

Betriebliche Anwendungen

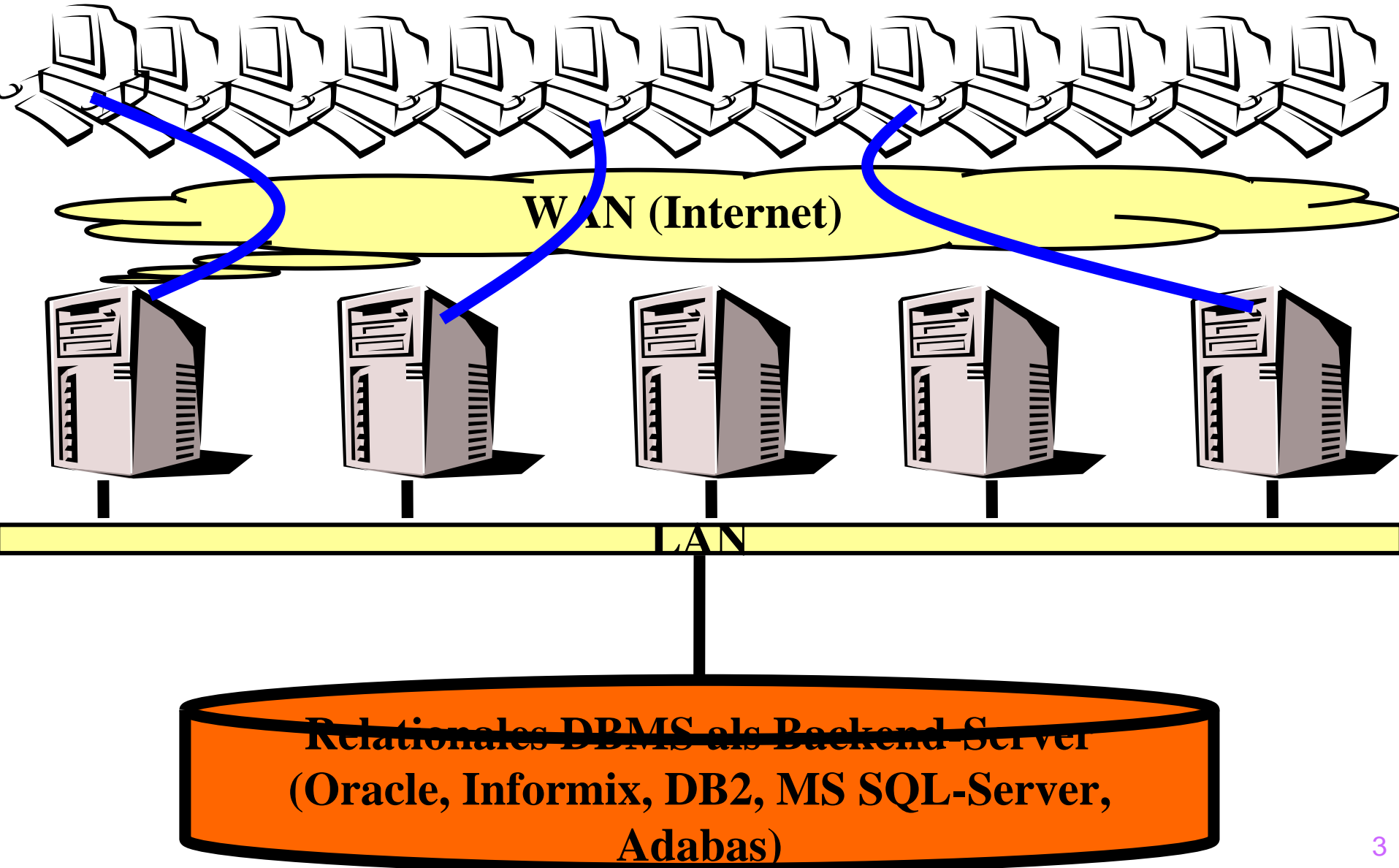
- OLTP
- Data Warehouse
- Data Mining



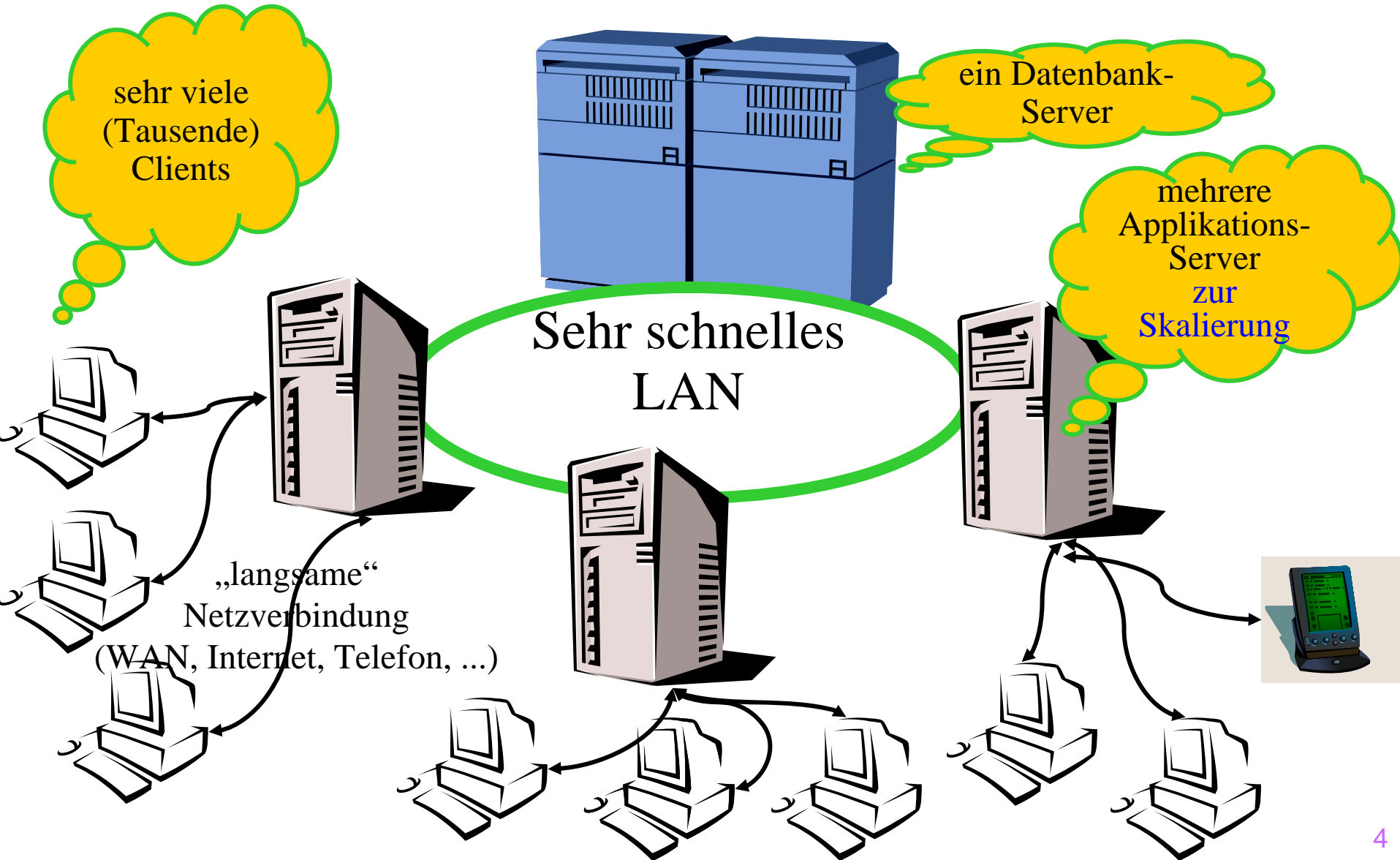
OLTP: Online Transaction Processing

- Beispiele
 - Flugbuchungssystem
 - Bestellungen in einem Handelsunternehmen
- Charakterisierung
 - Hoher Parallelitätsgrad
 - Viele (Tausende pro Sekunde) kurze Transaktionen
 - TAs bearbeiten nur ein kleines Datenvolumen
 - „mission-critical“ für das Unternehmen
 - Hohe Verfügbarkeit muss gewährleistet sein
- Normalisierte Relationen (möglichst wenig Update-Kosten)
- Nur wenige Indexe (wegen Fortschreibungskosten)

SAP R/3: Enterprise Resource Modelling (ERP-System)

