidimensional („MOLAP“)
  - Metadaten (statisch / dynamisch)
  - Performance-Tuning (Indizes, Hashing, ...)
  - Administration, Back-up und Recovery

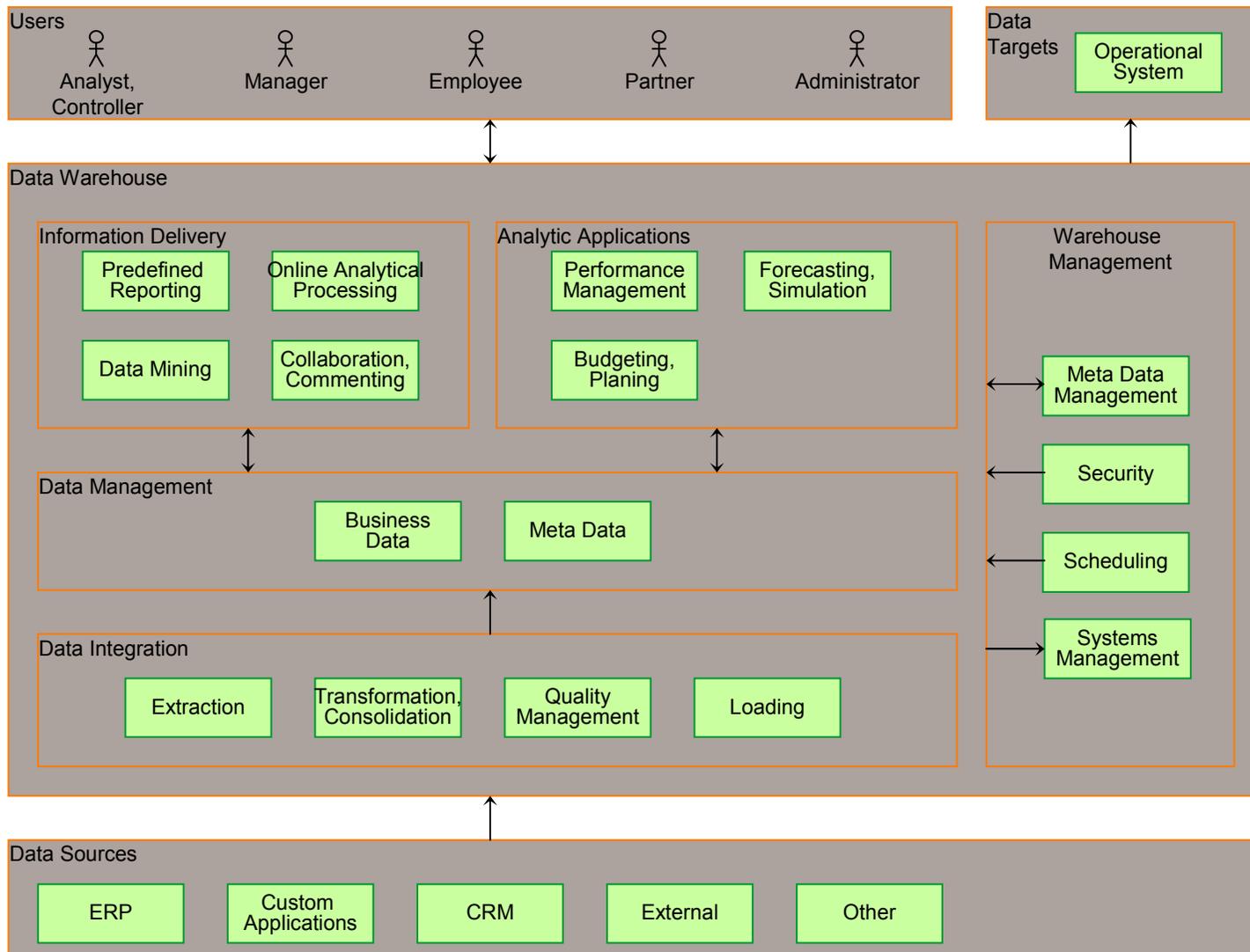
## Kernkomponenten eines Data Warehouse (2)

---

### Analyse und Präsentation

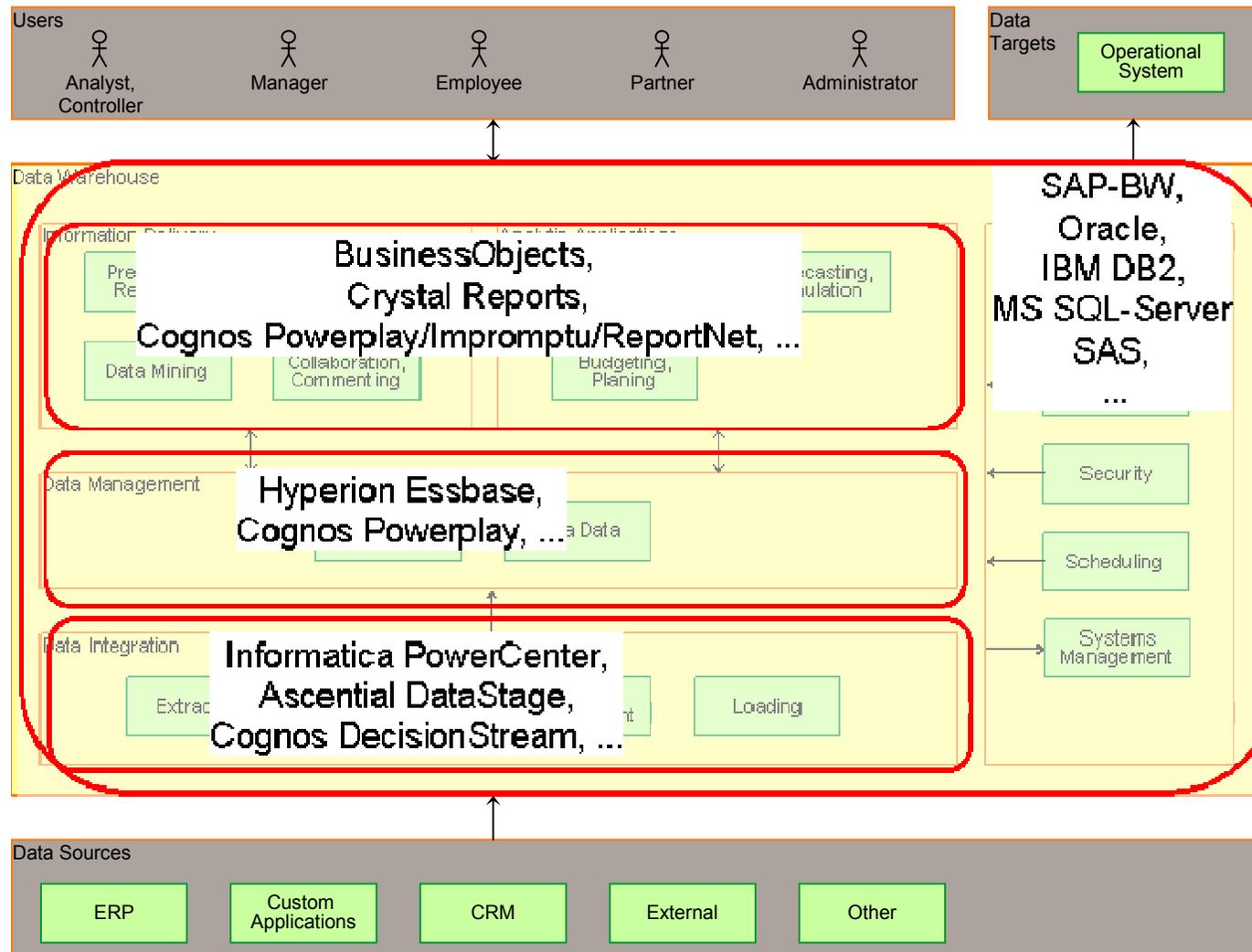
- Ad hoc - Auswertungen
  - Drill-down, Roll-up
  - Slicing and Dicing
- Vordefinierte Auswertungen (Scheduling)
- Graphische Darstellung
- Exception Highlighting
- Export in andere Dateiformate
- Intranet-Anbindung

