



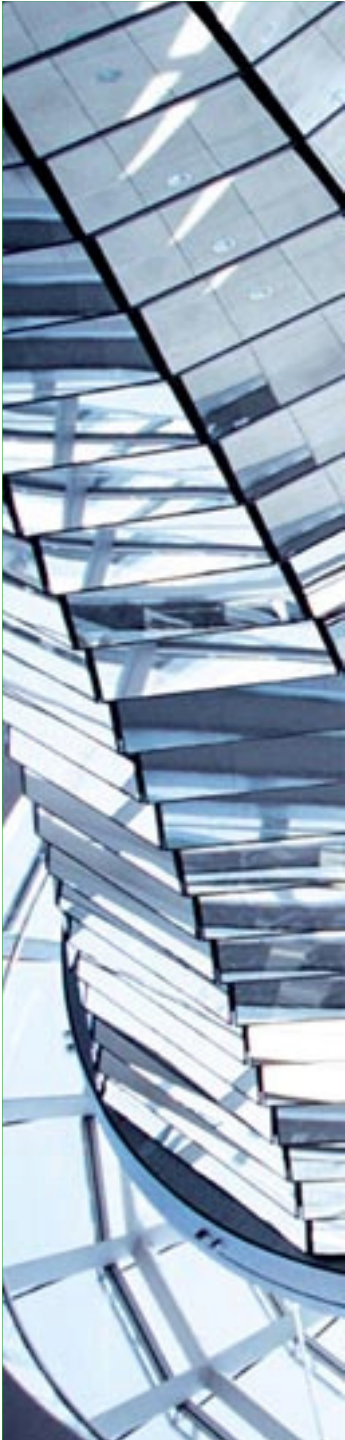
# 8. Business Intelligence Reference Architectures and Patterns

Winter Semester 2008 / 2009  
Prof. Dr. Bernhard Humm  
Darmstadt University of Applied Sciences  
Department of Computer Science

# The lecture in the context of the entire course

---

1. Introduction
2. A reference architecture for business information systems
3. Application kernel
4. Persistence and transaction
5. Authorization
6. Client architecture
7. Exception handling
8. Business Intelligence
9. Systems integration
10. Service-oriented architecture
11. Selected design patterns
12. Design for testability



# Agenda

---

## → Definitions

Reference Architecture

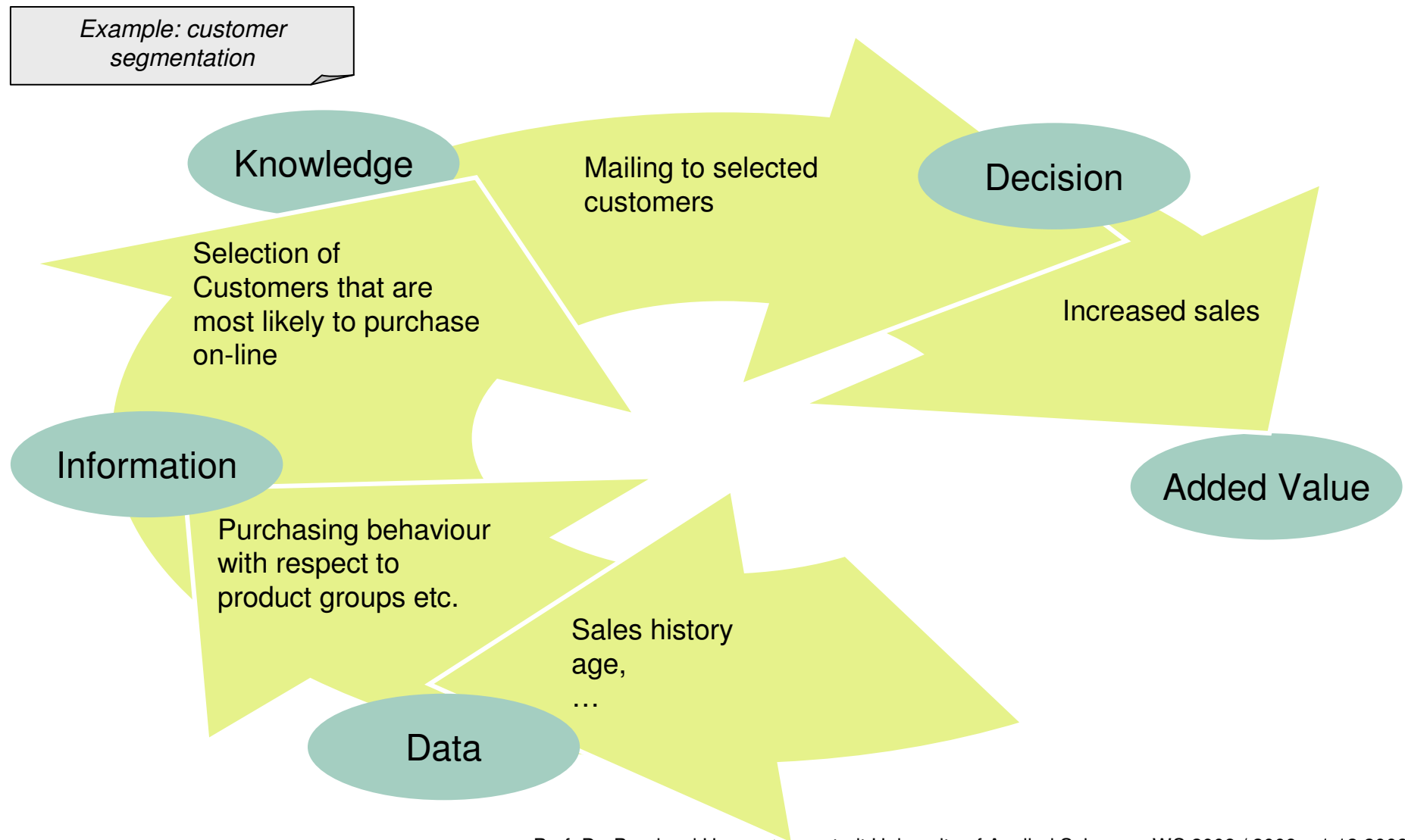
ETL

Aggregation

Products

Literature

# Business Intelligence (BI) is the process of transforming data into information and, furthermore, into knowledge



## Business Intelligence: a buzzword amongst many...

---

***Business Intelligence*** subsumes applications and technologies like, e.g., *Data Warehousing (DW), Data Mining, Online Analytical Processing (OLAP), and Analytical Applications.*

*Other related buzzwords / synonyms:*

*Analytical Customer Relationship Management (aCRM), Corporate Performance Management (CPM), Extraction – Transformation - Load (ETL), Right Time Analytics, Management Information System (MIS), Decision Support System (DSS)*

# Example of a BI application

**Country**

**Year**

**Top Selling Products**

Products	H1 FY 2004	H2 FY 2004	FY 2004
Mountain-200 Black, 46	\$121,955.73	\$121,644.83	\$243,600.56
Mountain-200 Black, 38	\$102,845.64	\$137,930.41	\$240,776.04
Mountain-200 Silver, 38	\$88,414.08	\$137,614.44	\$226,028.52
Mountain-200 Black, 42	\$102,675.48	\$121,692.38	\$224,367.86
Mountain-200 Silver, 46	\$94,009.80	\$123,272.61	\$217,282.42
Touring-1000 Blue, 54	\$12,324.69	\$36,113.89	\$48,438.58
Hitch Rack - 4-Bike	\$24,909.66	\$18,768.94	\$43,678.60
Touring-1000 Yellow, 54	\$5,996.41	\$35,057.27	\$41,053.69
Mountain-400-W Silver, 38	\$14,304.42	\$19,005.19	\$33,309.62
Touring-1000 Blue, 50	\$7,533.66	\$24,956.45	\$32,490.11
Mountain-400-W Silver, 40	\$15,098.35	\$15,771.47	\$30,869.82
Women's Mountain Shorts, S	\$14,582.82	\$14,622.16	\$29,204.98
Women's Mountain Shorts, L	\$14,673.36	\$13,544.42	\$28,217.78
Road-350-W Yellow, 44	\$2,527.00	\$24,608.57	\$27,135.56
Classic Vest, S	\$14,601.01	\$12,307.49	\$26,908.50
Mountain-400-W Silver, 42	\$10,976.21	\$13,635.02	\$24,611.23
Road-750 Black, 44	\$10,050.08	\$14,232.42	\$24,282.50
Road-350-W Yellow, 42	\$4,684.54	\$18,829.97	\$23,514.51
Mountain-400-W Silver, 46	\$8,945.79	\$13,928.43	\$22,874.22
Sport-100 Helmet, Blue	\$10,697.04	\$12,123.04	\$22,820.07
Touring-1000 Blue, 46	-\$5,295.97	\$27,997.57	\$22,701.60