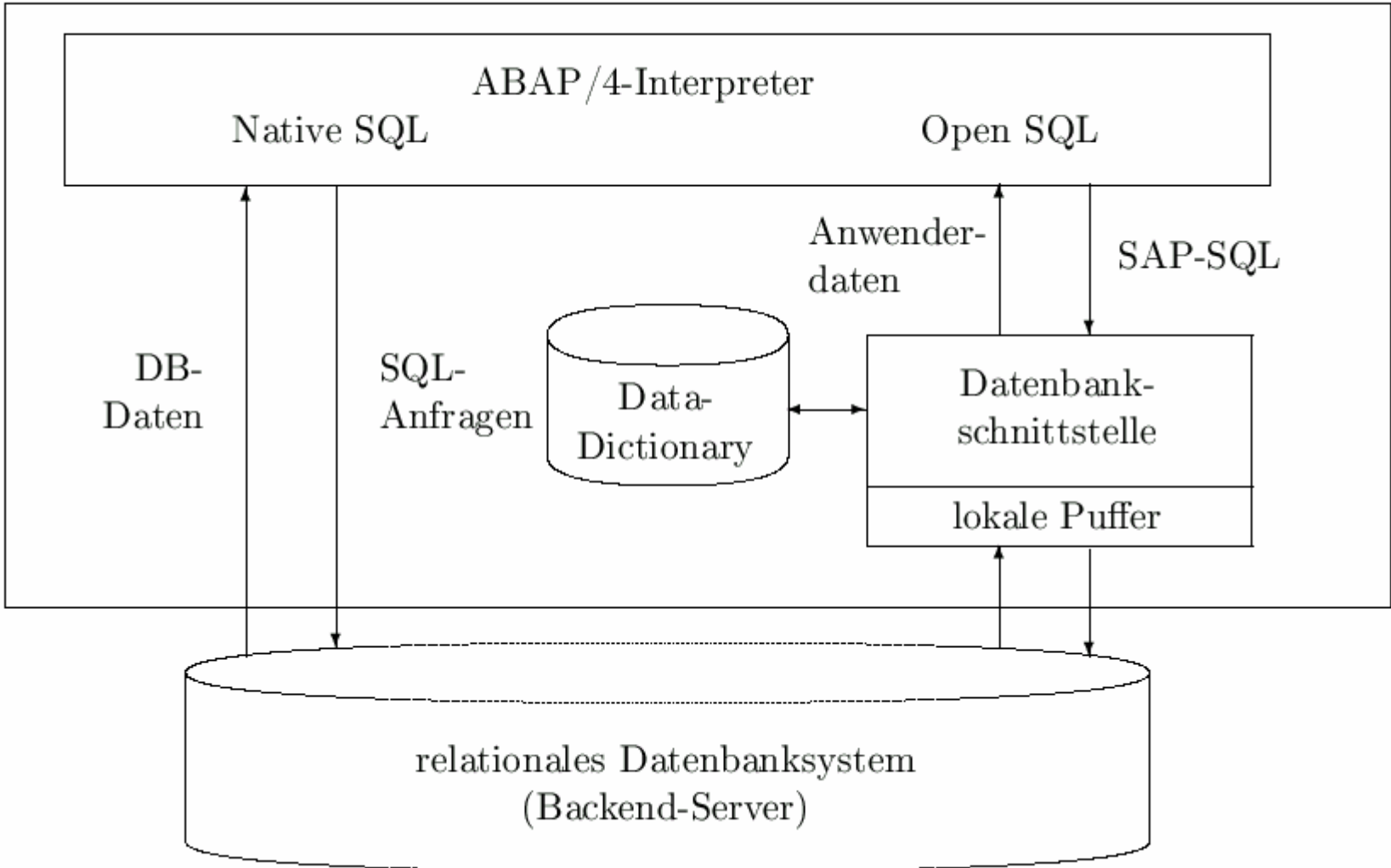


Dreistufige Client/Server-Architektur (3 Tier, SAP R/3)



Interne Architektur von SAP R/3

SAP R/3



Data Warehouse-Anwendungen: OLAP~Online Analytical Processing

- Wie hat sich die Auslastung der Transatlantikflüge über die letzten zwei Jahre entwickelt?

oder

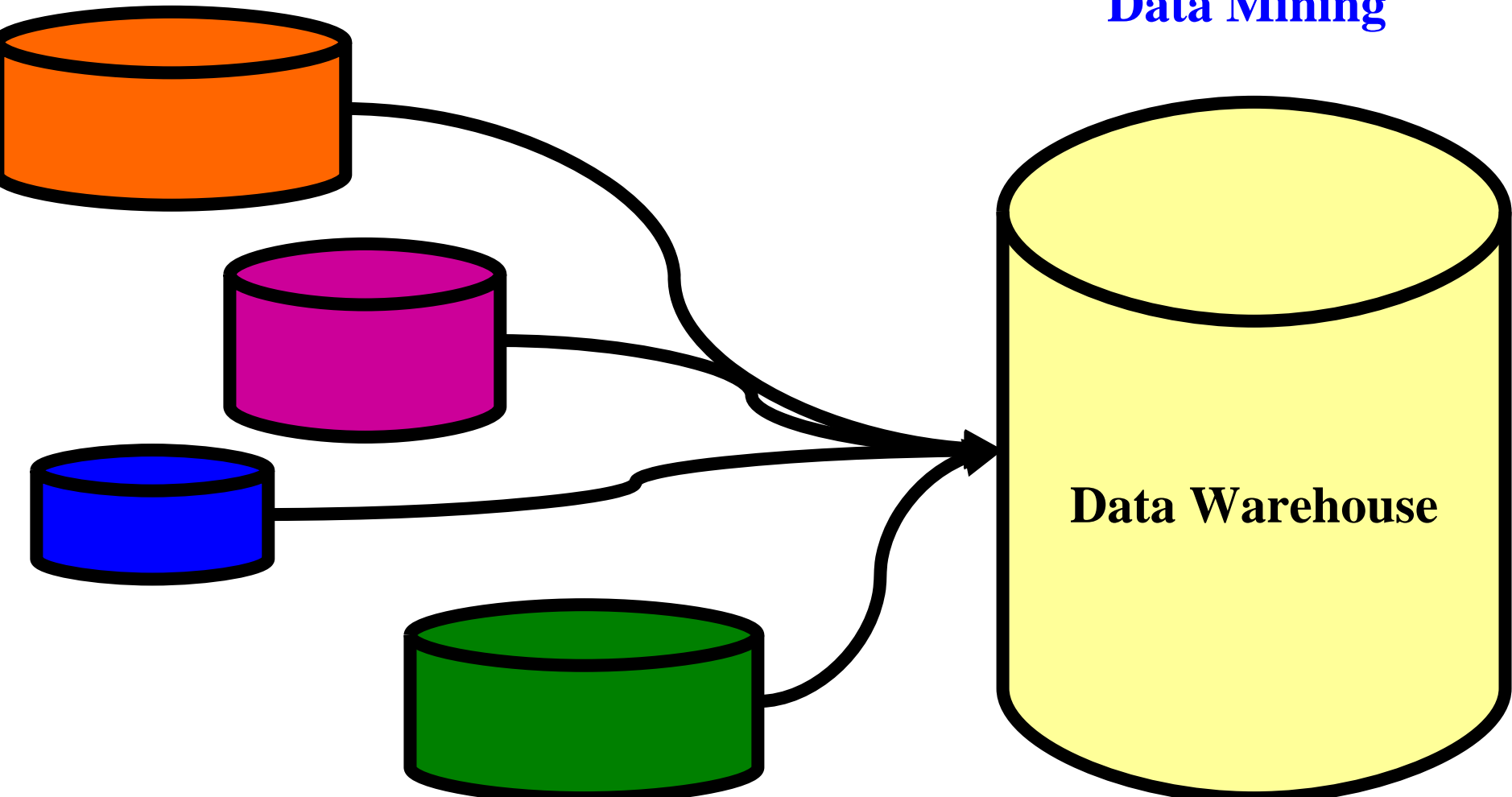
- Wie haben sich besondere offensive Marketingstrategien für bestimmte Produktlinien auf die Verkaufszahlen ausgewirkt?