# sd&m-Referenz-Architektur „i-BI“



- modular und serviceorientiert
- Services können von unterschiedlichen Produkten / Individuallösungen abgedeckt werden

# Mit Hilfe der Referenzarchitektur können die Produkte in einer BI-Produktlandkarte gegenübergestellt werden.



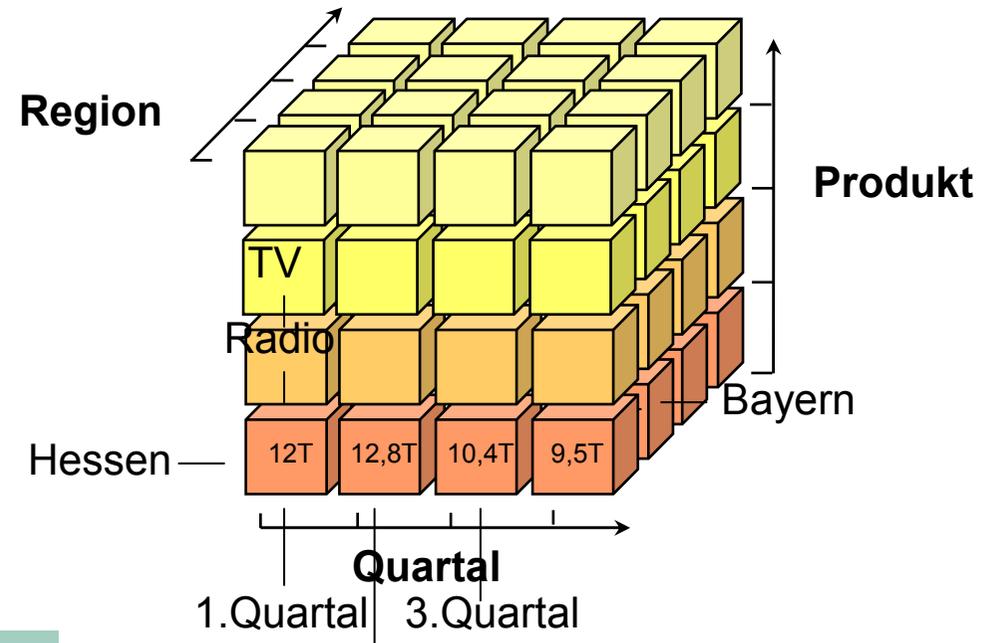
# Multidimensionales Datenmodell

Beispielauswertung:

Produkt	Region	Quartal	Absatz
Radio	Hessen	1	12000
Radio	Hessen	2	12800
Radio	Hessen	3	10400
Radio	Hessen	4	9500
Radio	Bayern	1	17300
Radio	Bayern	2	18200
Radio	Bayern	3	17900
Radio	Bayern	4	17100

