Moduldokumentation

Modul Distributed Data Management (ddm)

Simon Wächter

2019

Inhalt

[1 Einleitung 4](#_Toc2105870)

[1.1 Einleitung 4](#_Toc2105871)

[1.2 Lernziele 4](#_Toc2105872)

[1.3 Prüfungen 4](#_Toc2105873)

[2 Woche 1 5](#_Toc2105874)

[2.1 Gliederung 5](#_Toc2105875)

[2.2 Fachvertiefung 5](#_Toc2105876)

[2.3 Verteiltes Rechnen 5](#_Toc2105877)

[2.3.1 Verteiltes Rechnen (Distributed Computing) 5](#_Toc2105878)

[2.3.2 Was wird verteilt 6](#_Toc2105879)

[2.4 Einordnung der Thematik 6](#_Toc2105880)

[2.4.1 Verteilt, mobil, parallel … 6](#_Toc2105881)

[2.4.2 … Big Data 6](#_Toc2105882)

[2.5 Verteilte Datenbanksysteme 7](#_Toc2105883)

[2.5.1 Verteilte Datenbanksysteme 7](#_Toc2105884)

[2.5.2 Klassifizierung 7](#_Toc2105885)

[2.5.3 DDBS (P2P) Architecture 8](#_Toc2105886)

[2.5.4 MDBS Architecture 9](#_Toc2105887)

[2.5.5 Date’s 12 Regeln 9](#_Toc2105888)

[2.5.6 Aspekte Verteilter Datenbanksysteme 10](#_Toc2105889)

[2.6 Parallele Datenbanksysteme 10](#_Toc2105890)

[2.6.1 Parallele Datenbanksysteme 10](#_Toc2105891)

[2.6.2 Architekturen 10](#_Toc2105892)

[2.6.3 Aspekte paralleler Datenbanksysteme 11](#_Toc2105893)

[2.7 NoSQL Systeme 11](#_Toc2105894)

[2.7.1 NoSQL Systeme 11](#_Toc2105895)

[2.7.2 Konzepte NoSQL Systeme 11](#_Toc2105896)

[2.8 Mobile Datenbanksysteme 12](#_Toc2105897)

[2.9 Semantic Web 12](#_Toc2105898)

[2.10 Gliederung 13](#_Toc2105899)

[2.11 Was sind Trigger? 13](#_Toc2105900)

[2.12 Wozu dienen Trigger? 13](#_Toc2105901)

[2.13 Trigger Konzept 14](#_Toc2105902)

[2.14 Triggertypen 14](#_Toc2105903)

[2.14.1 Triggertypen: Ereignis 14](#_Toc2105904)

[2.14.2 Triggertypen: Timing 14](#_Toc2105905)

[2.14.3 Triggertypen: Granulat 15](#_Toc2105906)

[2.15 DML-Trigger auf Tabellen 15](#_Toc2105907)

[2.16 PL/SQL Block 16](#_Toc2105908)

[2.16.1 Einfacher PL/SQL Block 16](#_Toc2105909)

[2.16.2 Bildschirmausgabe 16](#_Toc2105910)

[2.16.3 Datentypen 17](#_Toc2105911)

[2.16.4 Anweisung 17](#_Toc2105912)

[2.17 1. Erfahrung mit Trigger 18](#_Toc2105913)

[2.17.1 Beispiel-Datenbasis 18](#_Toc2105914)

[2.17.2 1. Erfahrung mit Trigger 19](#_Toc2105915)

[2.18 Konsistent Prüfen 20](#_Toc2105916)

[3 Woche 2 22](#_Toc2105917)

[3.1 Inhalt 22](#_Toc2105918)

[3.2 Einführung 22](#_Toc2105919)

[3.2.1 Ausganglage 22](#_Toc2105920)

[3.2.2 Datenbank Beispiel 22](#_Toc2105921)

[3.2.3 Verteilte Anwendungen 23](#_Toc2105922)

[3.2.4 Entwurfsproblem 23](#_Toc2105923)

[3.2.5 Alternative Entwurfsstrategien 24](#_Toc2105924)

[3.2.6 Top-down Entwurf 24](#_Toc2105925)

[3.3 Verteilungsentwurf 25](#_Toc2105926)

[3.3.1 Kernpunkte des Verteilungsentwurfs 25](#_Toc2105927)

[3.3.2 Grad der Fragmentierung 25](#_Toc2105928)

[3.3.3 Arten der Fragmentierung 25](#_Toc2105929)

[3.3.4 Möglichkeit der Fragmentierung – Horizontal 26](#_Toc2105930)

[3.3.5 Möglichkeit der Fragmentierung – Vertikal 26](#_Toc2105931)

[3.3.6 Korrektheit der Fragmentierung 27](#_Toc2105932)

[3.4 PHF 27](#_Toc2105933)

[3.4.1 PHF – Information über Daten 27](#_Toc2105934)

[3.4.2 PHF – Information über Anwendungen 28](#_Toc2105935)

[3.4.3 Minterm Beispiele 28](#_Toc2105936)

[3.4.4 PHF – Information über Anwendungen 28](#_Toc2105937)

[3.5 Verfahren 29](#_Toc2105938)

[3.5.1 PHF Verfahren, 1. Schritt 29](#_Toc2105939)

[3.5.2 Vollständigkeit und Minimalität 29](#_Toc2105940)

[3.5.3 PHF – Verfahren, 2. Schritt 29](#_Toc2105941)

[3.5.4 PHF – Verfahren, 3. Schritt 29](#_Toc2105942)

[3.5.5 PHF – Beispiel 30](#_Toc2105943)

[3.6 DHF 31](#_Toc2105944)

[3.6.1 Abgeleitete horizontale Fragmentierung 31](#_Toc2105945)

[3.6.2 DHF - Beispiel 32](#_Toc2105946)

[3.7 Tafelmitschrift 32](#_Toc2105947)

[4 Woche 3 34](#_Toc2105948)

# Einleitung

## Einleitung

Dieses Dokument stellt die Moduldokumentation für das Modul ddm dar. Allfällige Unterlagen sind im Modulordner zu finden.

## Lernziele

Skalierbarkeit, Verfügbarkeit und Ausfallstoleranz bilden Schlüsselanforderungen an heutige verteilte Informationssysteme. In diesem Modul vertiefen sich die Studierenden in die grundlegenden Konzepte und Technologien verteilter Datenhaltung und -verarbeitung. Dabei wird sowohl auf die klassischen Ansätze verteilter Datenbanksysteme eingegangen als auch neuartige Paradigmen der NoSQL Systeme besprochen.

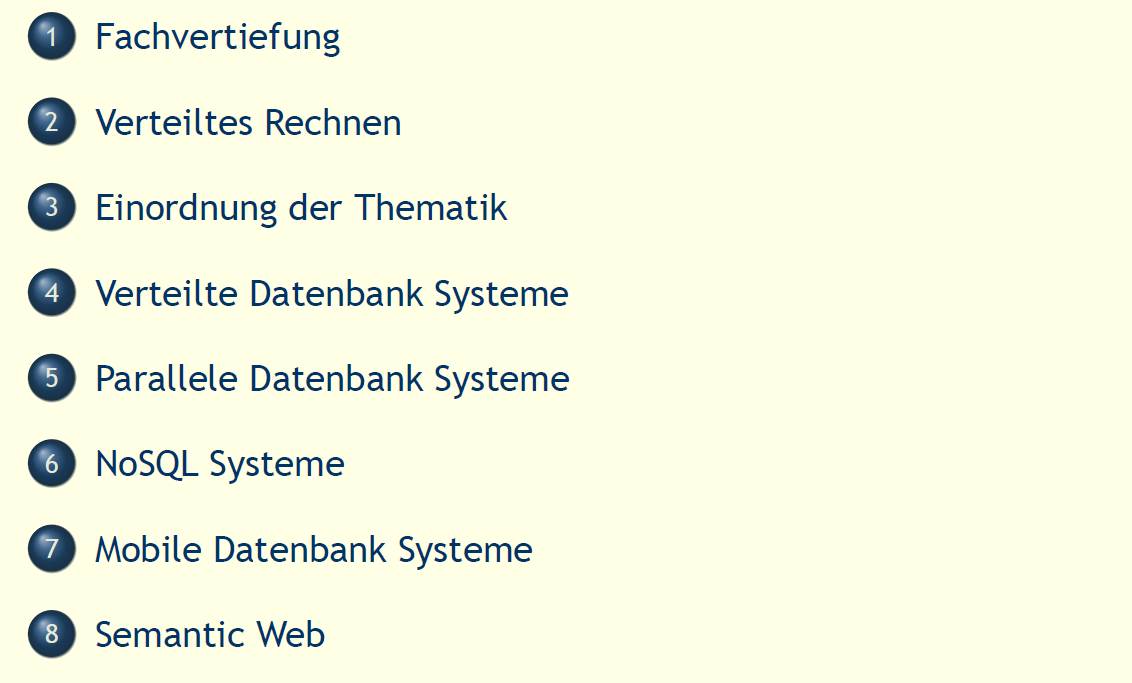
* Aktive Datenbanken:
  + Stored Procedures
  + ECA Prinzip
  + Trigger
* Verteilte Datenbanksysteme:
  + Architektur verteilter Datenbanksysteme
  + verteilter Datenentwurf
  + verteilte Anfrageverarbeitung
  + verteilte Transaktionen
  + Replikation
* NoSQL:
  + Daten-, Speicher- und Anfragemodelle
  + Key/Value, Wide Column, Document, Graph
  + Map/Reduce
  + Partitionierung (Sharding), verteilte Hashtabellen
  + Konsistenz- und Replikationsmodelle
  + CAP Theorem, BASE
* NoSQL Systeme: Cassandra, MongoDB, HBase, Neo4j