Wenige, aber dafür sehr komplexe Anfragen!

Sammlung und periodische Auffrischung der Data Warehouse-Daten

**OLTP-Datenbanken
und andere Datenquellen**

**OLAP-Anfragen
Decision Support
Data Mining**



Das Stern-Schema

Verkäufe					
VerkDatum	Filiale	Produkt	Anzahl	Kunde	Verkäufer
25-Jul-00	Passau	1347	1	4711	825
...

Filialen			
Filialenkennung	Land	Bezirk	...
Passau	D	Bayern	...
...

Kunden			
KundenNr	Name	wiealt	...
4711	Kemper	43	...
...

Verkäufer					
VerkäuferNr	Name	Fachgebiet	Manager	wiealt	...
825	Handyman	Elektronik	119	23	...
...

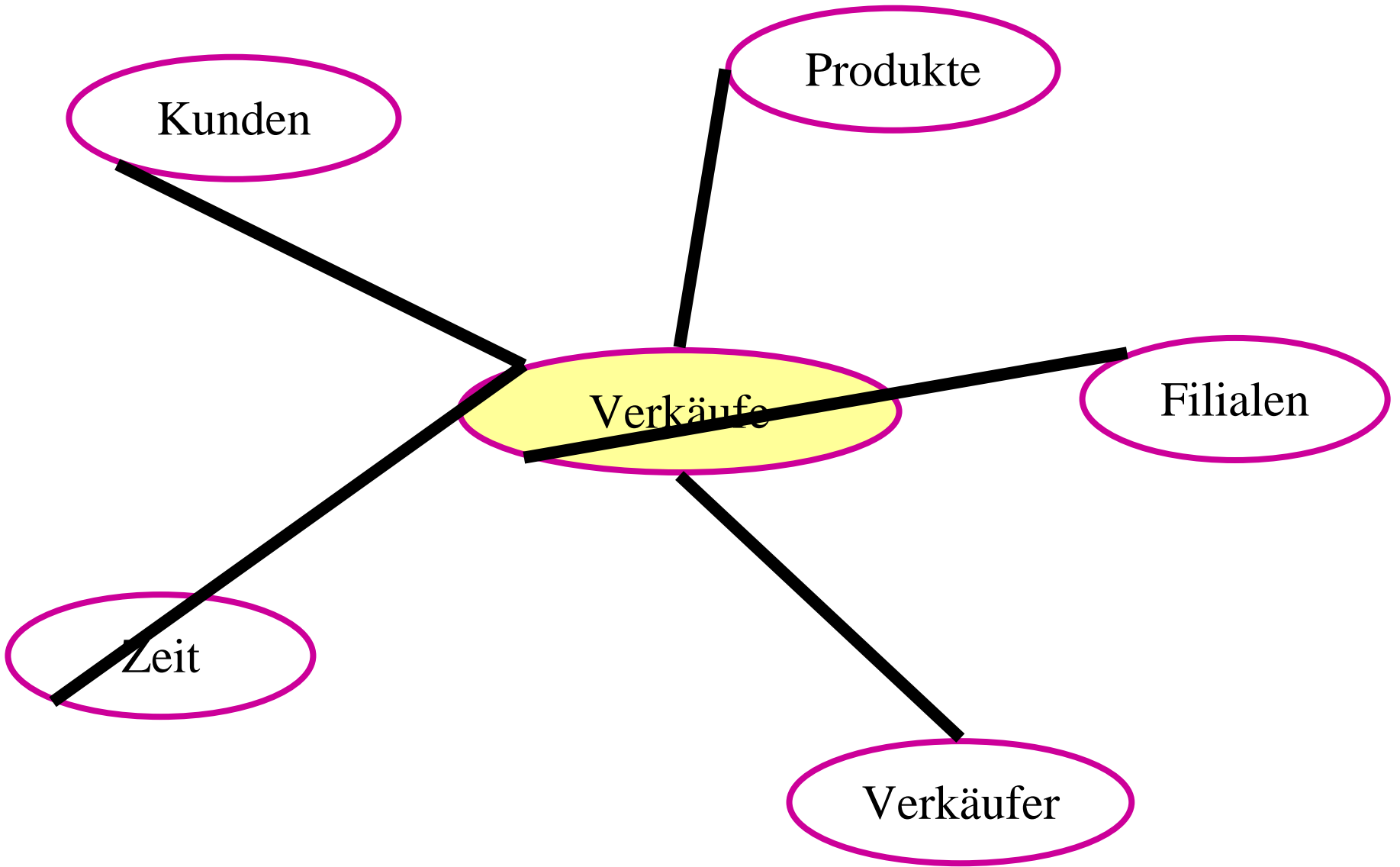
Zeit								
Datum	Tag	Monat	Jahr	Quartal	KW	Wochentag	Saison	...
...
25-Jul-00	25	Juli	2000	3	30	Dienstag	Hochsommer	...
...
18-Dec-01	18	Dezember	2001	4	52	Dienstag	Weihnachten	...
...