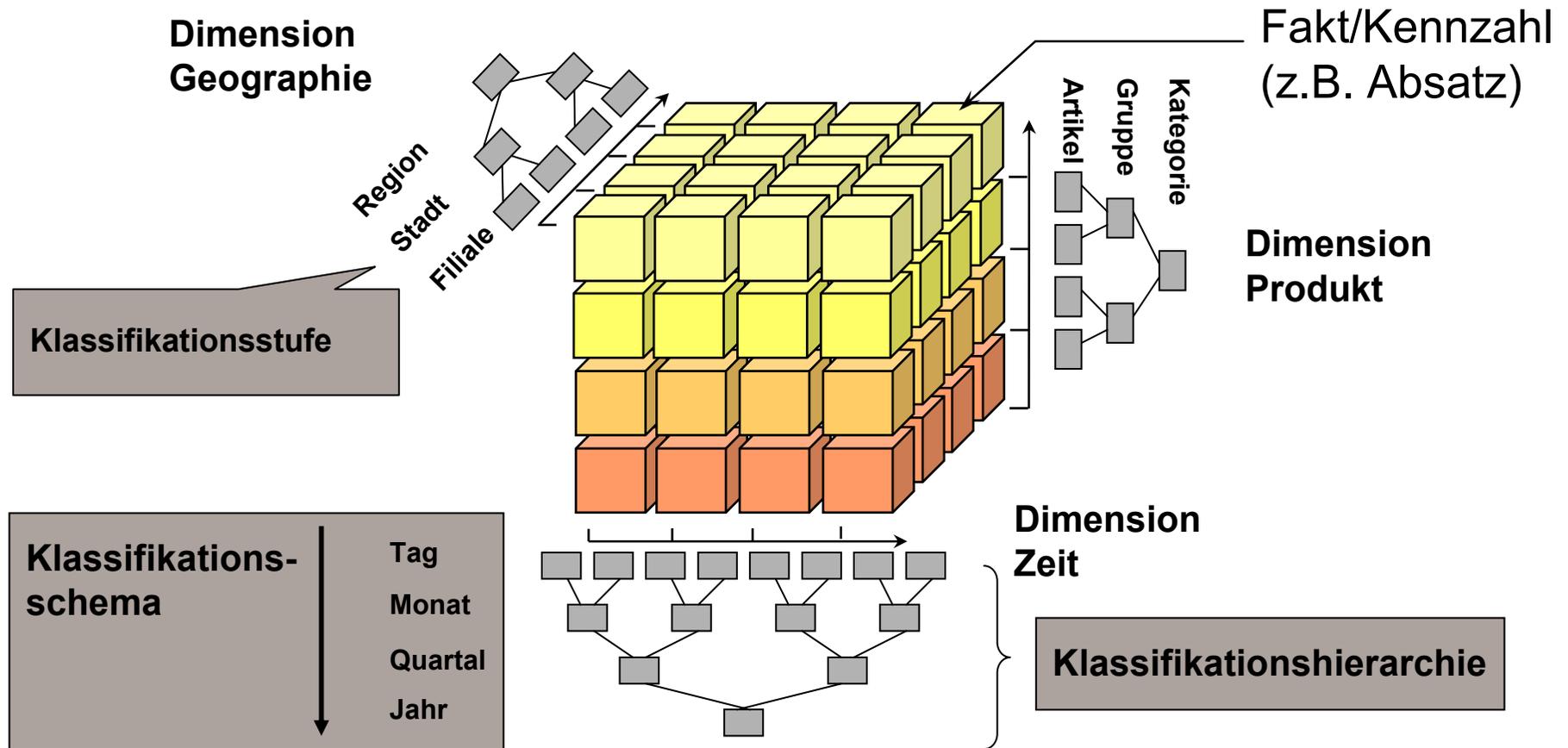
→ Fakten, Kenngrößen (z.B. Absatz)

→ Betrachtungsperspektiven (z.B. Region)



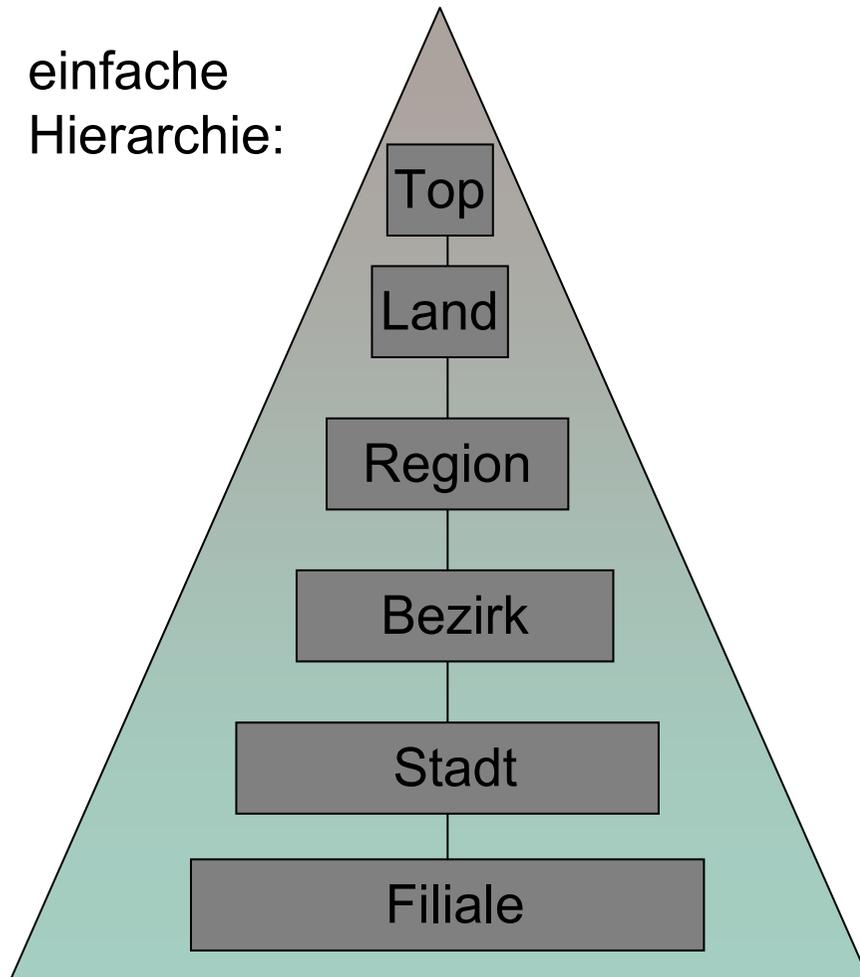
Faktenwürfel wird durch Dimensionen aufgespannt.