## Prüfungen

Die Modulnote setzt sich aus einer Erfahrungsnote zu 50% mit zwei Assessments zu je 25% und einer Modulschlussprüfung zu 50% zusammen.

# Woche 1

## Gliederung

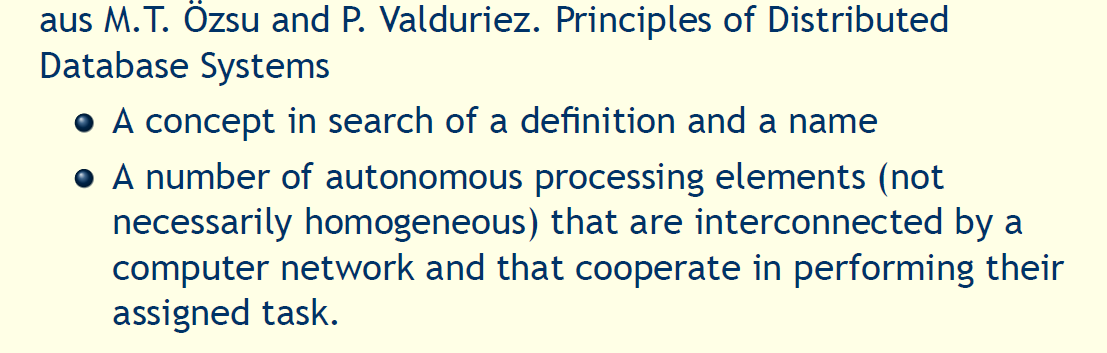


## Fachvertiefung



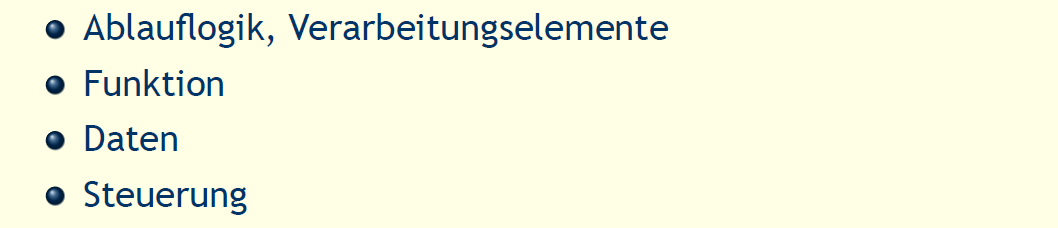
## Verteiltes Rechnen

### Verteiltes Rechnen (Distributed Computing)



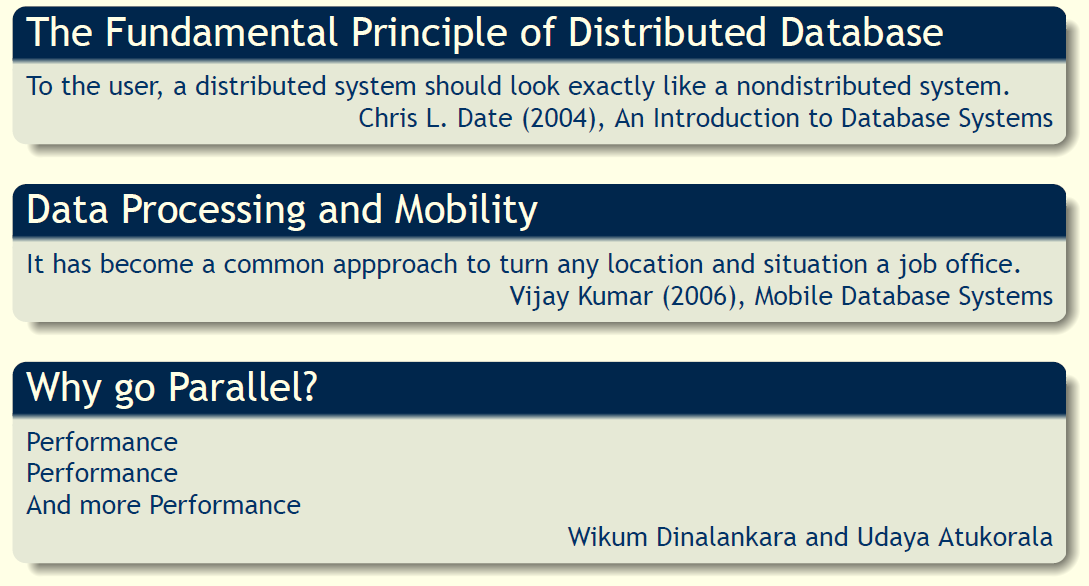
In verteilten Systemen kann die Ablauflogik, Daten, Steuerung verteilt sein. Für einen Benutzer verhält sich ein verteiltes System wie ein nicht verteiltes System. Warum sollte man Parallelisieren? -> Performance, Performance, Performance