Produkte					
ProduktNr	Produkttyp	Produktgruppe	Produkthauptgruppe	Hersteller	...
1347	Handy	Mobiltelekom	Telekom	Siemens	...
...

Stern-Schema bei Data Warehouse-Anwendungen

- Eine sehr große Faktentabelle
 - Alle Verkäufe der letzten drei Jahre
 - Alle Telefonate des letzten Jahres
 - Alle Flugreservierungen der letzten fünf Jahre
 - normalisiert
- Mehrere Dimensionstabellen
 - Zeit
 - Filialen
 - Kunden
 - Produkt
 - Oft nicht normalisiert

Das Stern-Schema: Handelsunternehmen



Stern-Schema

Verkäufe					
VerkDatum	Filiale	Produkt	Anzahl	Kunde	Verkäufer
25-Jul-00	Passau	1347	1	4711	825
...

Faktentabelle (SEHR groß)

Filialen			
FilialenKennung	Land	Bezirk	...
Passau	D	Bayern	...
...

Kunden			
KundenNr	Name	wieAlt	...
4711	Kemper	43	...
...

Dimensionstabellen (relativ klein)

Verkäufer					
VerkäuferNr	Name	Fachgebiet	Manager	wieAlt	...
825	Handyman	Elektronik	119	23	...
...

Stern-Schema (cont'd)

Zeit							
Datum	Tag	Monat	Jahr	Quartal	KW	Wochentag	Saison
25-Jul-00	25	7	2000	3	30	Dienstag	Hochsommer
...		
18-Dec-01	18	12	2001	4	52	Dienstag	Weihnachten
...

Produkte					
ProduktNr	Produkttyp	Produktgruppe	Produkthauptgruppe	Hersteller	..
1347	Handy	Mobiltelekom	Telekom	Siemens	..
...

Nicht-normalisierte Dimensionstabellen: effizientere Anfrageauswertung

Zeit							
Datum	Tag	Monat	Jahr	Quartal	KW	Wochentag	Saison
25-Jul-00	25	7	2000	3	30	Dienstag	Hochsommer
...		
18-Dec-01	18	12	2001	4	52	Dienstag	Weihnachten
...

Datum → Monat → Quartal

Produkte					
ProduktNr	Produkttyp	Produktgruppe	Produkthauptgruppe	Hersteller	..
1347	Handy	Mobiltelekom	Telekom	Siemens	..
...

ProduktNr → Produkttyp → Produktgruppe → Produkthauptgruppe

Anfragen im Sternschema

select sum(v.Anzahl), p.Hersteller

from Verkäufe v, Filialen f, Produkte p, Zeit z, Kunden k

where z.Saison = 'Weihnachten' and

z.Jahr = 2001 and k.wieAlt < 30 and

p.Produkttyp = 'Handy' and f.Bezirk = 'Bayern' and

v.VerkDatum = z.Datum and v.Produkt = p.ProduktNr and

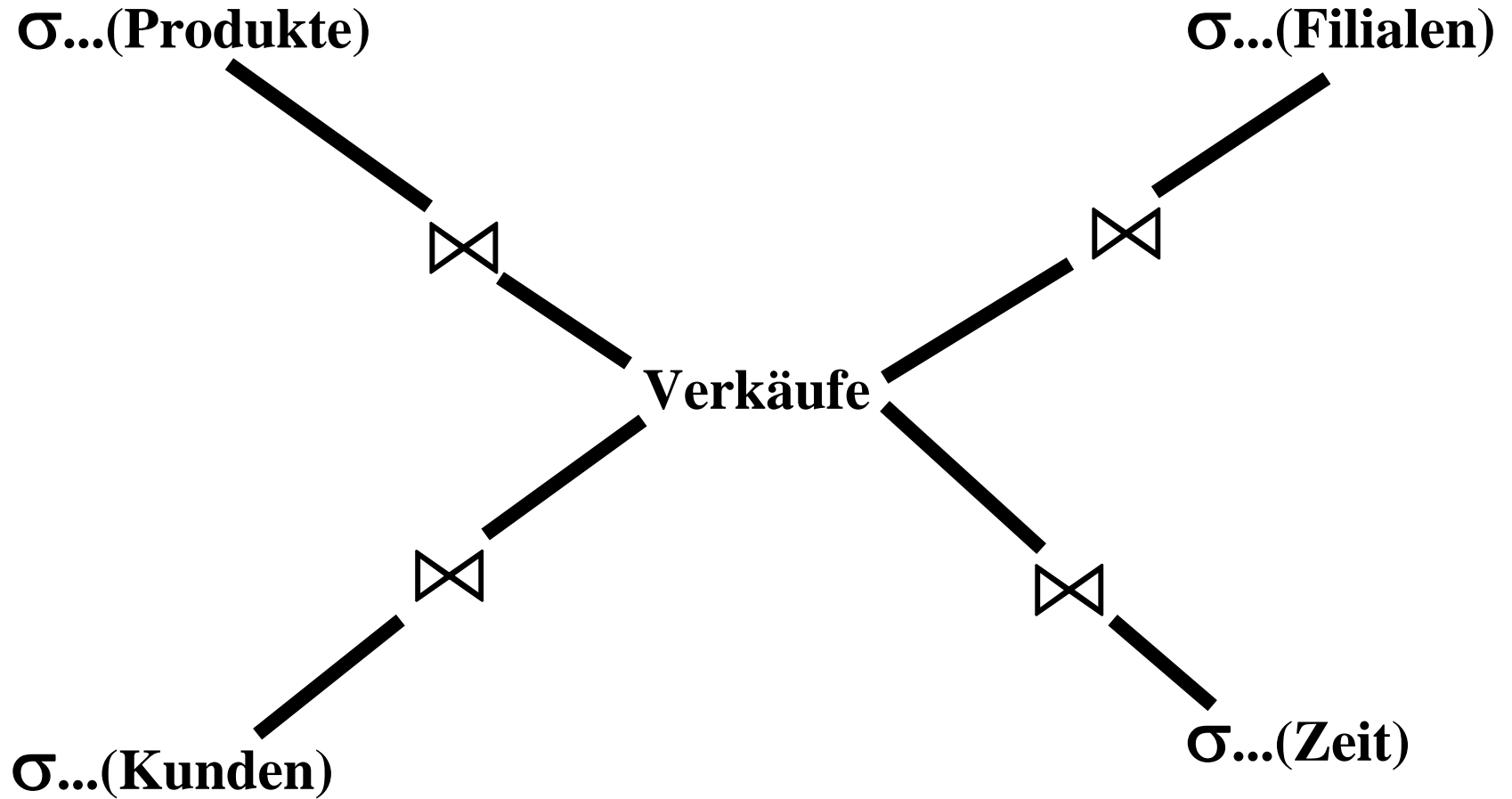
v.Filiale = f.FilialenKennung and v.Kunde = k.KundenNr