**Top 5 Cities by Profit**

City	Sales	Profit
Seattle	\$1,494,860.53	\$129,944.49
Phoenix	\$685,695.21	\$67,498.32
Loveland	\$816,333.19	\$62,737.64
Tooele	\$877,107.19	\$57,834.06
Saint Louis	\$648,952.98	\$57,834.06

**Facts and measures (possibly aggregated)**

**Graphic representation**

Source: MSDN

© 2008 / 2009. 1.12.2008

# Facts and Measures

## ■ Measure

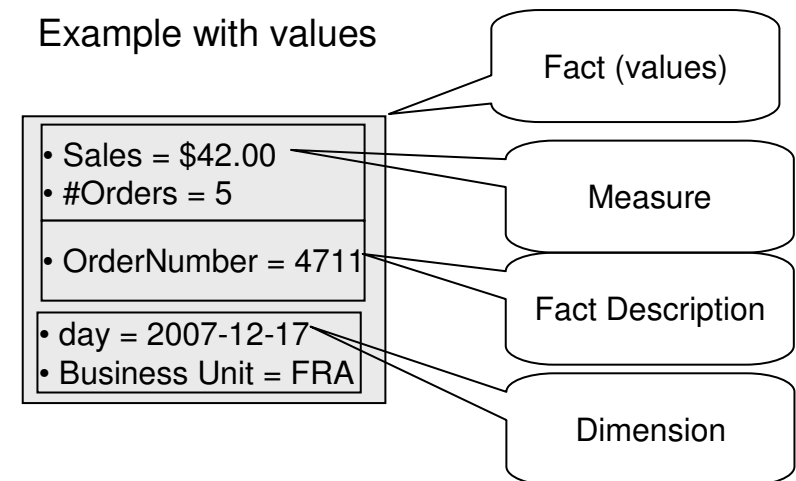
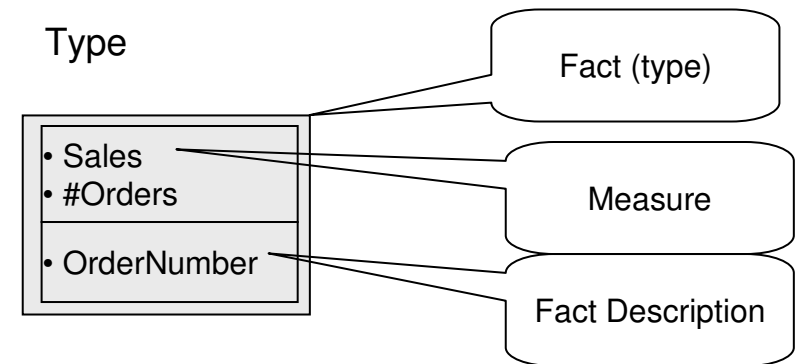
- The smallest unit of information in a DW
- Always numerical
- Can be aggregated (sum, average, etc.)
- Distinguish between measure types (e.g., sales) and measure values (e.g., \$42.00)

## ■ Fact Description

- Provide additional information concerning measures
- Will not be aggregated; need not be numerical

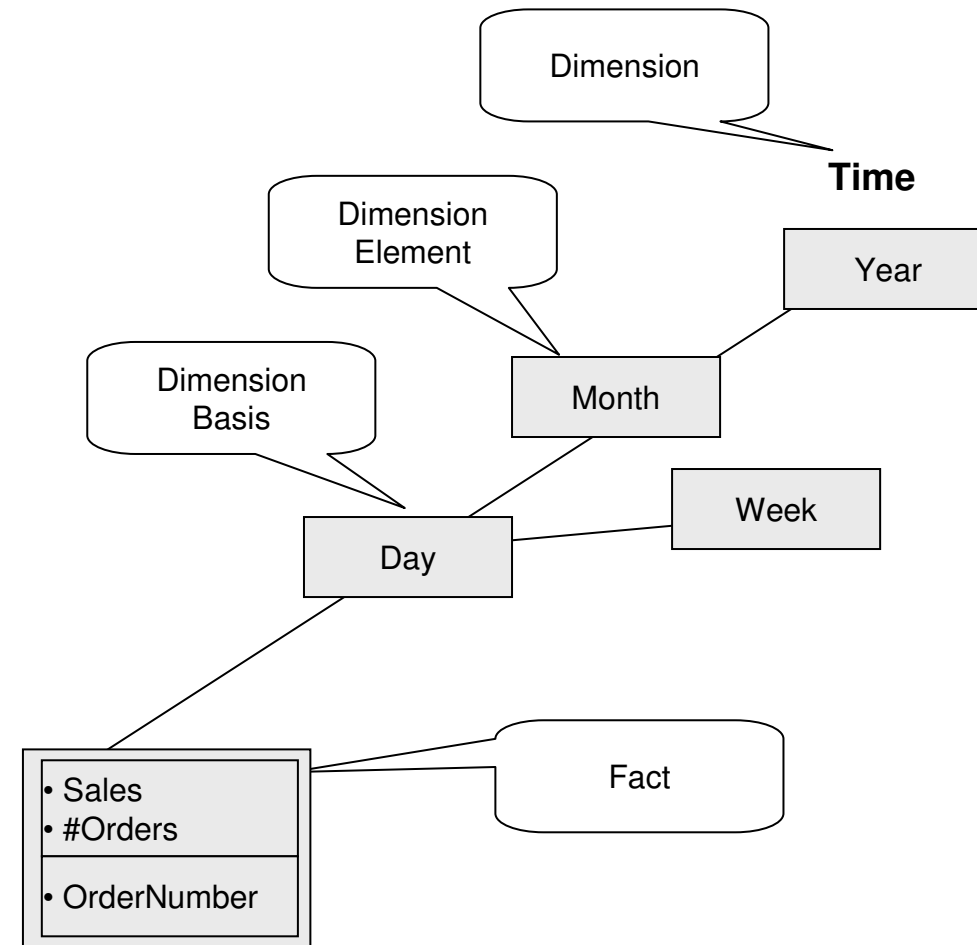
## ■ Fact

- An entity consisting of measures and fact descriptions as attributes
- Associated to dimensions