### Was wird verteilt

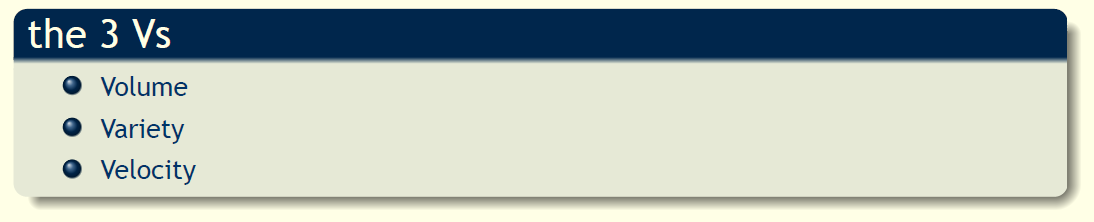


## Einordnung der Thematik

### Verteilt, mobil, parallel …



### … Big Data

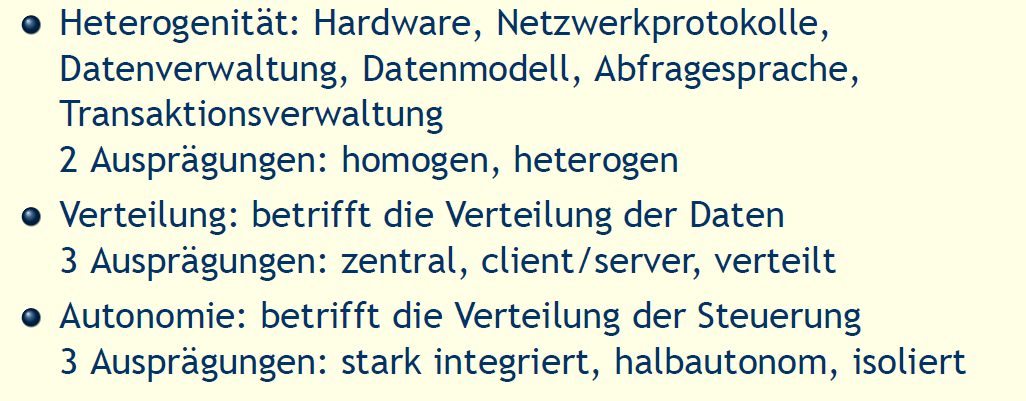


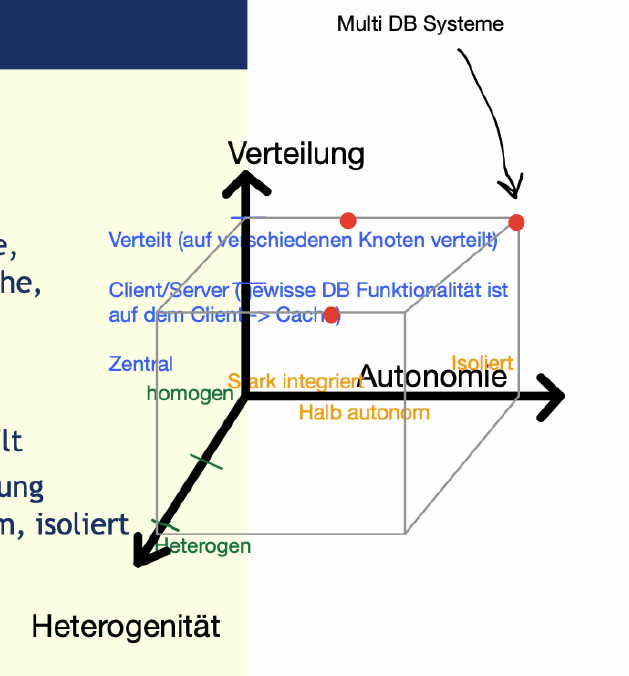
## Verteilte Datenbanksysteme

### Verteilte Datenbanksysteme

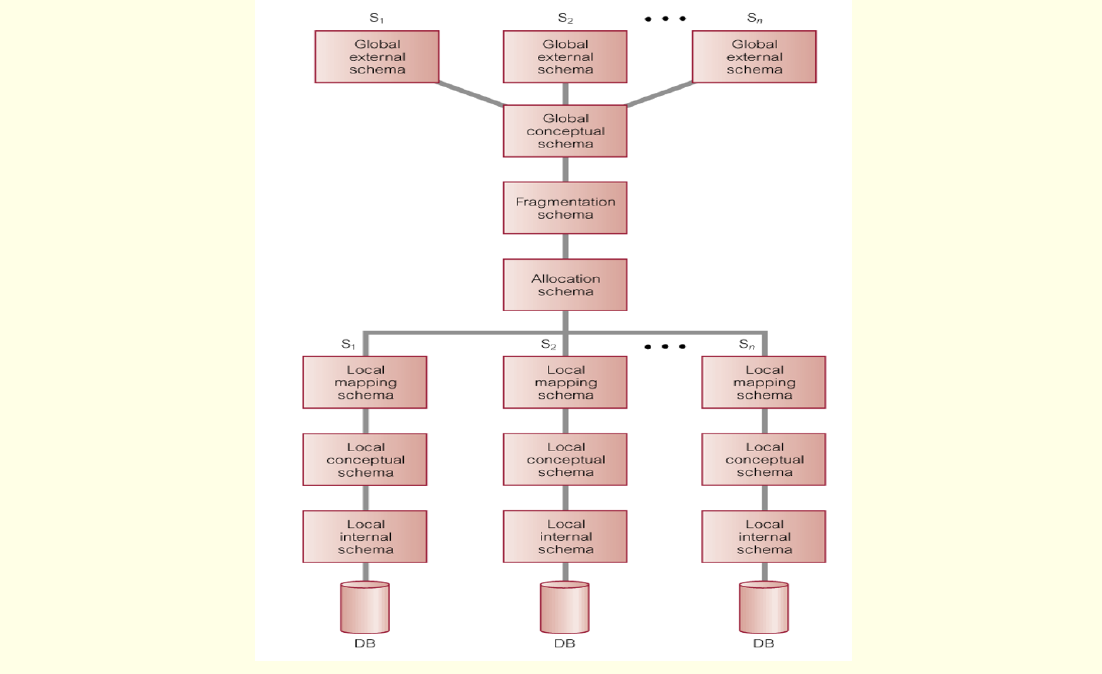


### Klassifizierung

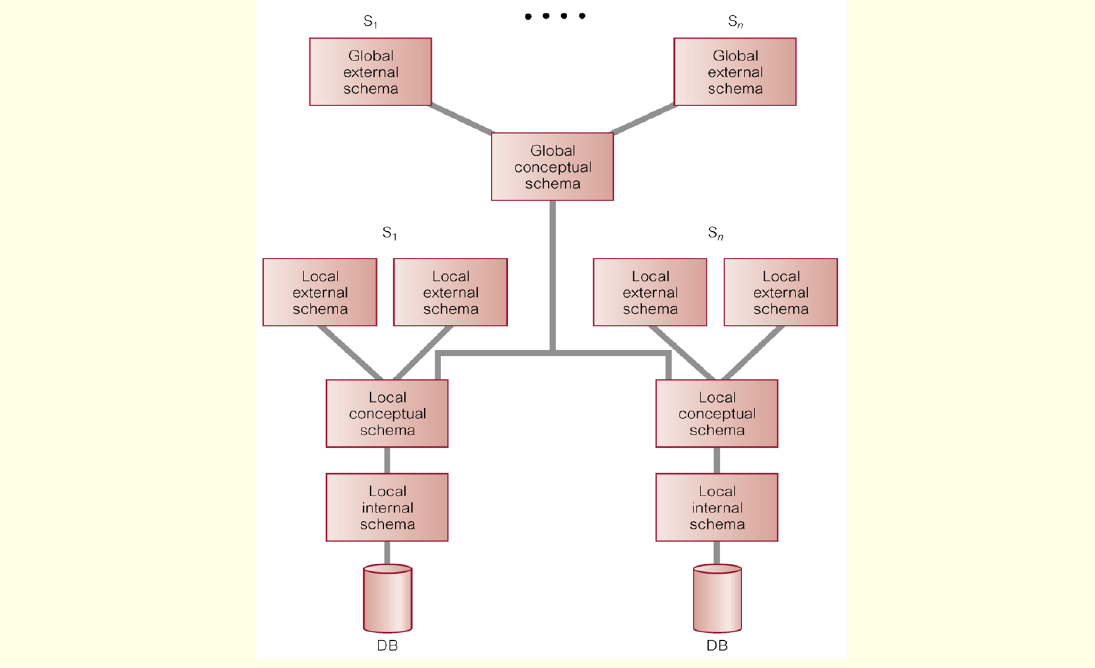




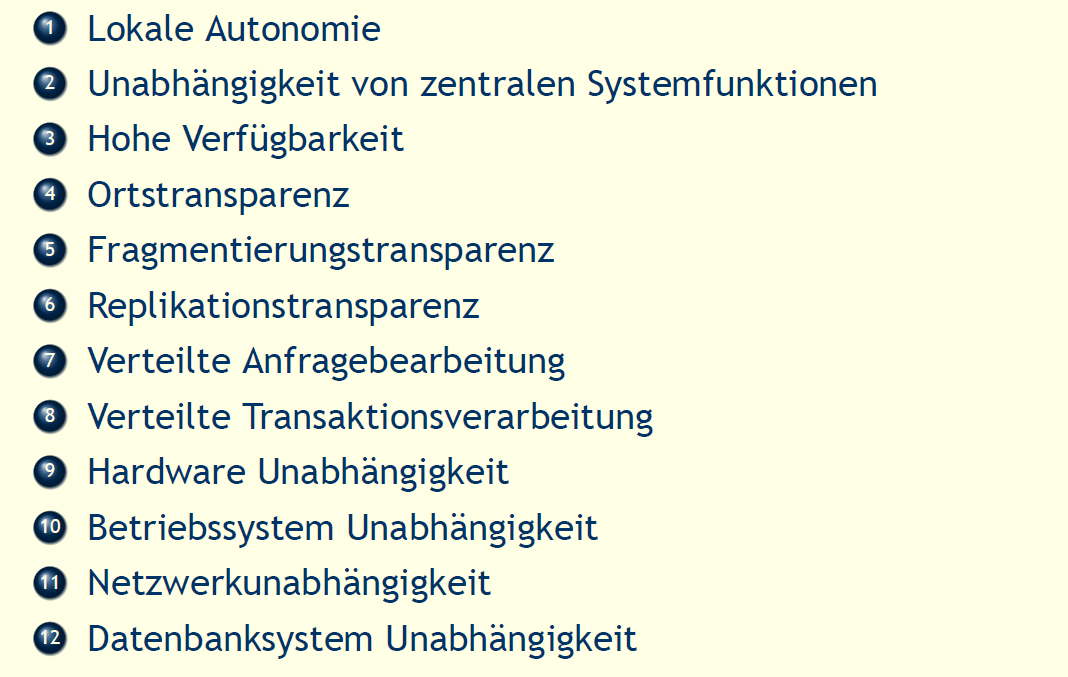
### DDBS (P2P) Architecture



### MDBS Architecture



### Date’s 12 Regeln

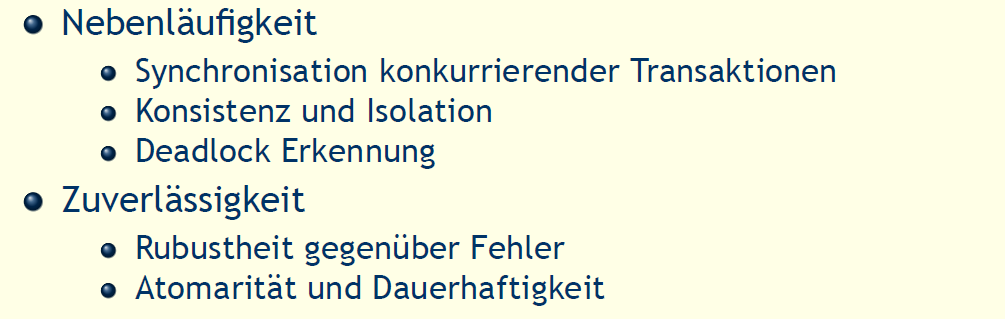


Ortstransparenz und Namenstransparenz:

* Kann mittels View oder Synonymen erstellt werden.
* Damit die Orts- und Namenstransparenz auf jeder DB genutzt werden kann müssen die Links
* symmetrisch sein. Als von jeder DB ein Link auf die anderen erstellt werden.

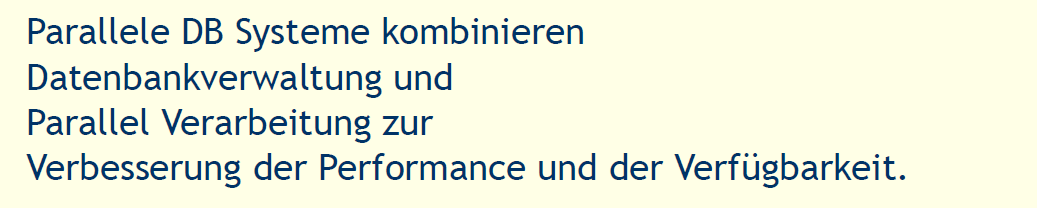
### Aspekte Verteilter Datenbanksysteme





## Parallele Datenbanksysteme

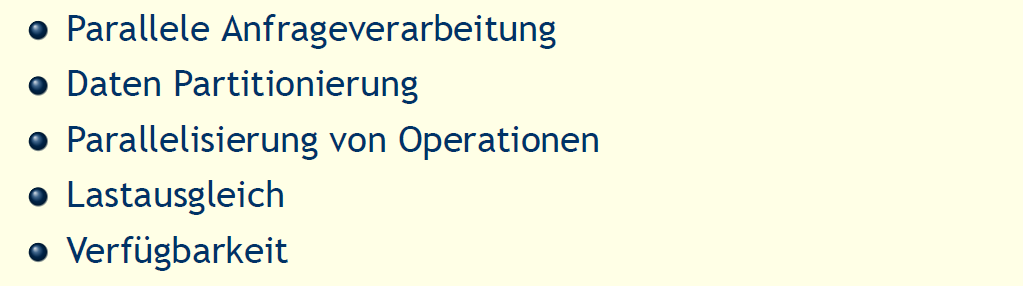
### Parallele Datenbanksysteme



### Architekturen

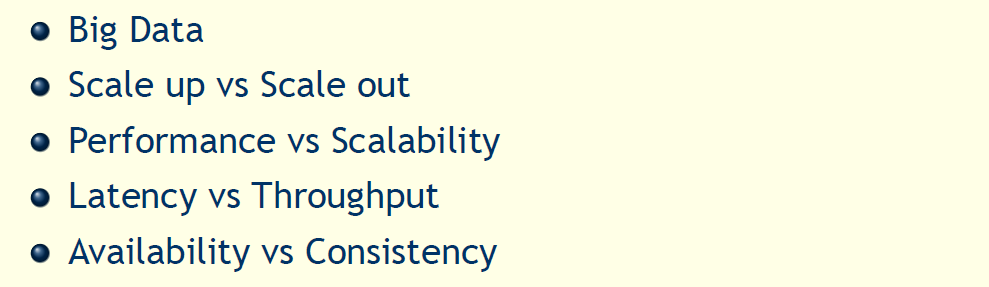


### Aspekte paralleler Datenbanksysteme



## NoSQL Systeme

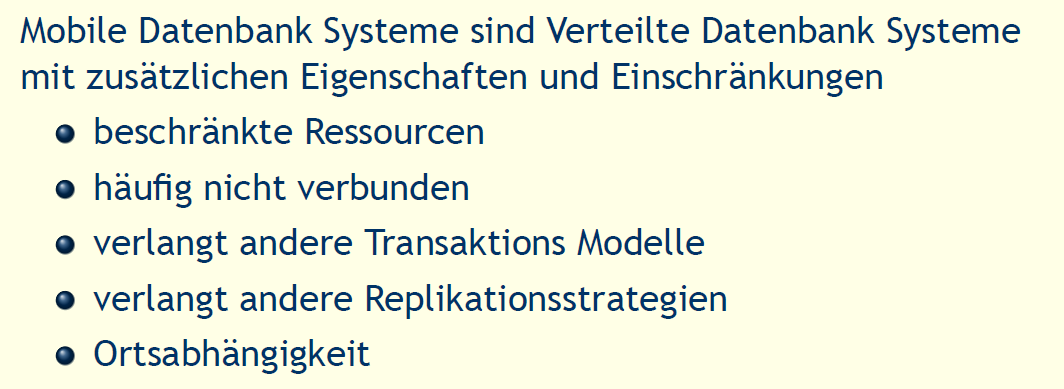
### NoSQL Systeme



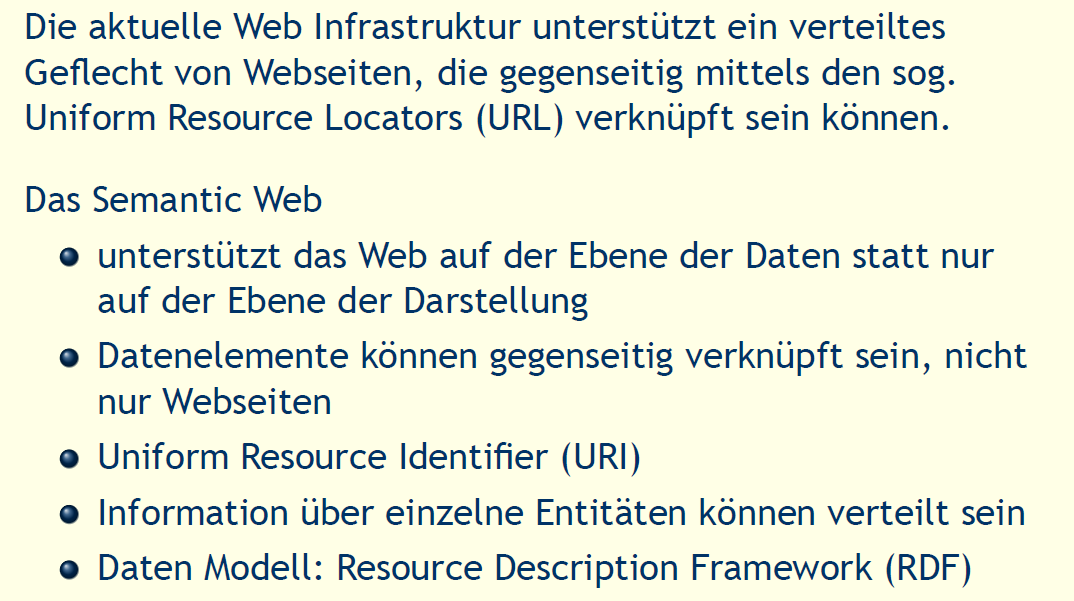
### Konzepte NoSQL Systeme



## Mobile Datenbanksysteme



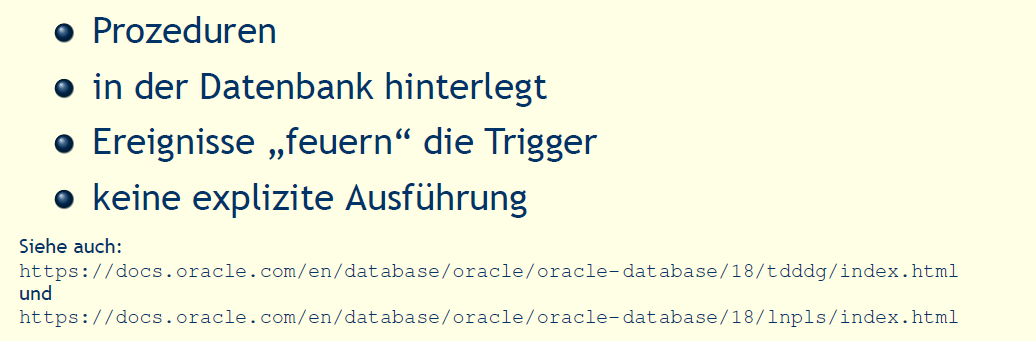
## Semantic Web



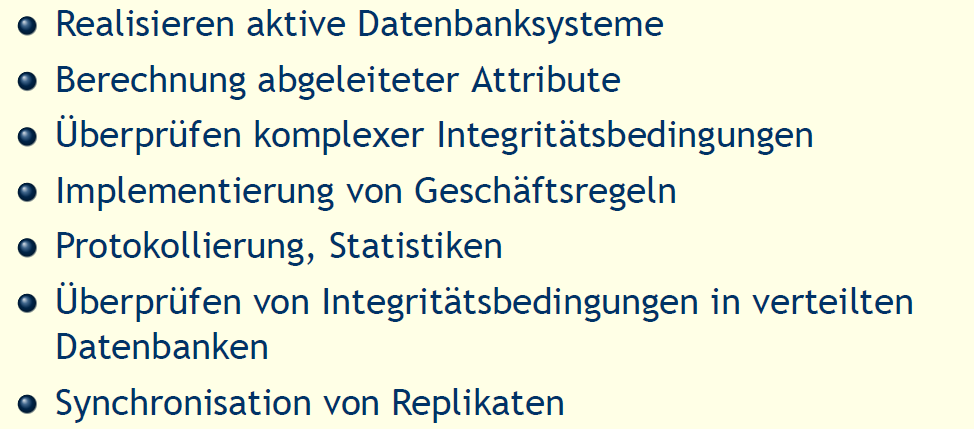
## Gliederung



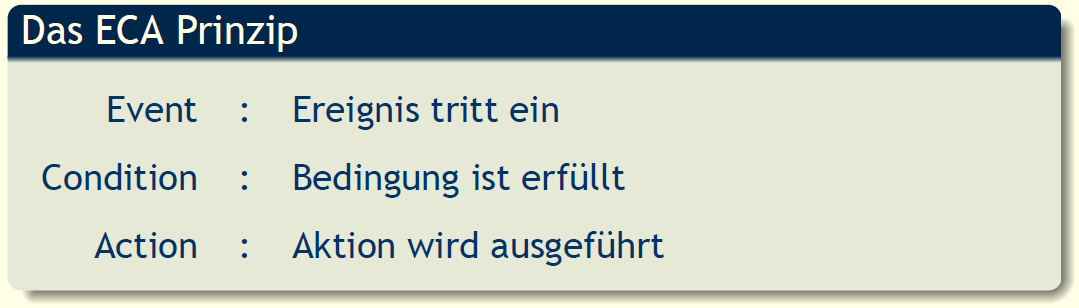
## Was sind Trigger?