group by p.Hersteller;

Einschränkung
der Dimensionen

Join-Prädikate

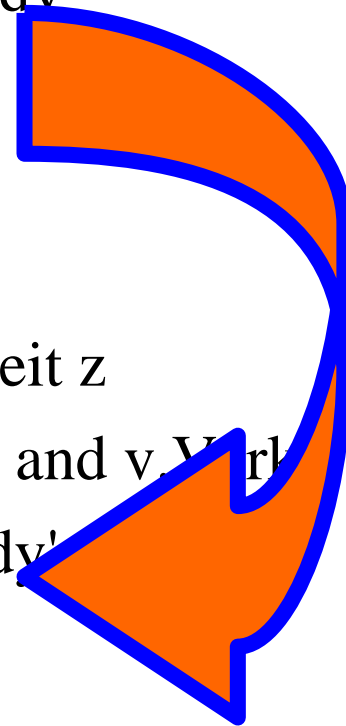
Algebra-Ausdruck



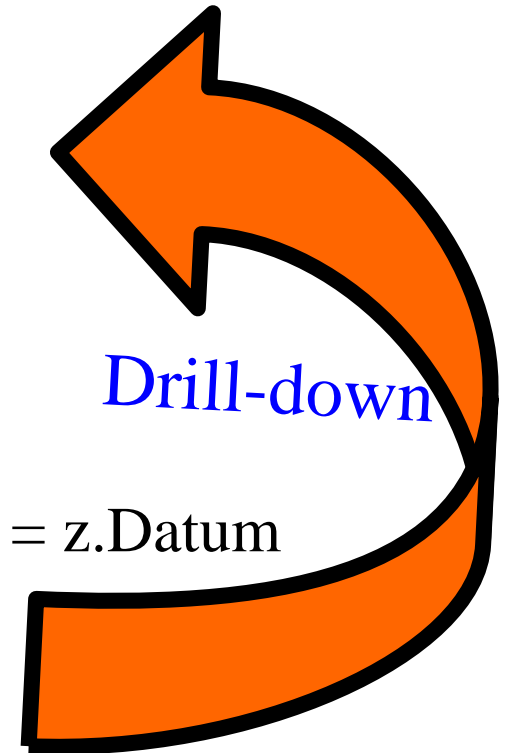
Roll-up/Drill-down-Anfragen

```
select Jahr, Hersteller, sum(Anzahl)
from Verkäufe v, Produkte p, Zeit z
where v.Produkt = p.ProduktNr and v.VerkDatum = z.Datum
        and p.Produkttyp = 'Handy'
group by p.Hersteller, z.Jahr;
```

```
select Jahr, sum(Anzahl)
from Verkäufe v, Produkte p, Zeit z
where v.Produkt = p.ProduktNr and v.VerkDatum = z.Datum
        and p.Produkttyp = 'Handy'
group by z.Jahr;
```



Roll-up



Drill-down

Handyverkäufe nach Hersteller und Jahr		
Hersteller	Jahr	Anzahl
Siemens	1999	2.000
Siemens	2000	3.000
Siemens	2001	3.500
Motorola	1999	1.000
Motorola	2000	1.000
Motorola	2001	1.500
Bosch	1999	500
Bosch	2000	1.000
Bosch	2001	1.500
Nokia	1999	1.000
Nokia	2000	1.500
Nokia	2001	2.000