# Multidimensionale Daten

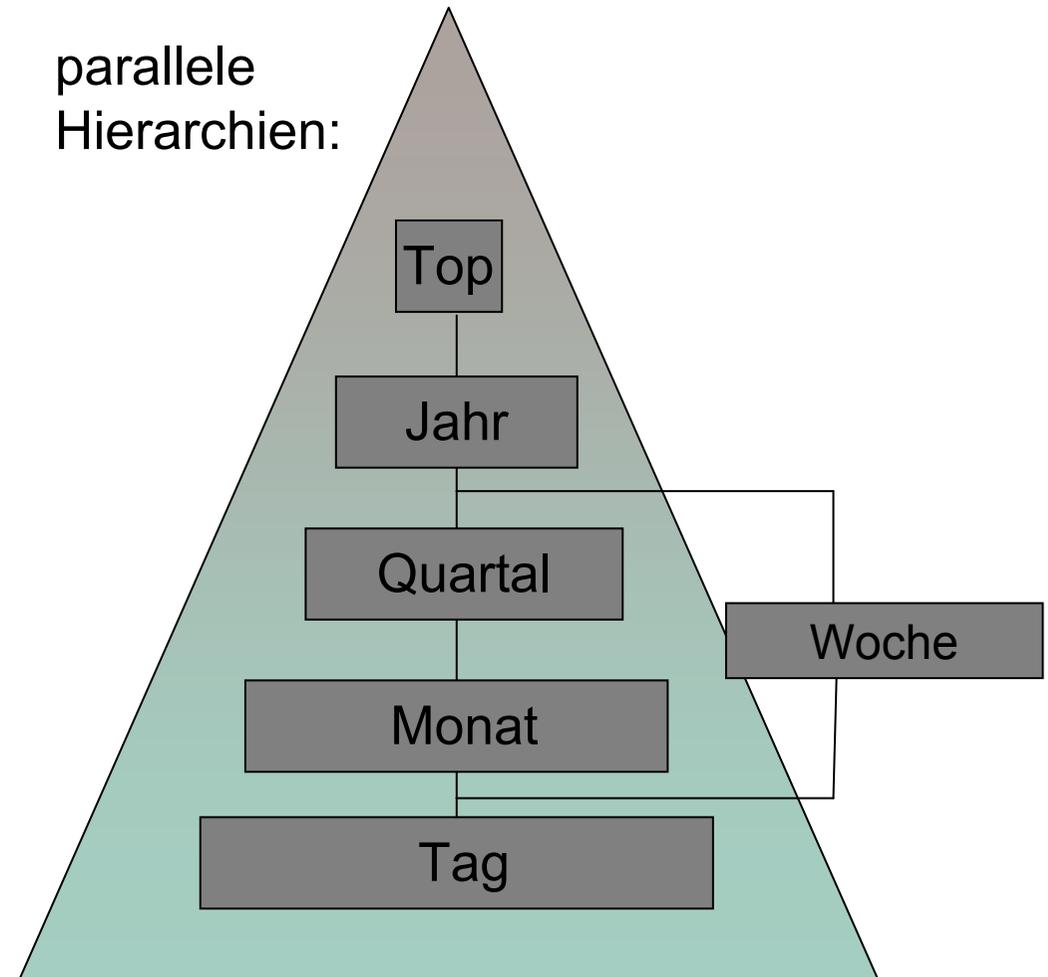


## Dimensionen: einfache vs. parallele Hierarchien

einfache Hierarchie:



parallele Hierarchien:



## Dimensionen und Fakten

1/3

- Fakten entsprechen Kennzahlen, die für eine bestimmte Kombination von Dimensionswerten Gültigkeit haben. Beispiel: „Erlös der Produktgruppe Kühlschränke im Monat Dezember 2000 in München“
- Eine typische Abfrage auf einem Data Warehouse-Datenbestand besteht darin, eine oder mehrere Kennzahlen bezüglich einer bestimmten Dimensionskombination darzustellen, wobei
  - bzgl. einiger Dimensionen eingeschränkt wird
  - bzgl. anderer Dimensionen aufgerissen wird (Zeilen oder Spalten)
  - bzgl. der restlichen Dimensionen aggregiert wird

## Dimensionen und Fakten

2/3

Dimensionen sind endliche Wertebereiche

■ typische Dimensionen sind

- „Zeit“, einzelne Monate oder auch Kalendertage eines definierten Zeitraums
- „Ort“, z.B. eine Postadresse

Aber auch Kategorien, z. B.

- Produktgruppe:      Kühlgeräte, Fernseher, ...
- Werttyp:              Plan, Ist, Soll

■ Dimensionen sind häufig hierarchisch organisiert (Tag -> Monat -> Quartal -> Jahr) und können auch noch weitere Merkmale haben (Kalenderwoche, Wochentag, Feiertag)

## Dimensionen und Fakten

3/3

- Eine Kennzahl ist *aggregierbar bezüglich einer Dimension*, wenn es eine fachlich sinnvolle (kommutative und assoziative) Aggregationsfunktion gibt (zumeist Summe, gelegentlich auch Mittelwertbildung).
- Beispiel:
  - Der Umsatz einer Firma ist bezüglich der Dimensionen Zeit, Verkaufsorganisation und Verkaufsort sinnvoll aggregierbar durch Summenbildung.
  - Der Kundenbestand ist durch Summierung zwar bezüglich der Verkaufsorganisation aggregierbar, nicht jedoch bezüglich der Zeitdimension; hier ist jedoch eine Mittelwertbildung sinnvoll
- Entsprechend „verdichtete“ Ergebnisse heißen *Aggregate*. Das OLAP-Werkzeug (sowohl MOLAP als auch ROLAP) unterstützt den Designer bei der Bildung von Aggregaten.

# Das Stern-Schema ermöglicht durch hohe Redundanz schnelle Zugriffszeiten



© Ralph Kimball

aggregieren

Summe verkaufter Einheiten von Produkt x an Freitagen im Mai 2000, nach Verkäufern aufgeteilt:

```
SELECT  ClerkName,
        SUM (UnitsSold)
```

einschränken

```
FROM  Fact f,
      TimeDimension t,
      ClerkDimension c,
      ProductDimension p
```

aufreißen

```
WHERE f.Time_key = t.time_key
AND f.Product_key=p.Product_key
AND f.Clerk_Key=c.Clerk_key
AND t.DayOfWeek='Freitag'
AND p.Description = 'x'
AND t.FiscalPeriod = '2000/05'
```

```
GROUP BY c.ClerkName
```

## Begriffe

---

- **OLTP** = Online Transaction Processing  
“Die operativen Geschäftsprozesse und IT-Systeme eines Unternehmens“
- **OLAP** = Online Analytical Processing  
“Analyse und Auswertung von multidimensional aufbereiteten Daten, um Informationen für Unternehmensentscheidungen zu gewinnen”
- **MOLAP** = Multidimensional OLAP  
„OLAP auf Basis einer proprietären, multidimensionalen Datenbank“
- **ROLAP** = Relational OLAP  
„OLAP auf Basis einer relationalen Datenbank“
- **Dimension, Hierarchie, Würfel, multidimensionales Datenmodell, Drill Down, ...**