## Wozu dienen Trigger?

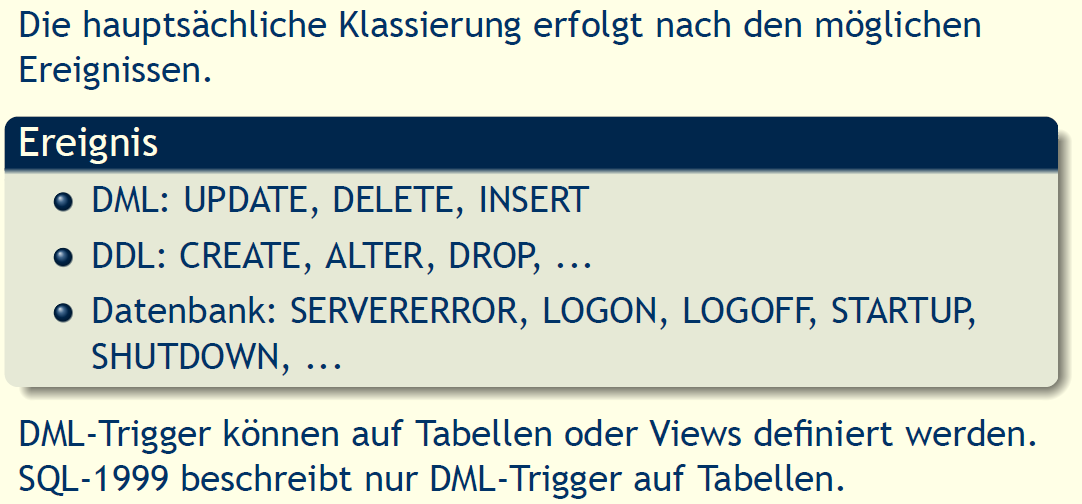


## Trigger Konzept

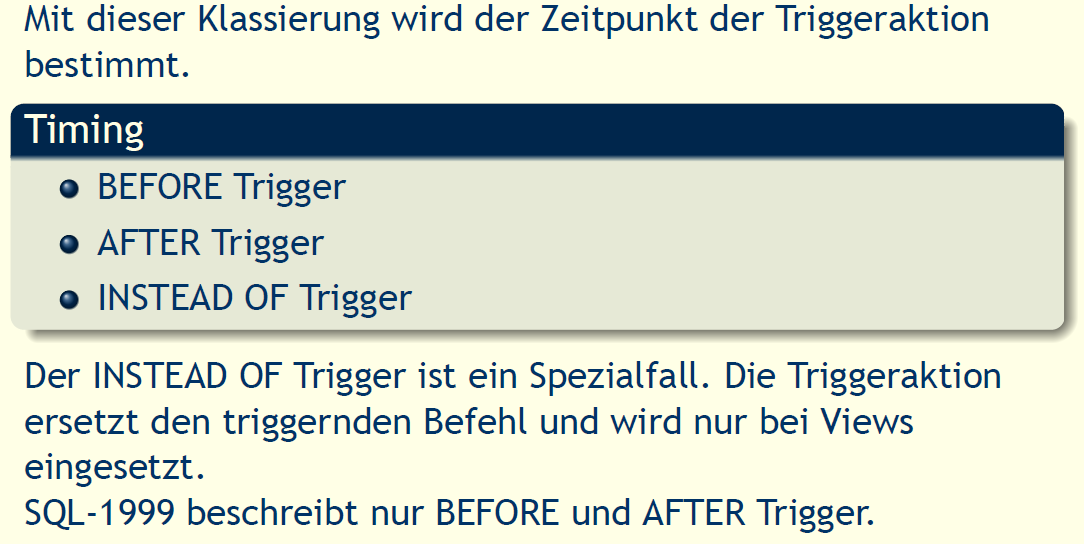


## Triggertypen

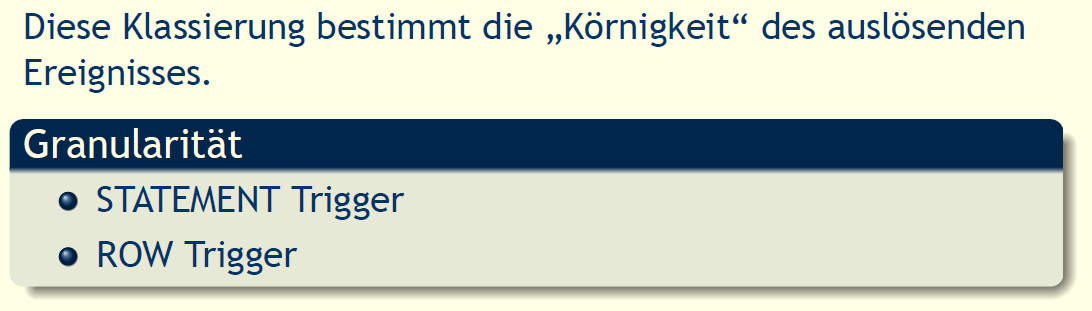
### Triggertypen: Ereignis



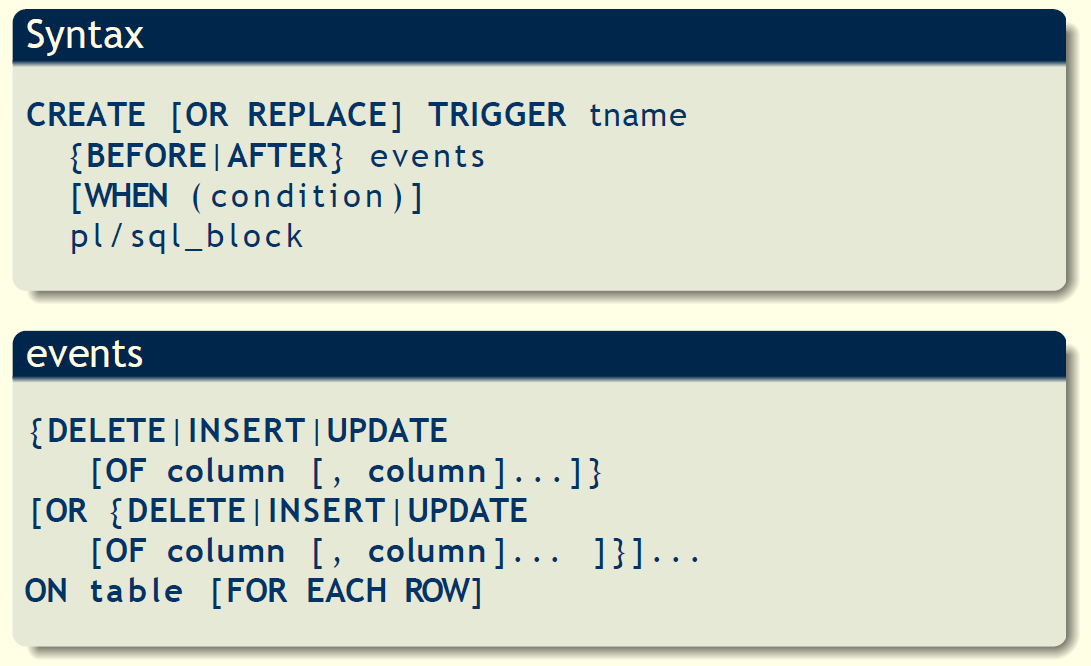