Handyverkäufe nach Jahr	
Jahr	Anzahl
1999	4.500
2000	6.500
2001	8.500

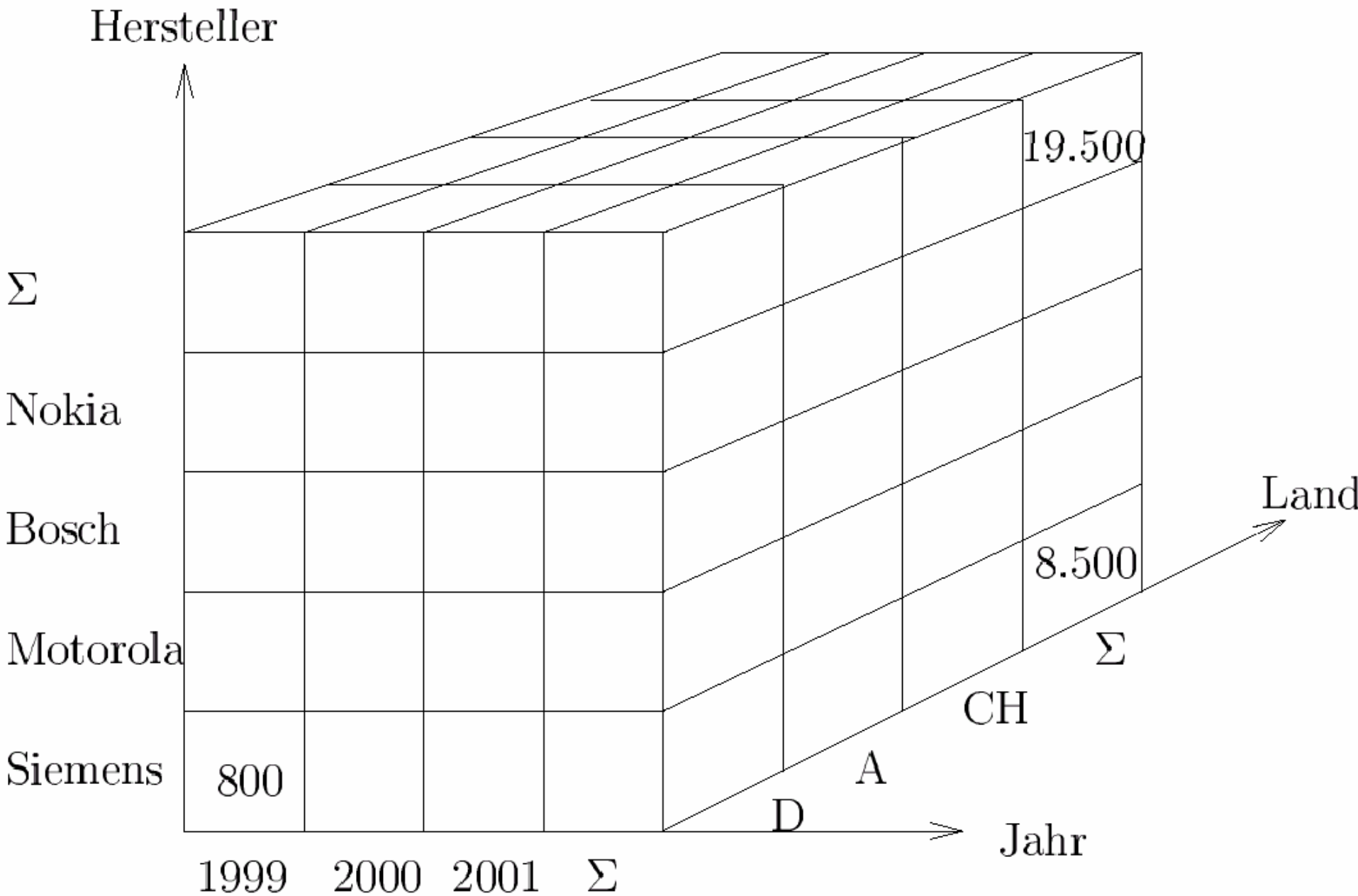
Handyverkäufe nach Hersteller	
Hersteller	Anzahl
Siemens	8.500
Motorola	3.500
Bosch	3.000
Nokia	4.500

Abb. 17.7: Analyse der Handyverkaufszahlen nach unterschiedlichen Dimensionen

Hersteller \ Jahr	1999	2000	2001	Σ
Siemens	2.000	3.000	3.500	8.500
Motorola	1.000	1.000	1.500	3.500
Bosch	500	1.000	1.500	3.000
Nokia	1.000	1.500	2.000	4.500
Σ	4.500	6.500	8.500	19.500

Abb. 17.8: Handyverkäufe nach Jahr und Hersteller

Würfeldarstellung



Der **cube**-Operator

select p.Hersteller, z.Jahr, f.Land, sum(v.Anzahl)

from Verkäufe v, Produkte p, Zeit z, Filialen f

where v.Produkt = p.ProduktNr and p.Produkttyp = 'Handy'

and v.VerkDatum = z.Datum and v.Filiale = f.Filialenkennung

group by cube (z.Jahr, p.Hersteller, f.Land);

Weiterführende Vorlesungen

Das Fachgebiet Wissensverarbeitung bietet regelmäßig an:

- **Vorlesungen**

- Knowledge Discovery
- Internet-Suchmaschinen
- Künstliche Intelligenz (Master)