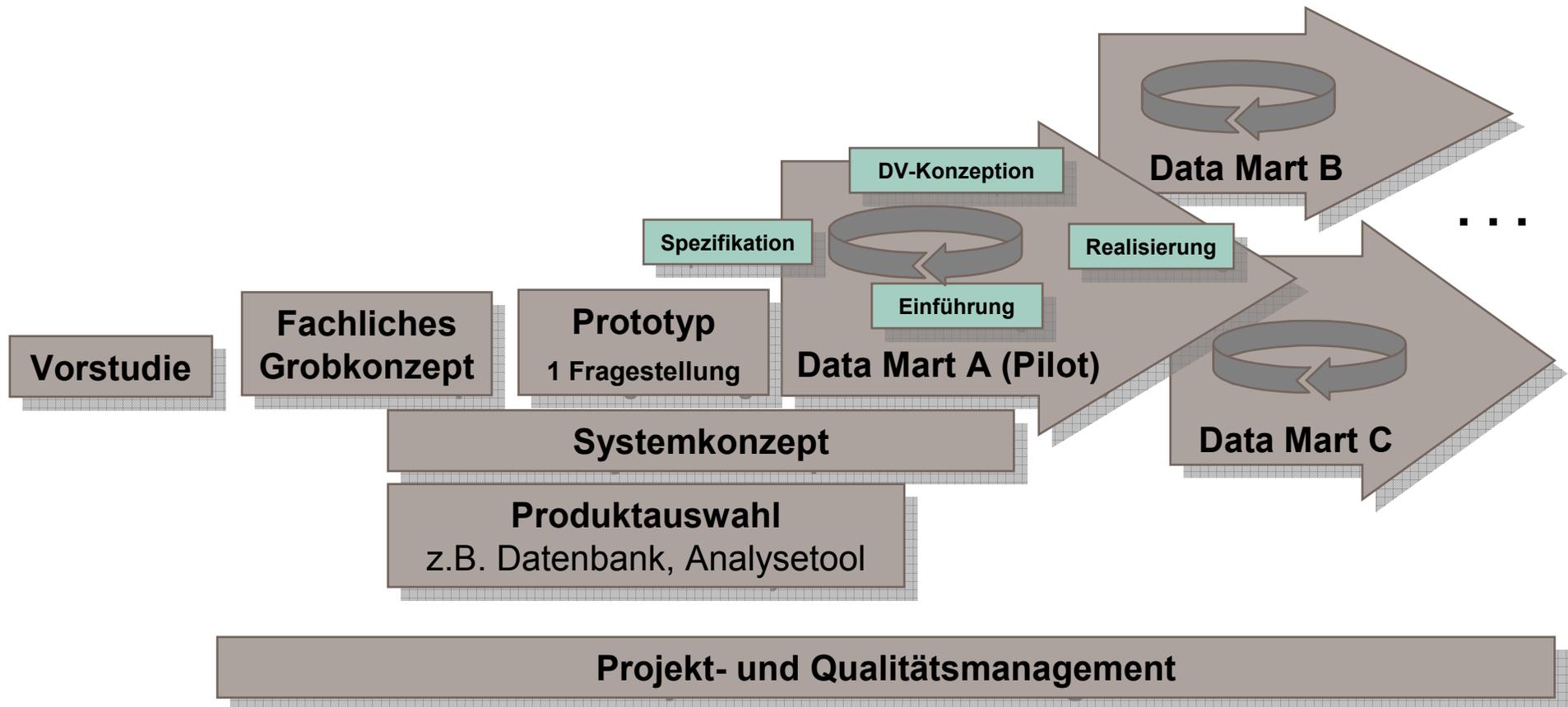


## Agenda

---

- Vorstellung
- Begriffsklärung: DWH und BI
- Technische Grundlagen
- **sd&m-Erfahrungen im Umgang mit BI-Projekten**
- Aktuelle Themen und Ausblick

# Data Warehouse-Projekte gemäß sd&m-Vorgehensmodell: Think big, start small ...



## Business-Intelligence-Projekte sind komplex ... gängige Fehler und wie sie zu vermeiden sind (1)

Beginne nicht mit der  
Tool-Auswahl

- Kläre erst, was du mit Business Intelligence machen willst
- Wichtig ist zunächst die fachliche Sicht
- Die Nutzungsweise kann z. B. über ROLAP vs. MOLAP entscheiden

Produkte sind nicht der  
Schlüssel zum Erfolg

- Gute und passende Tools sowie deren reibungsloses Zusammenspiel sind eine notwendige Voraussetzung
- Für den Erfolg ist jedoch die Erfüllung der Benutzeranforderungen maßgebend
- Der Trend geht von Best-of-Breed zu Tool-Suiten

Das Projektteam  
braucht mehr als Tool-  
Spezialisten

- Orientiere die Teambildung nicht nur an den eingesetzten Tools
- Flexibler sind kleine, fachlich orientierte Teilteams, die alle Ebenen der Architektur beherrschen, koordiniert von einem Chef-Designer
- Ein Team sollte ein Thema über alle Projektphasen bearbeiten

## Business-Intelligence-Projekte sind komplex ... gängige Fehler und wie sie zu vermeiden sind (2)

"Think big – start small"

- Streben nach der allumfassenden Lösung bremst das Projekt
- Ein Beginn mit globaler Erhebung und Spezifikation führt zu spätem Anwender-Feedback
- Ein schrittweises Vorgehen, bevorzugt nach fachlichen Teilbereichen, weniger nach technischen Gesichtspunkten ist sinnvoll

Bau nicht das alte Reporting nach

- Anwender wünschen zuerst immer, dass alles wie früher aussehen soll – schade um die Chance!
- Ein frischer Prototyp bringt alte Anwender auf neue Ideen
- Neue Geschäftsprozesse erfordern neue Analysen

Ein BI-Vorhaben ist kein Projekt, sondern ein Programm

- Die vielfältigen Möglichkeiten mit BI-Technologien sind bei Anwendern zunächst weitgehend unbekannt
- Der "Aha-Effekt" beflügelt die Fantasie, führt zu neuen Nutzern und Anforderungen
- "A Data Warehouse is a journey, not a destination...!"

## Business-Intelligence-Projekte sind komplex ... gängige Fehler und wie sie zu vermeiden sind (3)

Berichte für das Top-Management sind nicht alle automatisierbar

- Das Top-Management erwartet individuelle Zusammenstellungen vielfältiger Informationen, mit Bezug auf wechselnde Fragestellungen
- Diese stellt ein Sachbearbeiter individuell zusammen
- Basis hierfür können und sollten Daten aus einem DWH sein

Lass nur bewusst fremde Daten rein!