### Triggertypen: Timing



### Triggertypen: Granulat



## DML-Trigger auf Tabellen

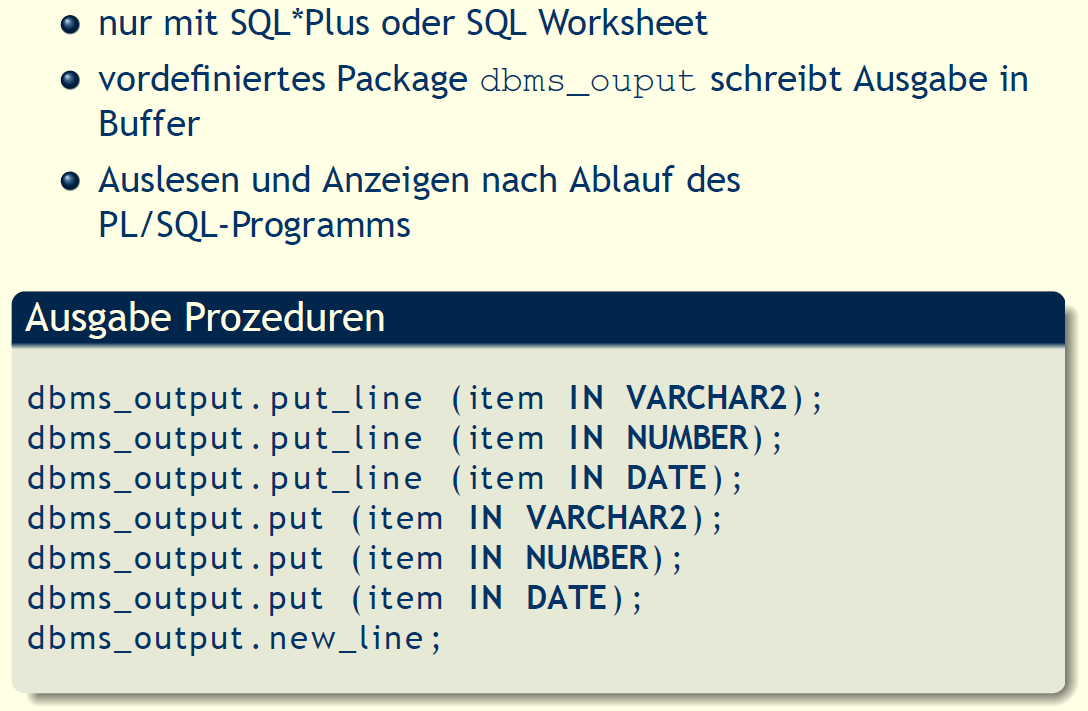


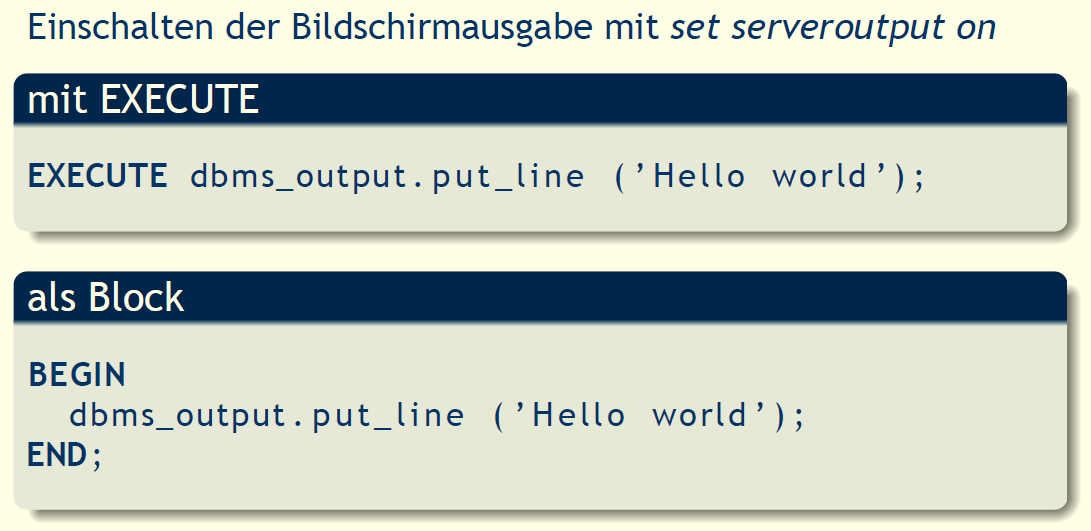
## PL/SQL Block

### Einfacher PL/SQL Block

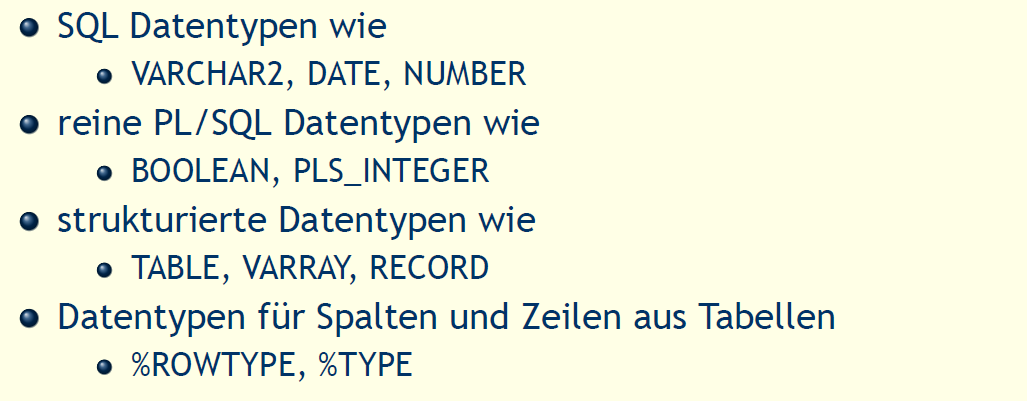