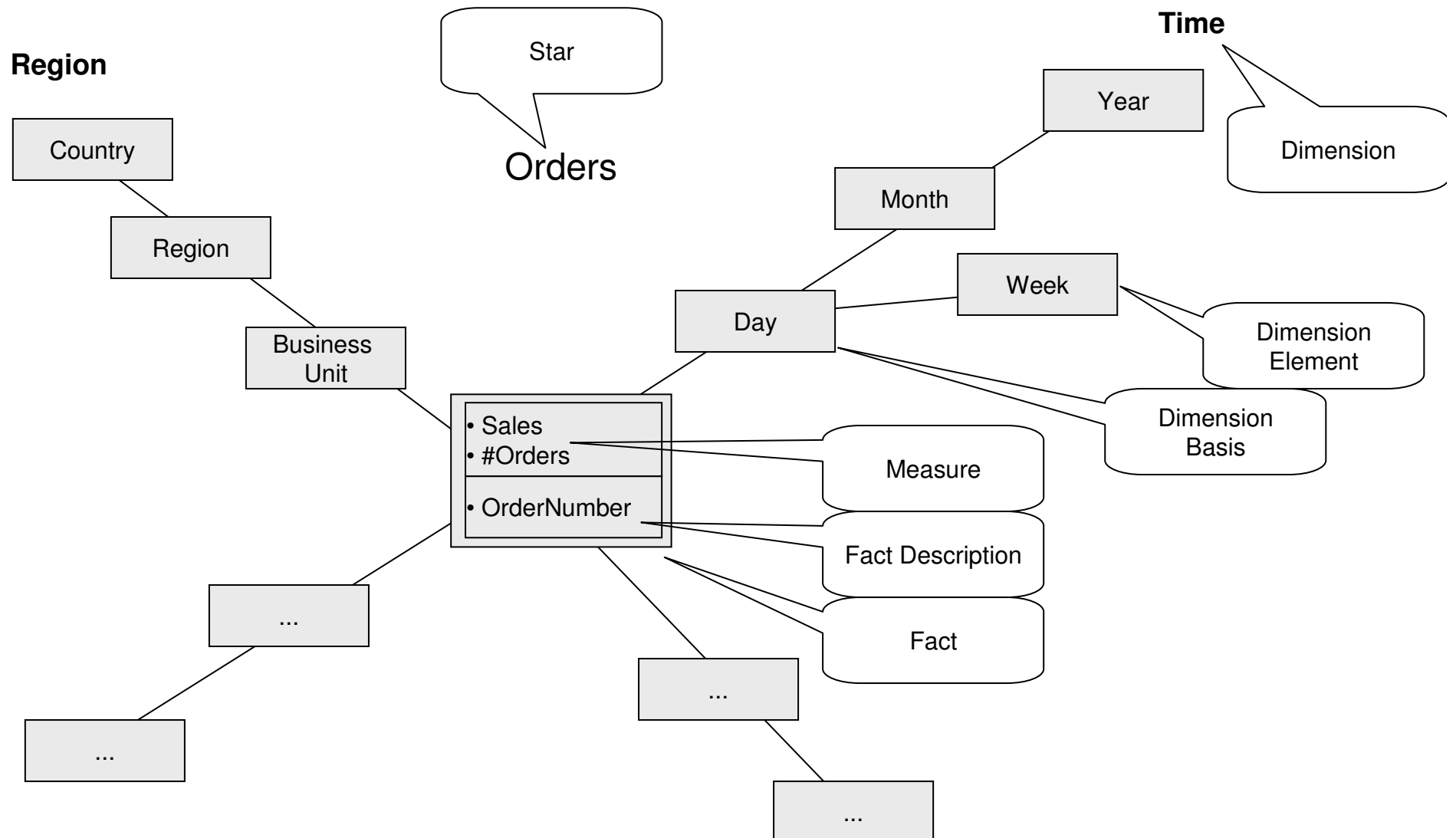
# Dimensions

- Dimension
  - Filter- and aggregation criterion for measures
  - Span a multi-dimensional space
  - Provide a coordinate system for navigating through measures
- Dimension element
  - Dimension can be hierarchically structured into several dimension elements (dimension hierarchy)
  - 1..n relationship between dimension elements
  - Form a list or – rarely – a tree resp. a directed acyclic graph (DAG)
  - Distinguish between dimension element types (e.g., day) and values (e.g., 2007-12-17)
- Dimension basis
  - Is a particular dimension element
  - Most concrete dimension element (innermost)

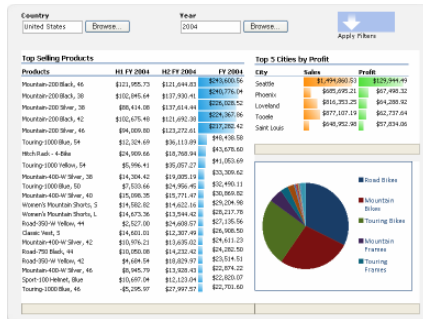


# Star (Cube) = Facts + Dimensions

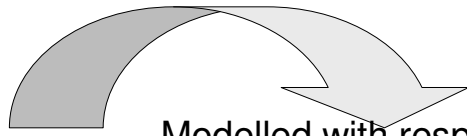
## Galaxy = Stars with common dimensions



# Modelling Reports and Stars (Cubes)

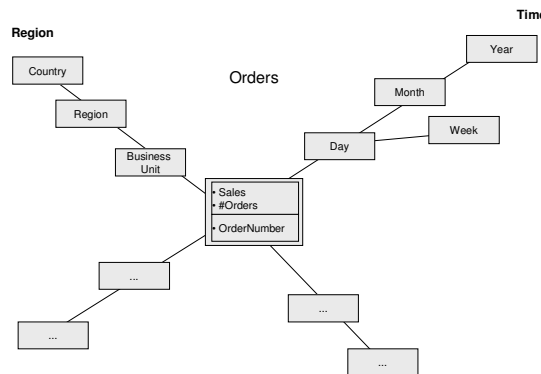


Report



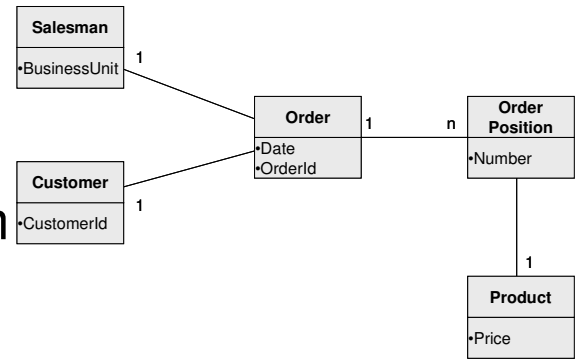
Modelled with respect to

Star (Cube) in DW  
(multi-dimensional)