- Die Bedeutung importierter Daten muss genau verstanden sein: Was sagt uns diese Zahl?
- Das gilt auch für die Transformationen auf dem Weg ins DWH!
- Auch der Anwender muss die Bedeutung der Zahl verstehen können

OLAP spielt eine geringere Rolle als angenommen

- OLAP-Funktionen (drill, slice, dice) werden von einer vergleichsweise kleinen Gruppe der Power User genutzt
- Viele DWH-Auszüge laufen per Standard-Reporting
- Der eigentliche DWH-Nutzen liegt in der Integration und Harmonisierung der Daten

## Business-Intelligence-Projekte sind komplex ... gängige Fehler und wie sie zu vermeiden sind (4)

Bau keine  
Monster-Würfel

- Mehr als 4 Dimensionen sind häufig schwer verständlich
- Multidimensionale DB haben Grenzen (Performance / Sparsity)
- Die fachliche Motivation für große Würfel ist gut zu hinterfragen

Garbage in –  
Garbage out

- Der Erfolg einer BI-Lösung steht und fällt mit der Qualität der Daten aus den Quellsystemen
- Prüfe diese frühzeitig und setze erforderliche Maßnahmen auf (Datenbereinigungen, Plausibilisierungen, ...)

Vermeide "Moving  
Targets"

- Ist das Quellsystem noch in der Entwicklung, läufst du jeder Änderung hinterher
- Beginne erst mit dem DWH, wenn sich das Quellsystem stabilisiert hat

Schnelligkeit geht oft  
vor Schönheit

- Ein schneller Proof-of-Concept macht Lust auf mehr
- Schlechte Performance bringt unzufriedene Anwender
- Achte auf Skalierbarkeit – die Datenbank wächst schnell!
- Komplexe Funktionalität, einfache Bedienbarkeit, Integration mit anderen Produkten sind eher zweitrangig

## Zentrale Erfolgsfaktoren für BI-Projekte

- ▶ Ohne einen von allen Parteien akzeptierten und entscheidungsfreudigen **Projektsponsor** aus dem Topmanagement sind BI-Vorhaben oft zum Scheitern verurteilt.
- ▶ BI-Lösungen können sehr verschieden sein und somit auch leicht am konkreten Bedarf vorbei zielen. Ein **klar definierter erwarteter Nutzen** hilft beim Erreichen realistischer Ziele.
- ▶ Die Anforderungsanalyse erfordert nicht nur die Aufnahme von Anforderungen, sondern auch eine **Hilfestellung für die Anwender** durch erfahrene DWH-Projektmitarbeiter.
- ▶ Bereits in der Projektplanung ist die **Betriebsphase zu berücksichtigen** (Wartungs- / Bedienungsfreundlichkeit, neue Betriebsaufgaben, Portaleinbindung, Support, Ablösung BI-Ressourcen, etc.).
- ▶ Am besten über einen **Projektleiter aus der Fachabteilung** sollten die Endanwender früh in Produktwahl und Systemdesign (Prototyping) einbezogen werden.
- ▶ Ein offensives **internes Marketing und klare Zuständigkeiten** helfen Interessenkonflikte aufzulösen (Monopol- bzw, Aufgabenverluste, mögliche Kritik an operativer Datenqualität, etc.).
- ▶ Es gibt eine anerkannte **DWH/BI-Standardarchitektur**. Es ist sinnvoll, diese als Vorlage zu verwenden und an die spezifischen Erfordernisse anzupassen.
- ▶ Ein unternehmensweiter BI-Standard sollte u.a. **eindeutige Begriffsdefinitionen** (single point of truth) liefern.



## Agenda

---

- Vorstellung
- Begriffsklärung: DWH und BI
- Technische Grundlagen
- sd&m-Erfahrungen im Umgang mit BI-Projekten
- Aktuelle Themen und Ausblick:  
Data Mining und Righttime Analytics**

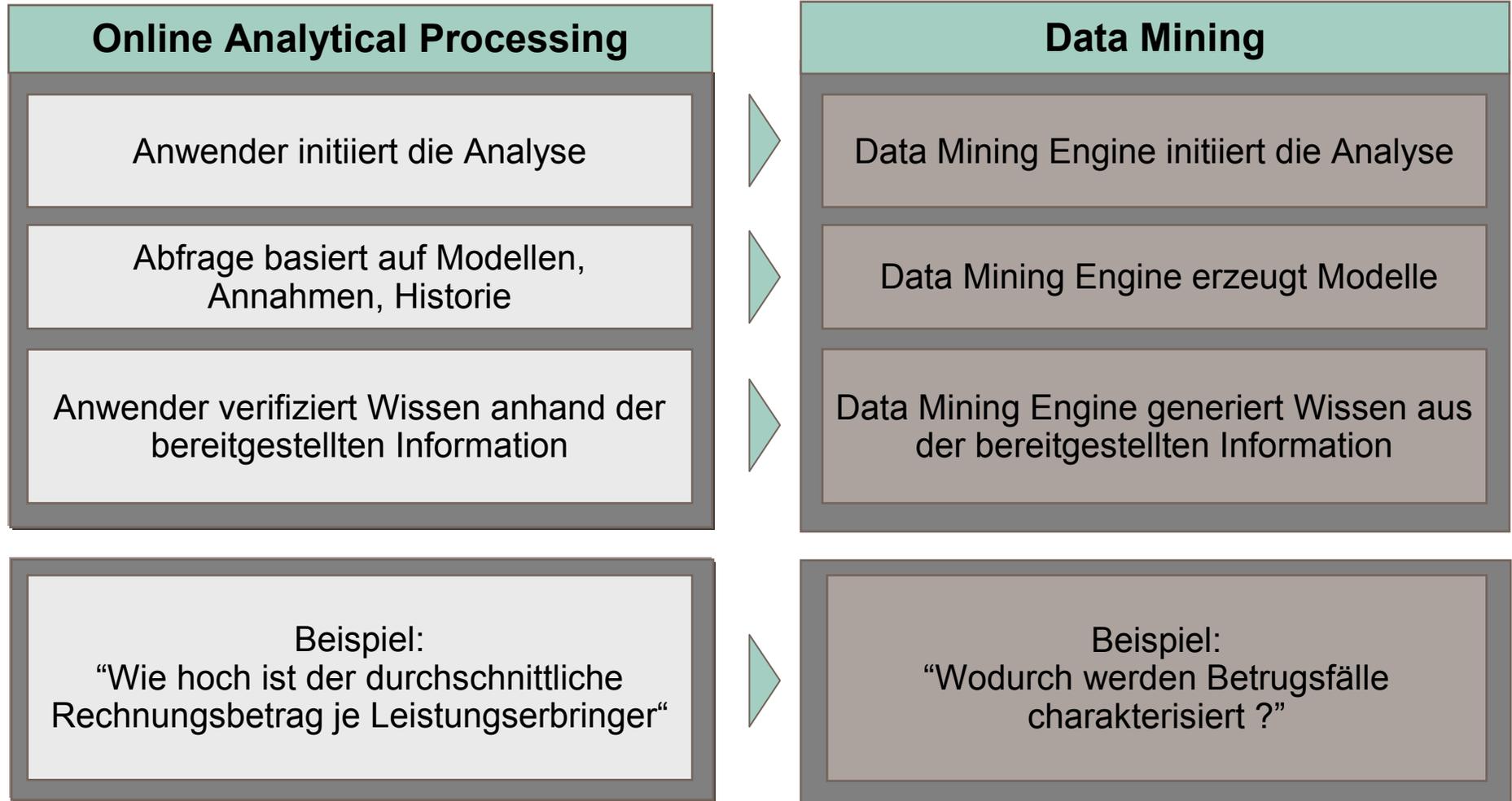
# Data Mining ist der dritte Schritt auf dem Weg zu Business Intelligence