### Bildschirmausgabe

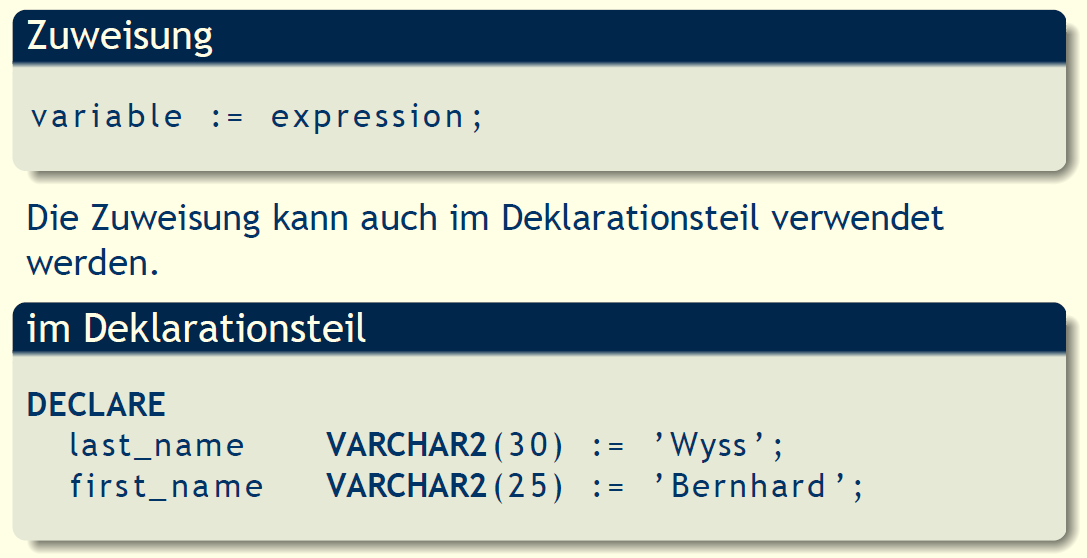


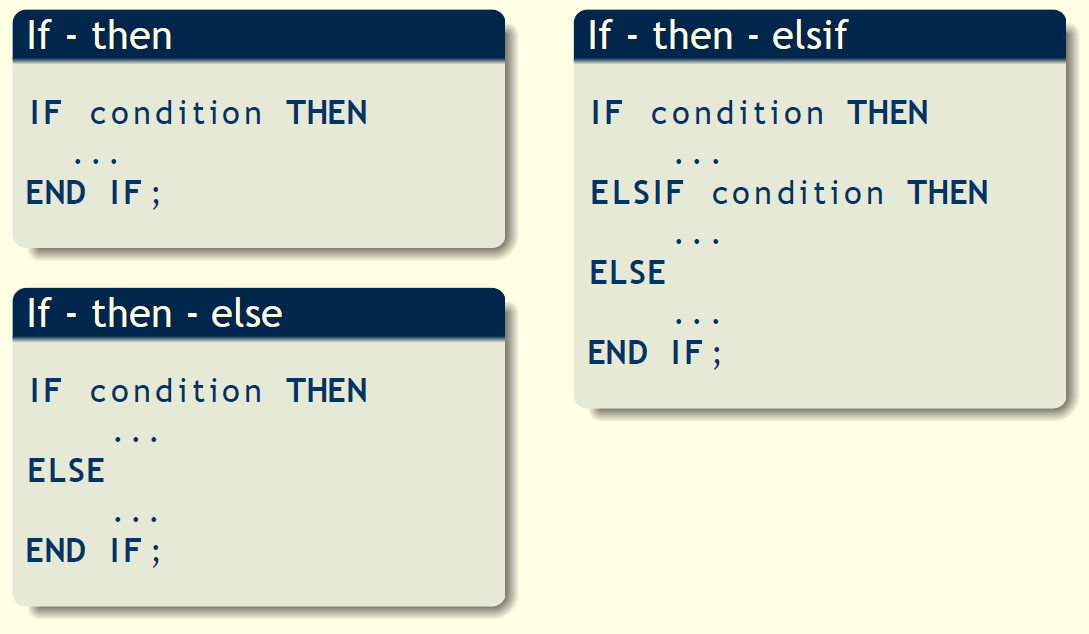


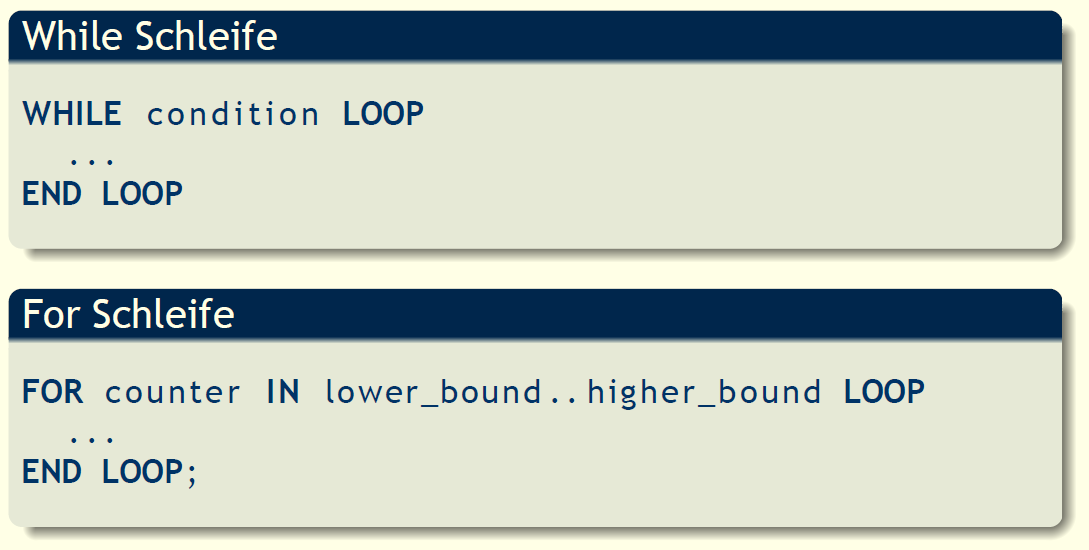
### Datentypen



### Anweisung

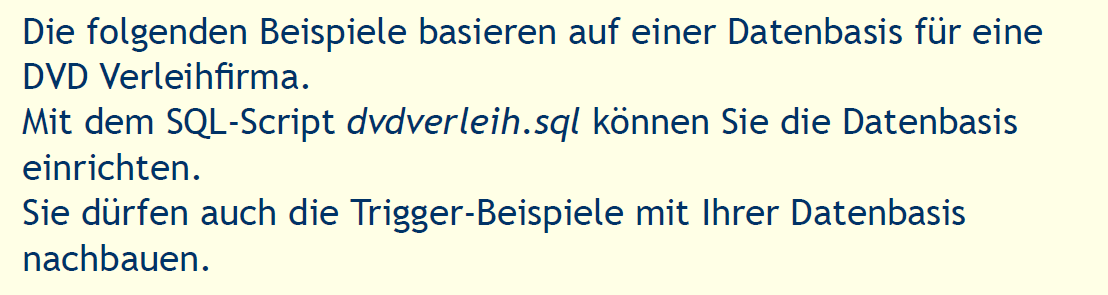




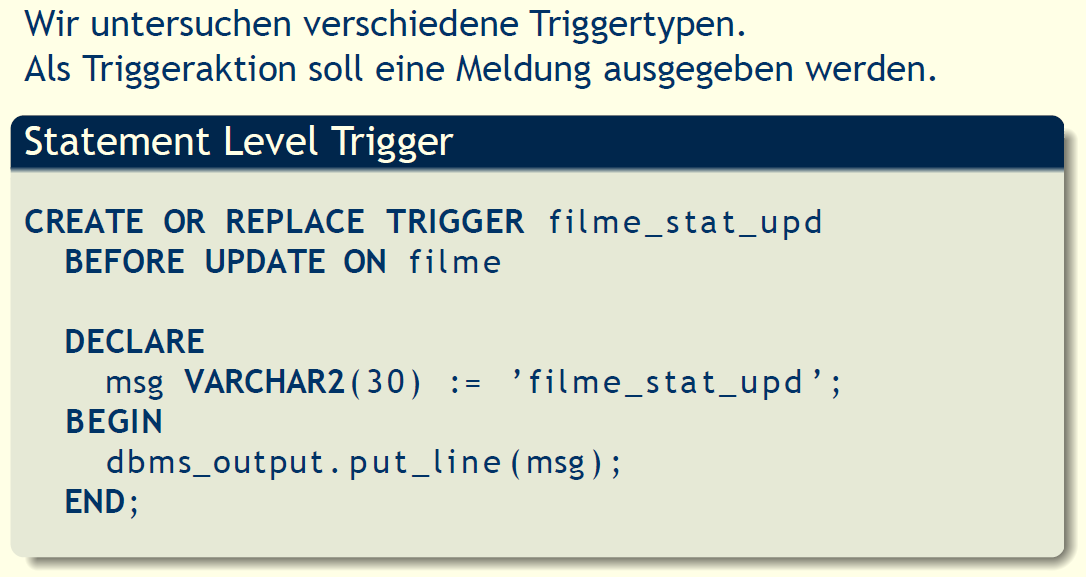


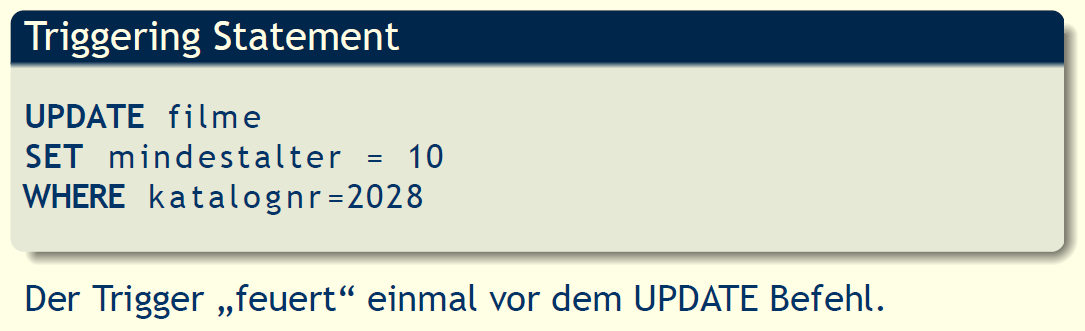