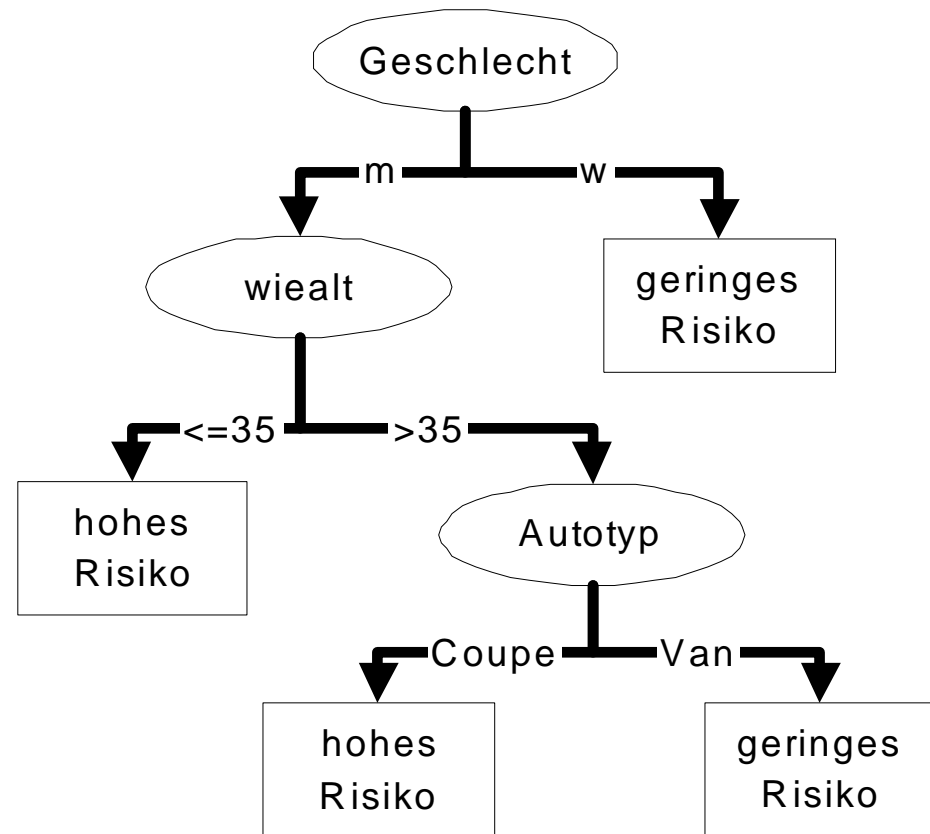
- **Projekte** und **Abschlussarbeiten** rund um unser Datenbank-basiertes soziale Lesezeichensystem BibSonomy:

<http://www.bibsonomy.org>

Vorlesung Knowledge Discovery

Bsp.: Klassifikation/Entscheidungsbaum

Schadenshöhe			
wiealt	Geschlecht	Autotyp	Schäden
45	w	Van	gering
18	w	Coupé	gering
22	w	Van	gering
38	w	Coupé	gering
19	m	Coupé	hoch
24	m	Van	hoch
40	m	Coupé	hoch
40	m	Van	gering
⋮	⋮	⋮	⋮



$(\text{wieAlt} > 35) \wedge (\text{Geschlecht} = \text{'m'}) \wedge (\text{Autotyp} = \text{'Coupé'}) \rightarrow (\text{Risiko} = \text{'hoch'})$

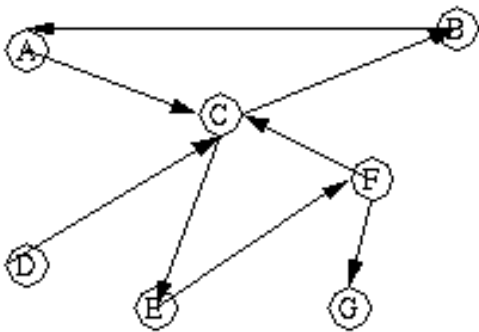
Wie werden Entscheidungs-/ Klassifikationsbäume erstellt

- Trainingsmenge
 - Große Zahl von Datensätzen, die in der Vergangenheit gesammelt wurden
 - Sie dient als Grundlage für die Vorhersage von „neu ankommenden“ Objekten
 - Beispiel: neuer Versicherungskunde wird gemäß dem Verhalten seiner „Artgenossen“ eingestuft
- Rekursives Partitionieren
 - Fange mit einem Attribut an und spalte die Tupelmengen
 - Jede dieser Teilmengen wird rekursiv weiter partitioniert, bis nur noch gleichartige Objekte in der jeweiligen Partition sind

Vorlesung Internet-Suchmaschinen

Bsp.: Ranking in Google

gecrawlter Web-Graph



Index erstellen



Indexterm	df	D_i, t_j
Computer	3	D _{7, 4} [] [] []
Datenbank	2	D _{1, 3} [] [] []
...		
Wissenschaft	4	D _{2, 4} [] [] [] [] []
System	1	D _{5, 2} [] [] [] [] []

Indexdatei

Posting-Listen

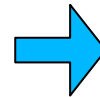


PageRank berechnen



PageRank

1. C
2. E
3. B
4. A
5. F
6. G
7. D



Vorlesung Künstliche Intelligenz

Wie können komplexere Informationen verwaltet werden?