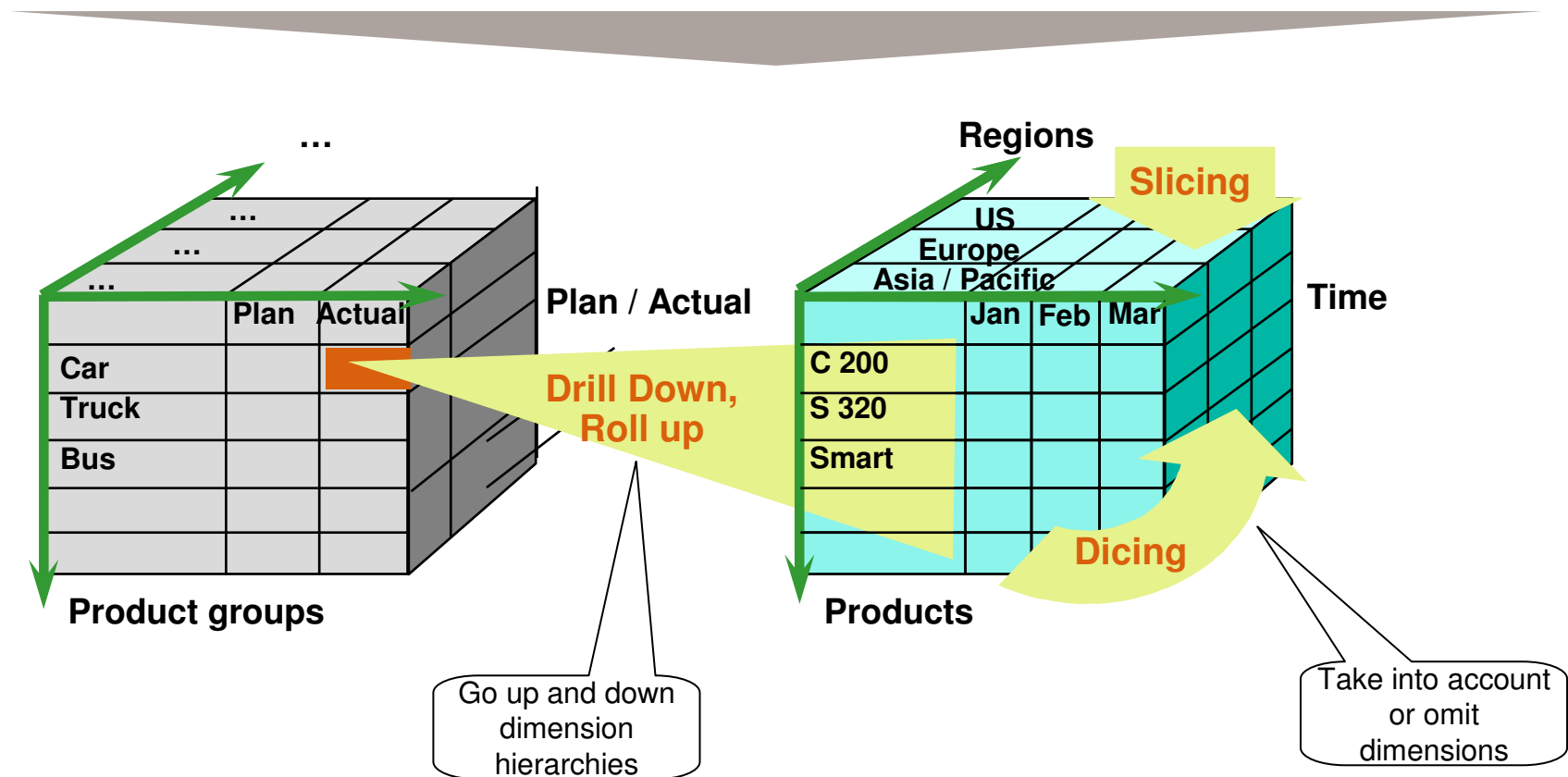
Modelled with respect to

Operational System  
(relational)