„We are drowning in data, but starving for knowledge!“



„Now that we have gathered so much data, what do we do with it?“

# OLAP verifiziert Wissen, Data Mining generiert Wissen



# Data Mining wird in den verschiedensten Branchen für verschiedenste Zwecke eingesetzt

**Data Mining Anwendungsgebiete**

	Banken	Versicherungen	Versorgungunternehmen	Telekommunikation	Handel	E-Commerce	Medizin
Kunden segmentieren	•	•	•	•	•	•	•
Kündigungsprävention		•	•	•			
Neukunden gewinnen	•	•	•	•	•	•	
Kampagnen optimieren	•	•	•	•	•	•	
Betrug aufdecken		•	•			•	
Cross- und Up-Selling	•	•		•	•	•	
Prognosen erstellen		•	•				•
Prozessüberwachung			•	•			•
Risiko prüfen	•	•				•	•

# Righttime Analytics schafft die Basis für zeitnahe Analysen der Geschäftsentwicklung

## Was ist unter Righttime Analytics zu verstehen?

- Die Quellsysteme übermitteln die in den operativen Prozessen entstandenen und für das DWH relevanten Daten ohne Zeitverzug (realtime) bzw. zeitnah (neartime) an das DWH.
- Auf Veränderungen im Geschäft kann schneller reagiert werden. Zum Teil können Reaktionen im operativen System proaktiv vorbereitet werden und dann ohne manuelle Interaktion ausgelöst werden (Active DWH).
- Beispiele:
  - Werden Ressourcenengpässen und Mengenschwankungen zeitnah erkannt, so können noch im Laufe des Tages Gegenmaßnahmen eingeleitet werden.
  - Sind die Verkaufszahlen beim neuen Produkt des Internet-Shops besser als geplant, kann eine zeitnahe Preisanpassung durchgeführt werden.
- Righttime Analytics schafft die Entscheidungsgrundlage für operative und taktische Entscheidungen.
- Das eröffnet weitere Themenfelder nah an den operativen Systemen:
  - Die operativen Systeme müssen die Durchführung solcher Entscheidungen unterstützen.
  - righttime decisioning: Auf welcher Datenbasis muss ich mich entscheiden? Wie bringe ich die Datenbasis dabei ein? Immer allerneueste Infos berücksichtigen? Beispiele: Kundenwert-Entscheidung beim eingehenden Call im Call-Center.
- Viele verwandte Begriffe: Realtime-Reporting, Nearthime Analytics, Active DWH

# Righttime Analytics ermöglicht eine vielseitigere Verwendung der im DWH gesammelten Informationen

## Righttime Analytics ermöglicht die zusätzliche operative Nutzung der DWH-Informationen

### Klassisches DWH

- Dient als Grundlage für strategische Entscheidungen
- Zwischen der Ermittlung der Daten und der Umsetzung von Management-Entscheidungen stehen manuelle Prozessschritte
- Information wird in Form von Reports und OLAP-Analysen angeboten
- Realtime-Alarming in den operativen System, isoliert von den Nachbarsystemen

### Righttime Analytics

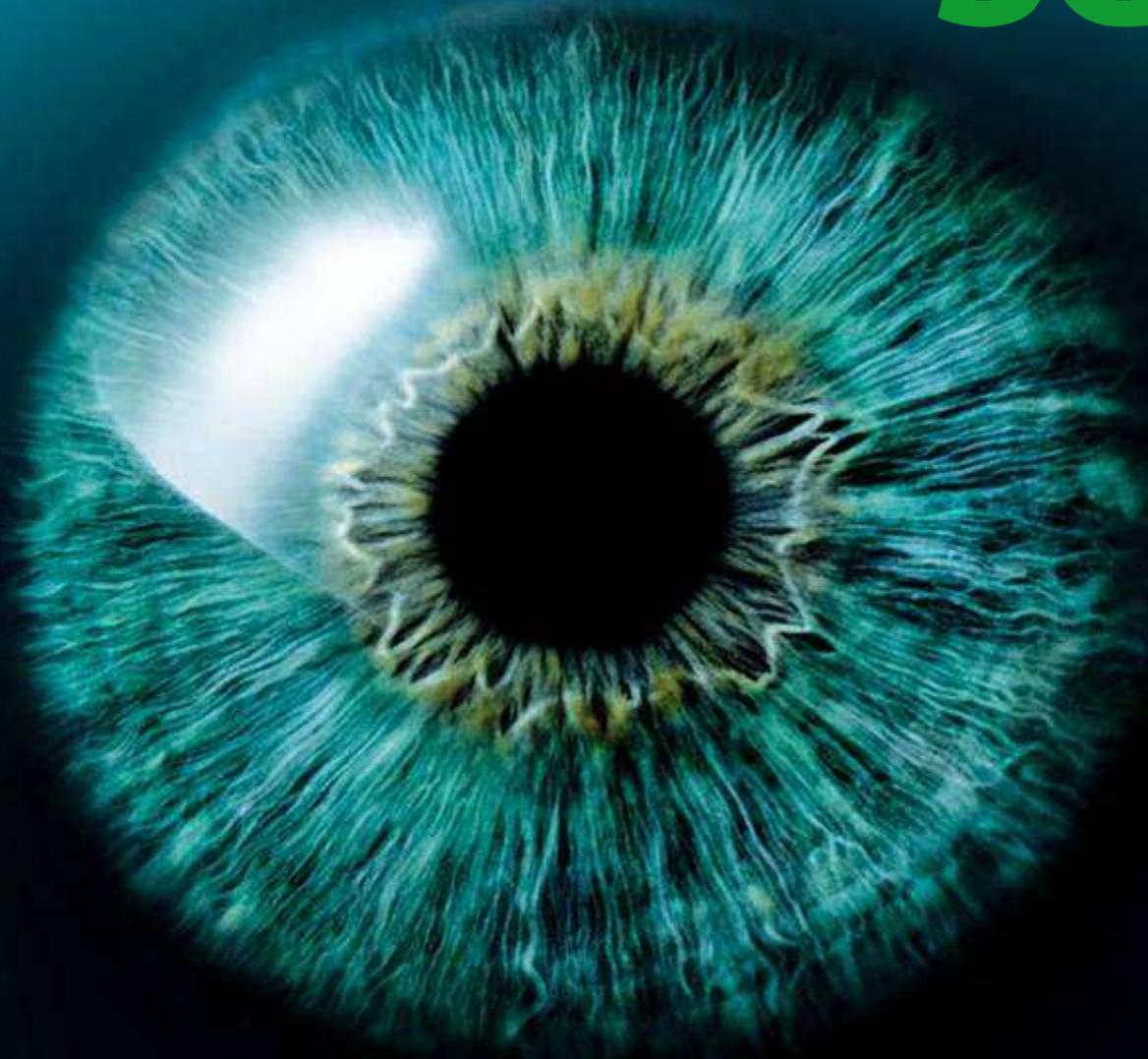
- Ist zudem Grundlage für operative und taktische Entscheidungen
- Closed loop möglich: Die Daten werden hochaktuell und automatisiert ins DWH eingeladen. Trigger können daraus automatische Informationen für das operative System erzeugen
- Zusätzliches Anbieten der Information für operative Systeme, z. B. in Form von Services
- Realtime-Alarming möglich in der 360°-Sicht im DWH über alle Systeme hinweg

# Zusammenfassung

- Zweck eines DWH
  - Zusammenführung und Homogenisierung von Daten aus div. Quellsystemen
  - für Berichtszwecke optimierte Art der Speicherung der Daten
  - Zielgruppe: Manager / Entscheidungsträger
- Technische Grundlagen
  - Kernkomponenten eines DWH: ETL, Datenbank, Analyse und Präsentation
  - Logische Struktur der Datenbank: multidimensionale Würfel (Fakten und Dimensionen)
  - Sternschema bildet multidimensionale Würfel auf relationale DB ab
- Projekterfahrungen
  - Essenziell: definierte Projektziele, früh sichtbarer Nutzen, Projektponsor, Vorgehensweise („think big – start small“)
  - Weniger wichtig ist die Toolauswahl, zu vermeiden ein „Nachprogrammieren“ eines vorhandenen Berichtswesens
- Data Mining dient der maschinellen Erkennung bisher unbekannter Zusammenhänge
- Righttime Analytics stellt operative Daten zeitnah im DWH bereit.

**Vielen Dank für die Aufmerksamkeit!**

**sd&m**



**André Speulmanns**

**0211 56623-208**

**[andre.speulmanns@sdm.de](mailto:andre.speulmanns@sdm.de)**



## Agenda

---

- Vorstellung
- Begriffsklärung: DWH und BI
- Technische Grundlagen
- sd&m-Erfahrungen im Umgang mit BI-Projekten
- Aktuelle Themen und Ausblick
- **Anhang: Literatur und Webadressen**

# Literatur: Bücher

---

- Berson/Smith: Data Warehousing, Data Mining & OLAP. McGraw-Hill 1997
- Kimball: The Data Warehouse Toolkit. Wiley 1996
- Kimball, Reeves, Ross, Thornthwaite: The Data Warehouse Lifecycle Toolkit. Wiley 1998
- Inmon: Building the Data Warehouse. Wiley 1996
- Inmon, Welch, Glassey: Managing the Data Warehouse. Wiley 1996
- Venerable, Adamson: Data Warehouse Design Solutions. Wiley 1998
- Martin (Hrsg.) : Data Warehousing - Data Mining - OLAP. Int. Thomson Publ. 1998
- Thomsen: OLAP Solutions - Building Multidimensional Information Systems. Wiley 1997
- Singh: Data Warehousing - Concepts, Technology, and Applications. Prentice-Hall 1997
- Devlin: Data Warehouse - From Architecture to Implementation. Addison-Wesley 1997
- Simon: 90 Days to the Data Mart. Wiley 1998
- Simon: Strategic Database Technology - Management for the Year 2000. Morgan Kaufmann 1995
- O'Neil: Oracle Data Warehousing Unleashed. Sams 1998
- Stonebraker, Hellerstein (Ed.) : Readings in Database Systems. Morgan Kaufmann 1998

# Web-Adressen (1)

---

<http://www.datawarehousing.com/>

Data Warehousing on the WWW

<http://www.rkimball.com>

Ralph Kimballs Homepage

<http://www.olapcouncil.org>

OLAP Council

<http://www.informationweek.de/>

Die Informationweek besitzt einen sehr interessanten Data Warehousing-Bereich

<http://www.dwinfocenter.org/>

The Data Warehousing Information Center

<http://www.dw-institute.com/>

Data Warehouse Institute: Hier gibt es eine gute Sammlung von technischen Whitepapers von verschiedenen Tool-Anbietern

<http://www.olapreport.com/>

The OLAP Report: Interessante und hilfreiche Marktanalysen.

<http://www.dmreview.com/>

Echte Fundgrube für Whitpapers, Produktbesprechungen und Auflistung von Anbietern für Spezialthemen.

<http://www.olapinfo.de/index.html>

Whitepapers und ausführliche Produktbeschreibung vieler OLAP-Werkzeuge.

## Web-Adressen (2)

---

- Arbor Software (Arbor Essbase) <http://www.arborsoft.com>
- Brio Technology (Brio Query) <http://www.brio.com>
- Business Objects <http://www.businessobjects.com>
- Cognos (Impromptu, Powerplay) <http://www.cognos.com>
- IBM <http://www.software.ibm.com/data/>
- Informix (MetaCube) <http://www.informix.com>
- Microstrategy (DSS Agent / Server) <http://www.strategy.com>
- NCR (Teradata) <http://www3.ncr.com/sdw/>
- Oracle (Oracle Express) <http://www.oracle.com/products/olap/>
- SAS Institute <http://www.sas.com/feature/4qdm/>
- Seagate Software (Seagate Holos) <http://seagatesoftware.com/holos>