# Navigating in a cube: slicing & dicing, drill down & roll up

The star is represented as a multi-dimensional cube





# Agenda

---

Definitions

→ **Reference Architecture**

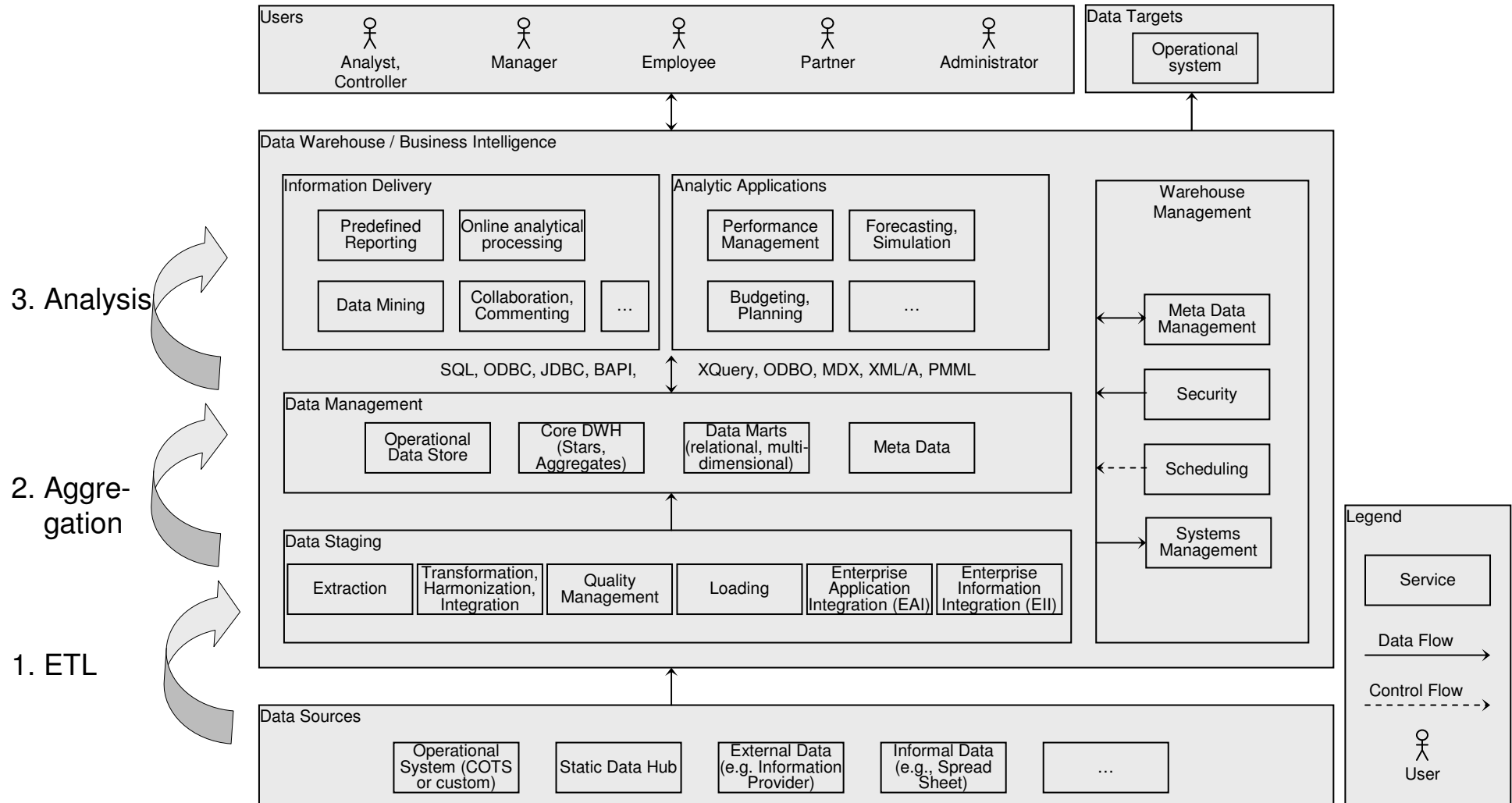
ETL

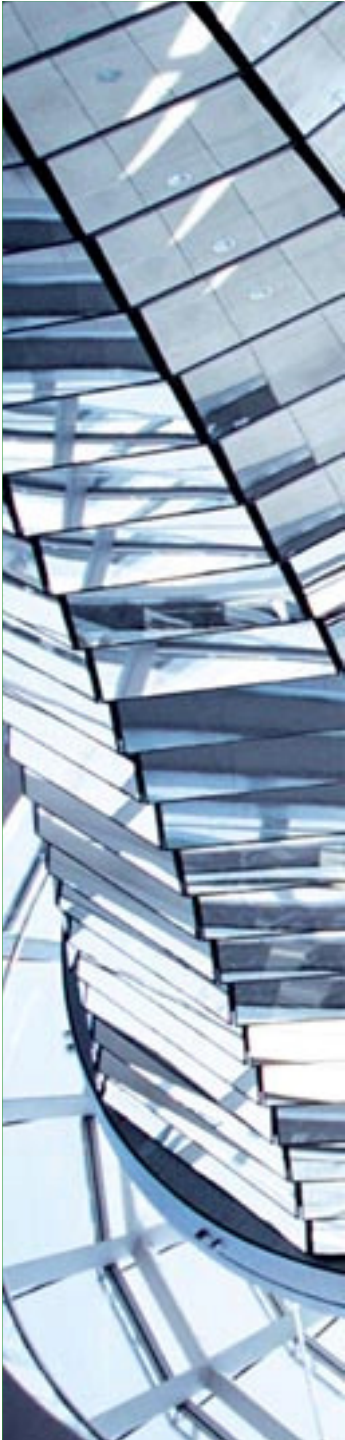
Aggregation

Products

Literature

# The reference architecture for Business Intelligence / Data Warehousing





# Agenda

---

Definitions

Reference Architecture

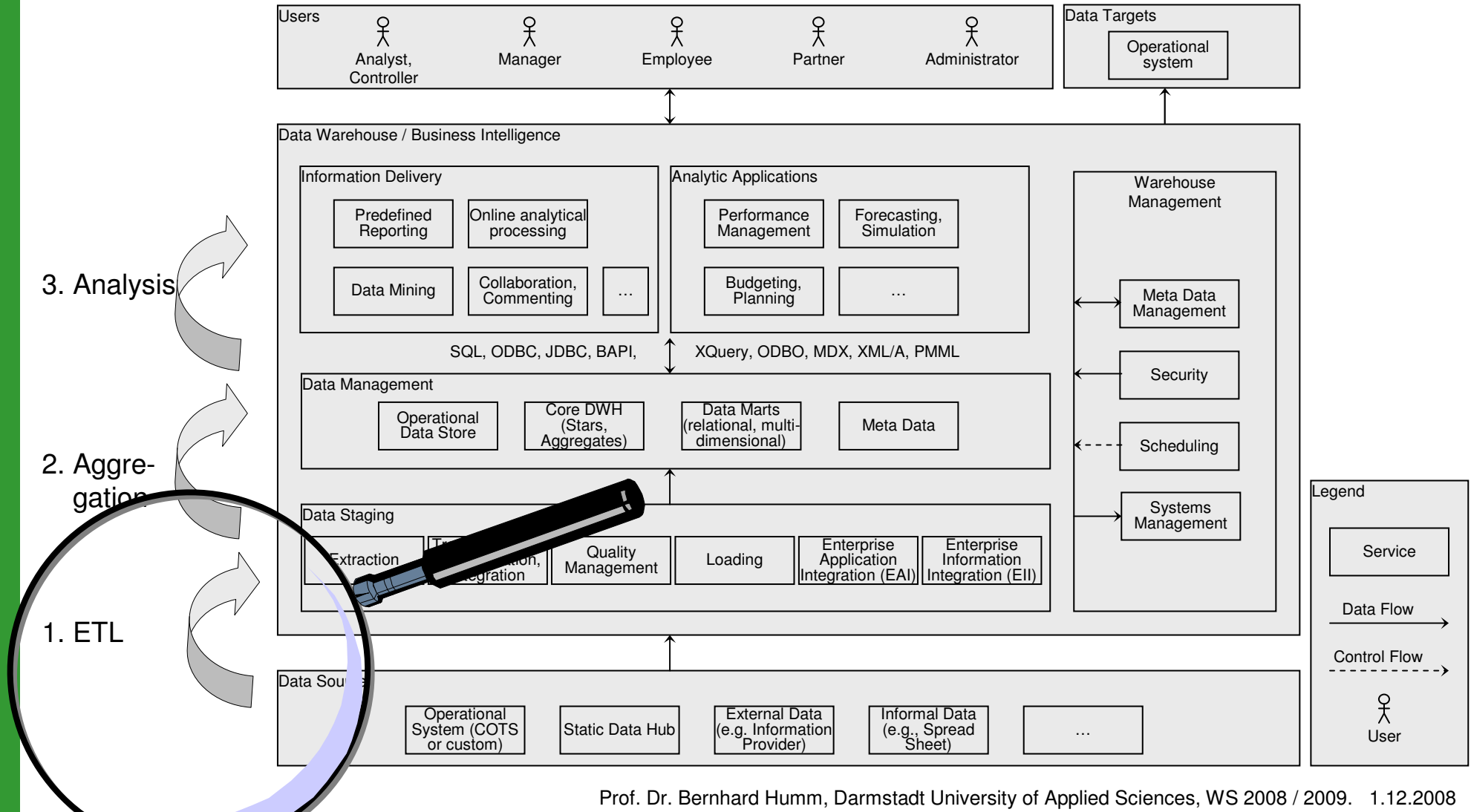
→ ETL

Aggregation

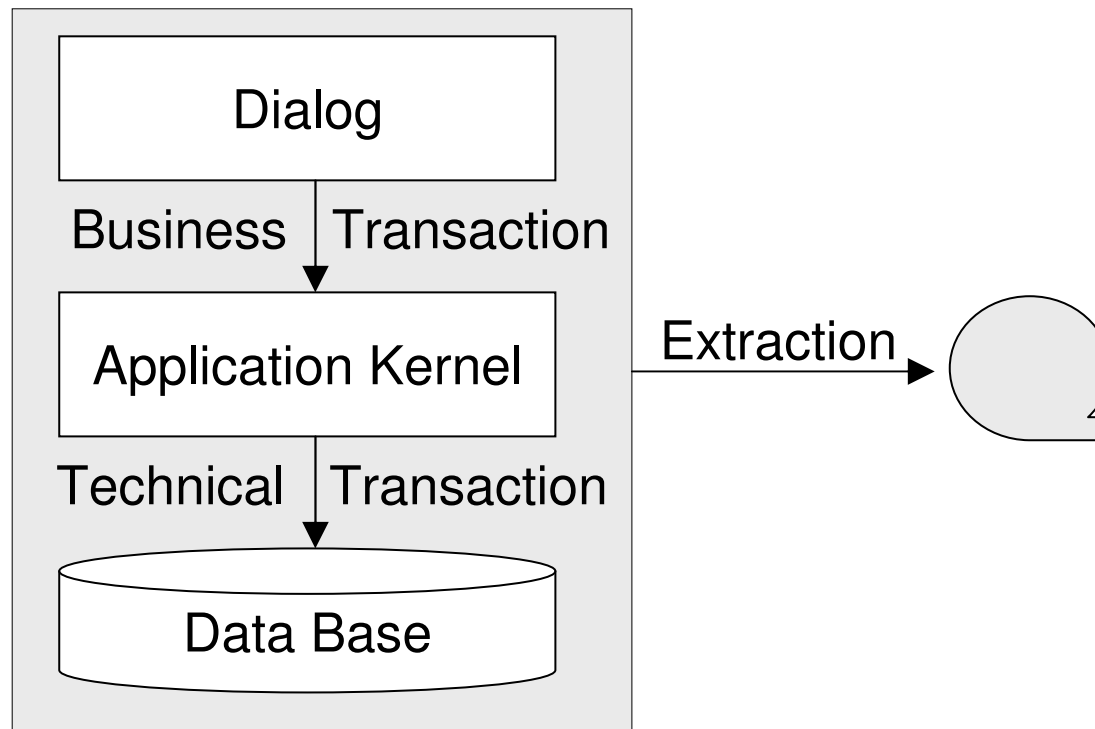
Products

Literature

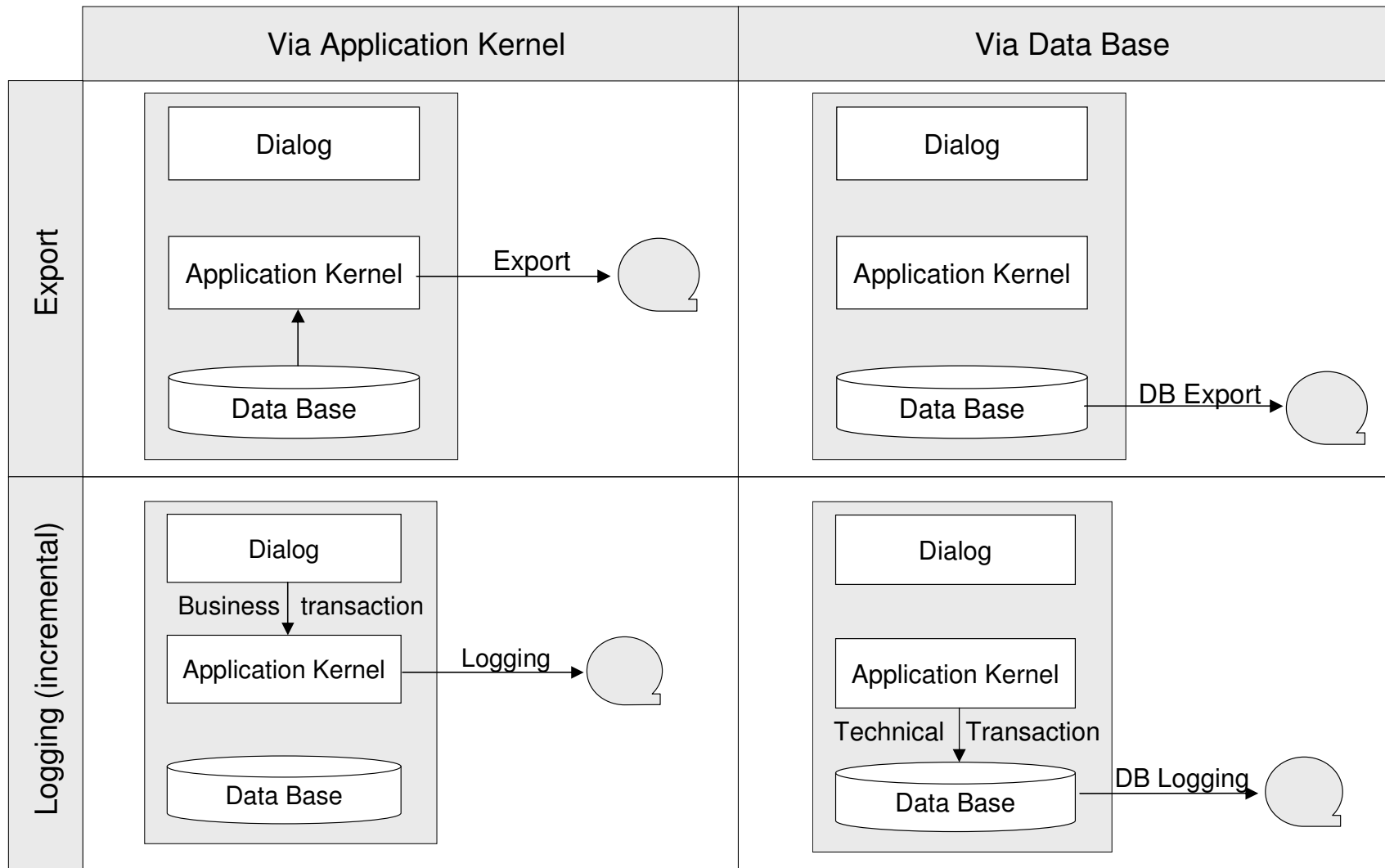
# Extraction / Transformation / Loading in the context of the reference architecture



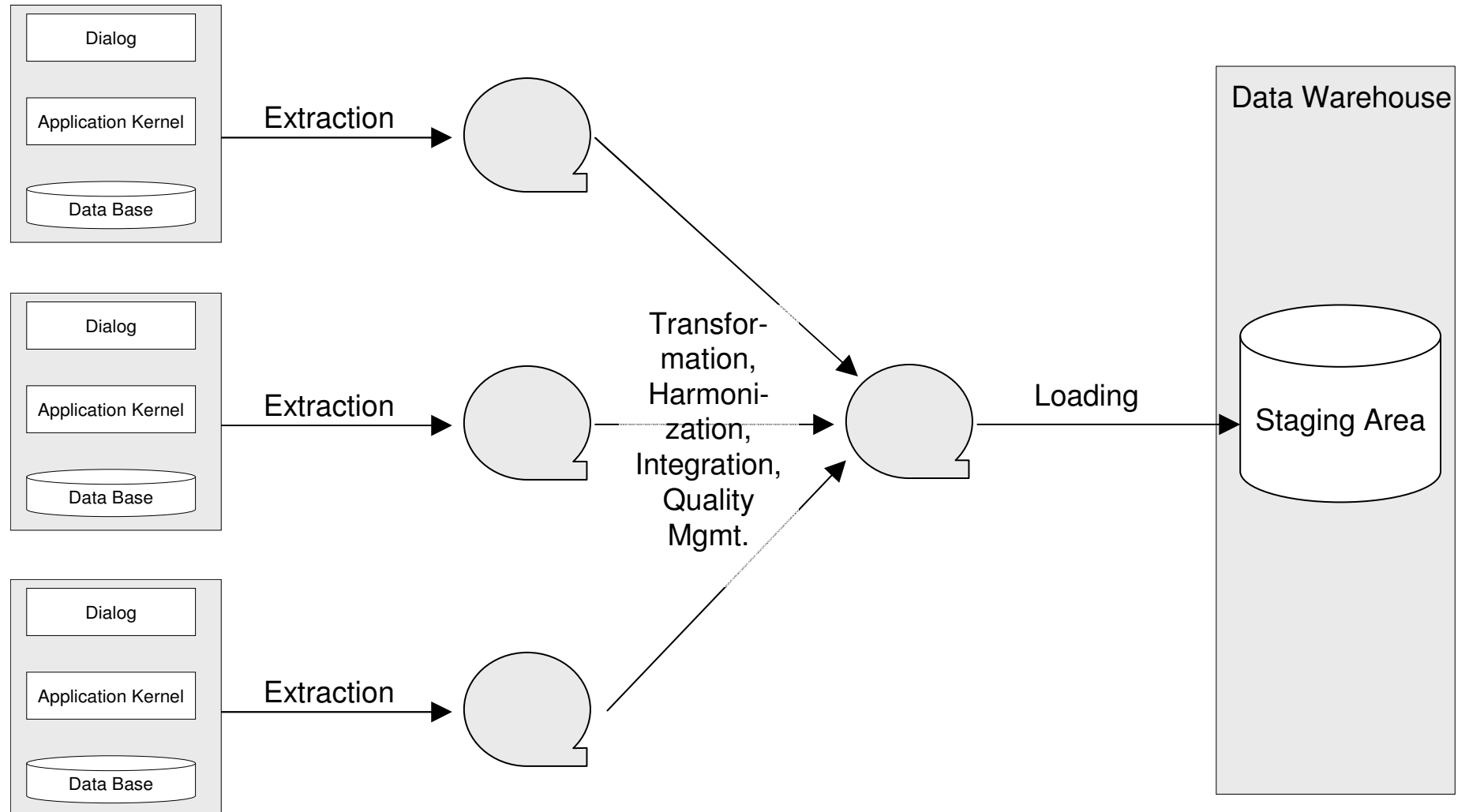
# Extraction: How to extract data from operational systems?



# 4 ways of extracting data from operational systems



# Transformation and Loading





# Agenda

---

Definitions

Reference Architecture

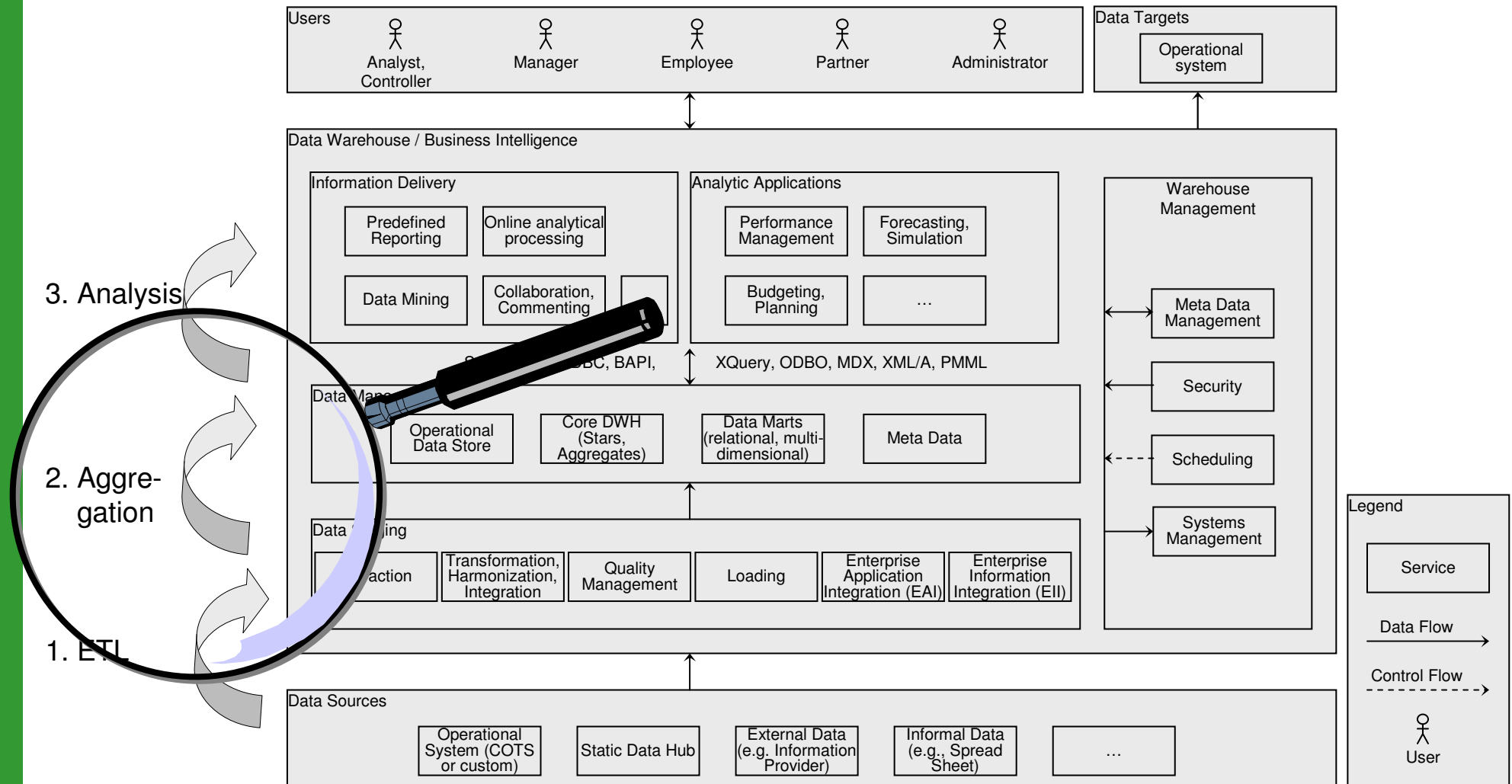
ETL

**→ Aggregation**

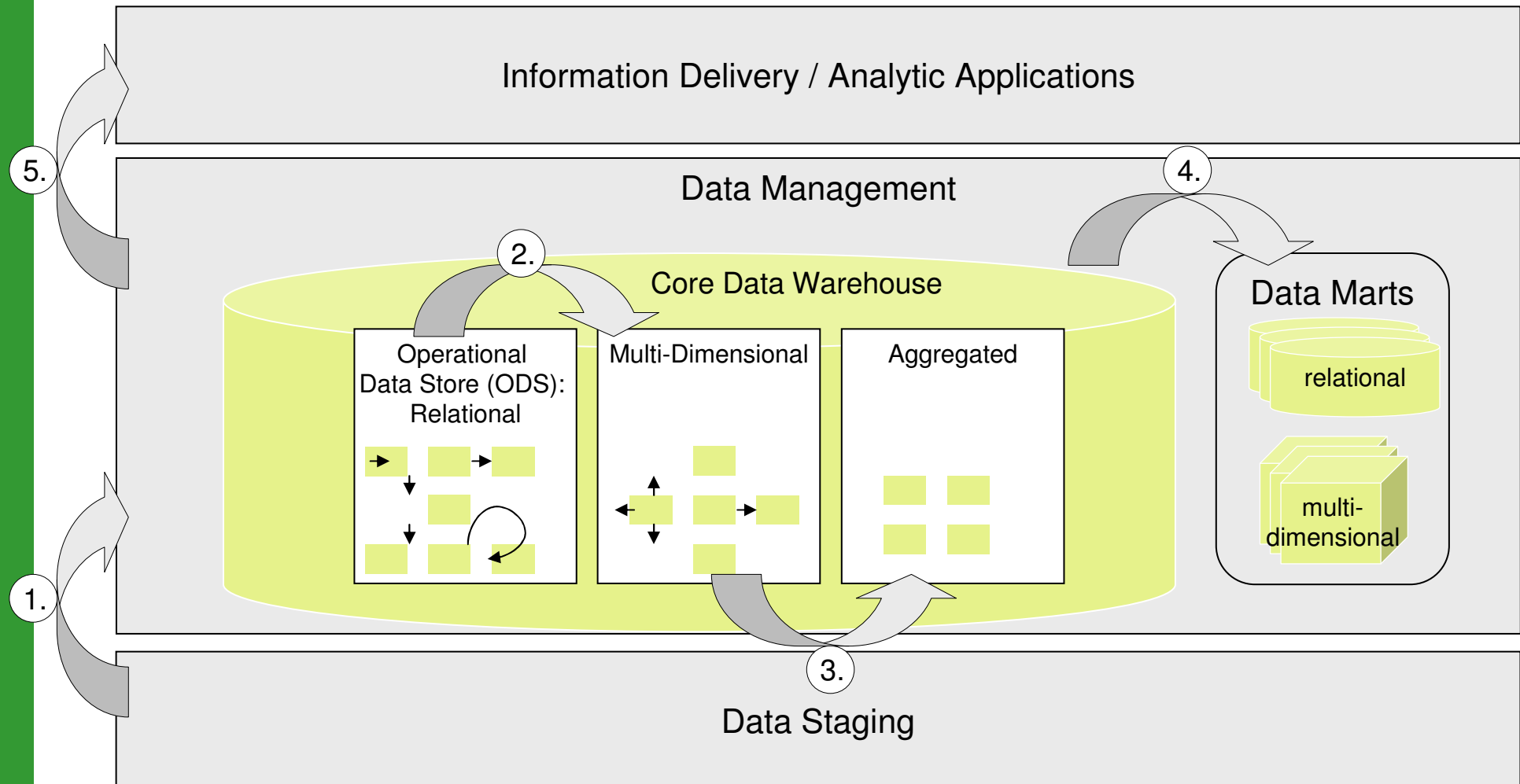
Products

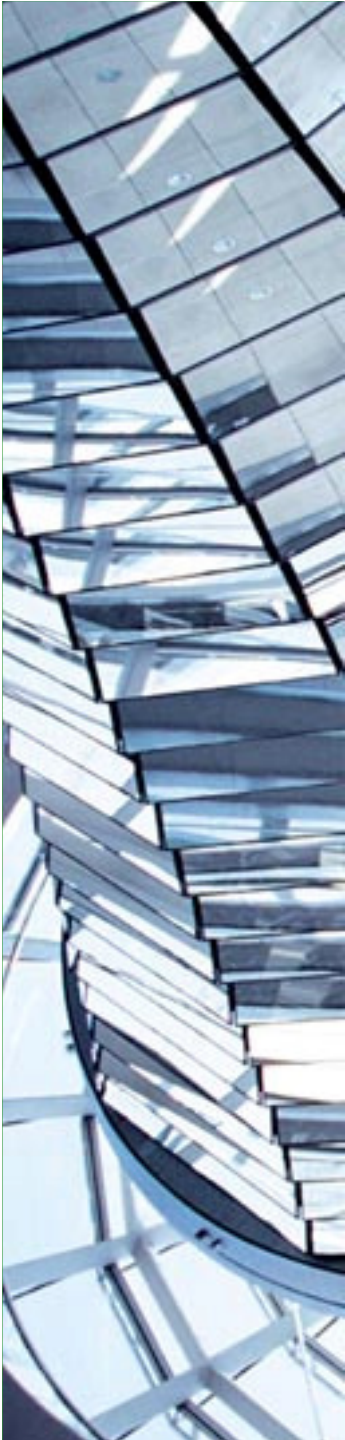
Literature

# Aggregation in the context of the reference architecture



# Aggregation





# Agenda

---

Definitions

Reference Architecture

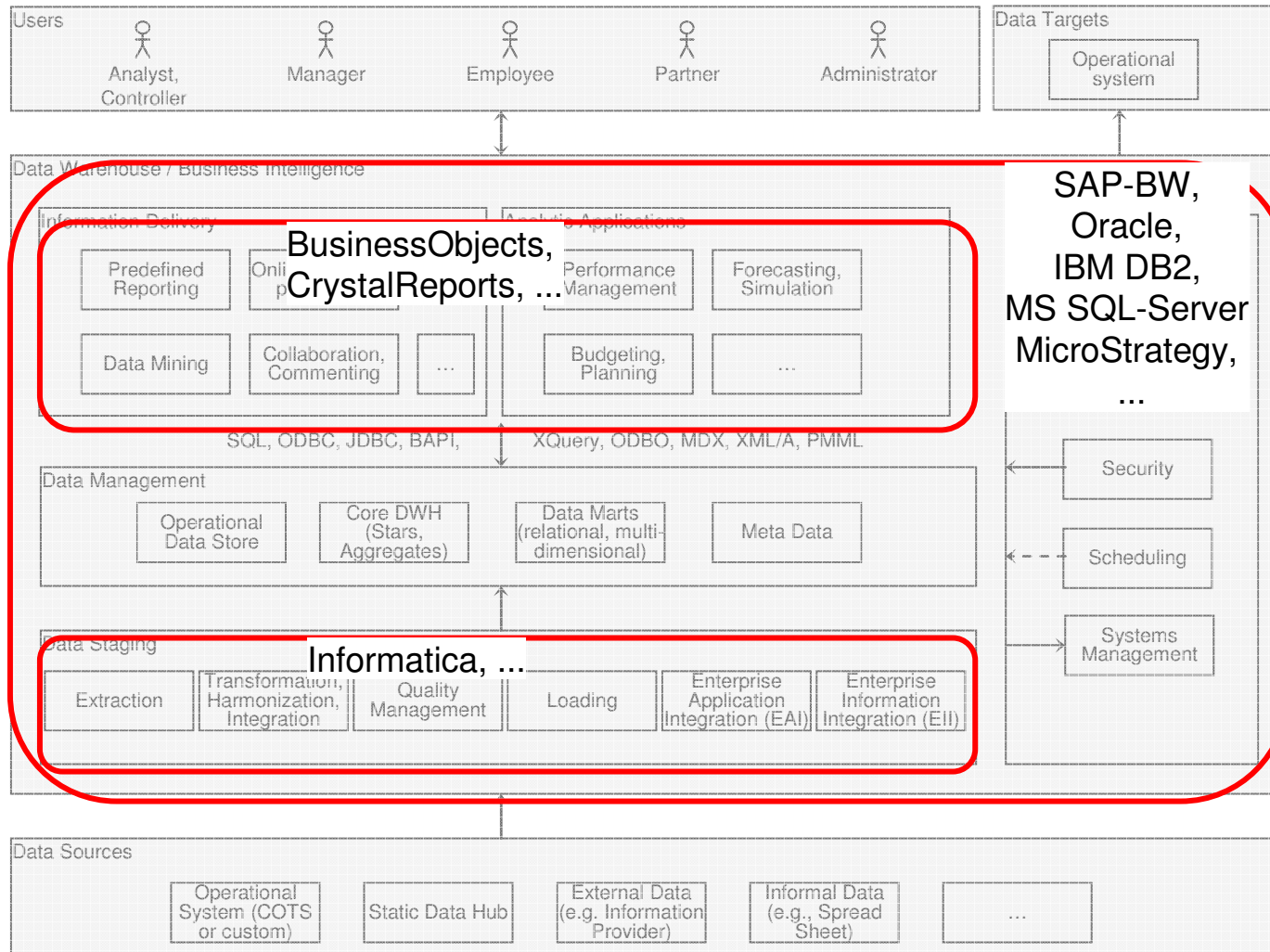
ETL

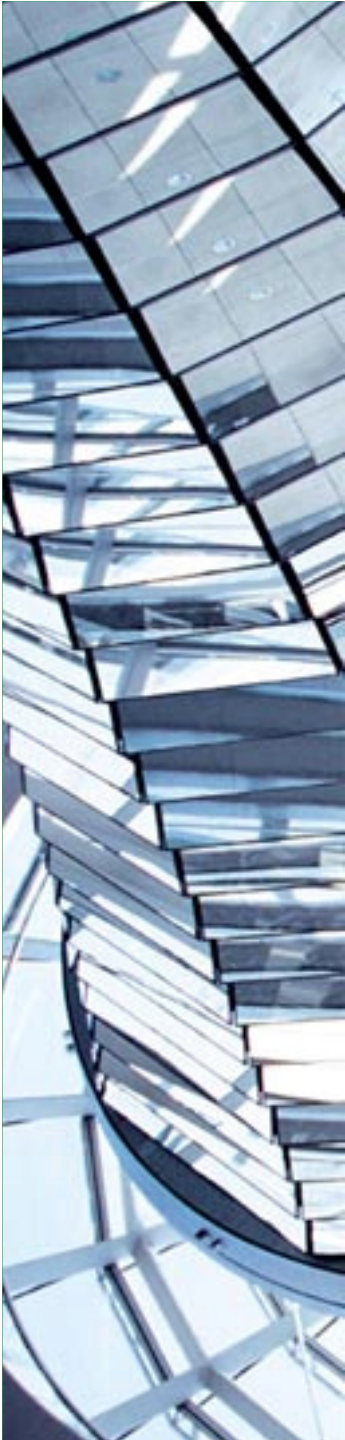
Aggregation

→ **Products**

Literature

# A product map assigns products to clusters of services in the reference architecture





# Agenda

---

Definitions

Reference Architecture

ETL

Aggregation

Products

→ Literature

# Literature

- Bernhard Humm, Frank Wietek: *Architektur von Data Warehouses und Business Intelligence Systemen*. Informatik Spektrum 3/05, S. 3-14, Springer Verlag. 2005 (download from my home page)

## Architektur von Data Warehouses und Business Intelligence Systemen

Bernhard Humm · Frank Wietek

**Business Intelligence (BI) ist der Prozess der Umwandlung von Daten in Informationen und weiter in Wissen. Entscheidungen und Prognosen stützen sich auf dieses Wissen und schaffen dadurch Mehrwert für ein Unternehmen. Ein Data Warehouse (DW) bildet in vielen Fällen die technische Basis zur Implementierung einer BI-Lösung.**

Architektur über die technische Architektur und Systemarchitektur bis zu Produkten. Den Abschluss bilden aktuelle Markttrends.

### Historie

*Entscheidungsunterstützende (dispositive/analytische) Systeme* haben eine lange Historie seit den 60er-Jahren und wurden im Verlauf der Jahrzehnte bei recht ähnlicher Funktionalität lediglich unterschiedlich betitelt: *Management Information Systems (MIS)*, *Decision Support Systems (DSS)*, *Executive Information Systems (EIS)*, *Data Warehouses (DW)* und schließlich *Business Intelligence (BI)-Lösungen*. Abbildung 1 gibt eine Übersicht über jeweilige Kernideen beziehungsweise Schlagworte.

Im Laufe dieser Zeit wurden das Maß der Integration von Daten und das Niveau der Entscheidungsunterstützung permanent verbessert. Während die Ansätze aus den 60er- und 70er-Jahren noch weitgehend – gemessen an ihren Ansprüchen – scheiterten, bauten die Werkzeuge späterer Genera-

Dieser Artikel stellt DW-Konzepte und -Architekturen dar, welche die Erprobung in mehrjähriger Praxis bestanden haben und somit als etabliert gelten. Er erläutert gängige Begriffe und beantwortet konkrete Fragestellungen des Systementwurfs. Der Bogen spannt sich von der Modellierung der fachlichen

tionen jeweils auf ihren Vorgängern auf, lernten aus deren Fehlern und wurden zunehmend erfolgreicher.

### Operative und dispositive Systeme: OLTP versus OLAP

Dienste, welche die Durchführung des operativen Geschäfts eines Unternehmens unterstützen, werden auch *OLTP (Online Transactional Processing)* genannt, dispositive Dienste auch *OLAP (Online Analytical Processing)*. OLAP und OLTP haben sehr unterschiedliche Charakteristiken, so dass sich eine Trennung in unterschiedliche Systeme anbietet. Tabelle 1 zeigt eine Gegenüberstellung.

### Data Warehouse

Ein *Data Warehouse* ist ein dispositives System. Wir zitieren eine klassische Definition von Inmon [12]: Ein Data Warehouse ist eine themenorientierte, zeitorientierte, integrierte und unveränderliche Datensammlung, deren Daten sich für Managemententscheidungen auswerten lassen. Insbesondere bedeuten:

- themenorientiert: alles über Kunden, Produkte etc.;
- zeitorientiert: periodische Ergänzung um aktuelle Daten, Verdichtung nach Zeitintervallen;
- integriert: Konsolidierung von Daten verschiedener operativer Systeme;
- unveränderlich: einmal gespeicherte Daten werden nicht mehr verändert.

DOI 10.1007/s00287-004-0450-5  
© Springer-Verlag 2005

Dr. B. Humm · Dr. F. Wietek  
sd8m Research  
Carl-Wery-Str. 42, 81739 München  
E-Mail: Bernhard.Humm@sdm.de  
E-Mail: Frank.Wietek@sdm.de