## 1. Erfahrung mit Trigger

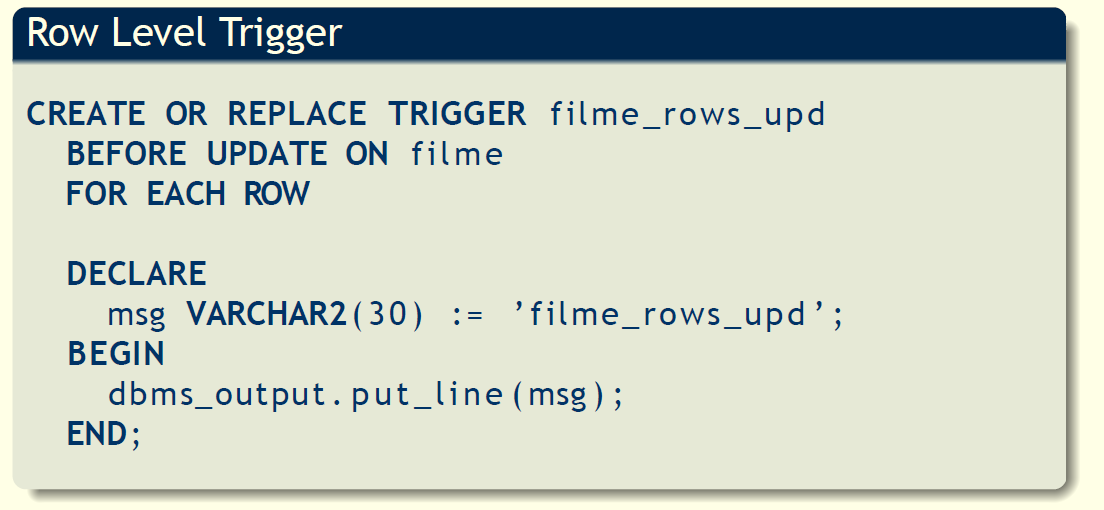
### Beispiel-Datenbasis

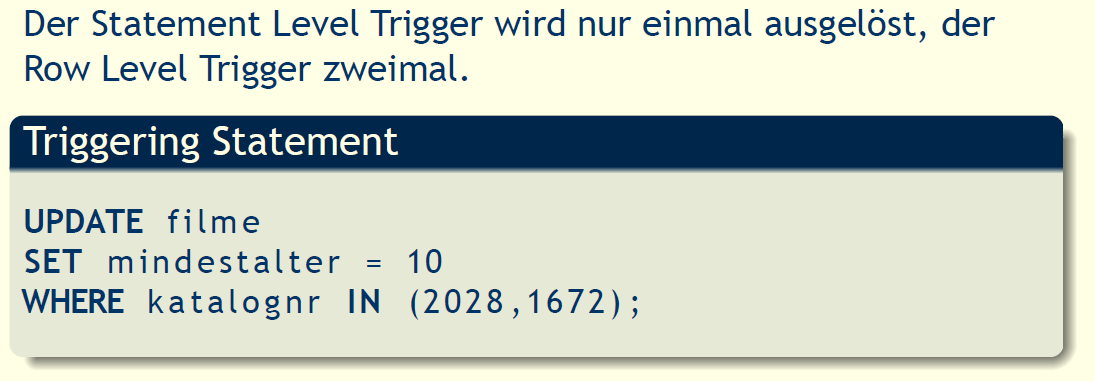


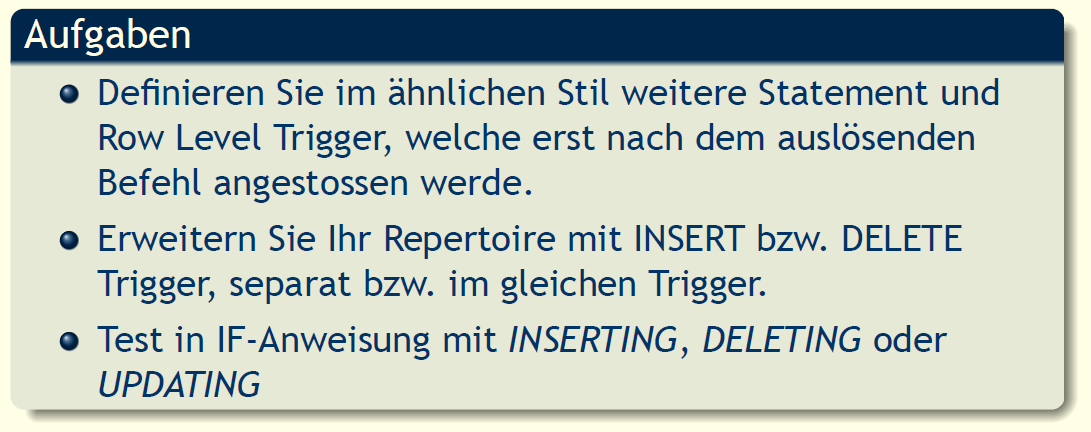
### 1. Erfahrung mit Trigger



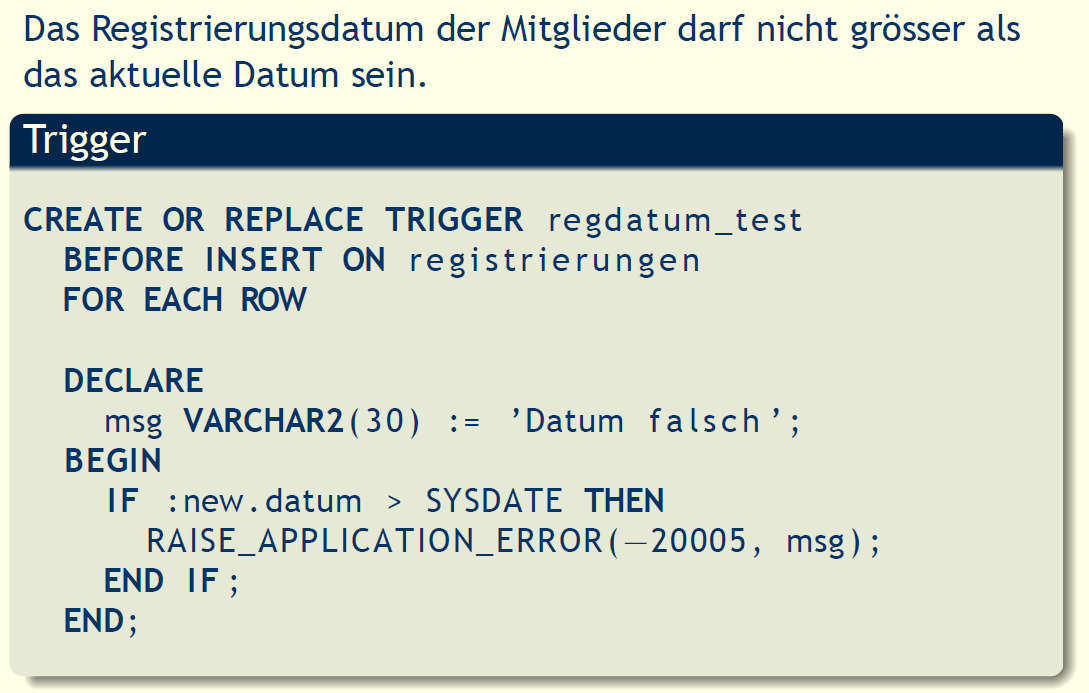


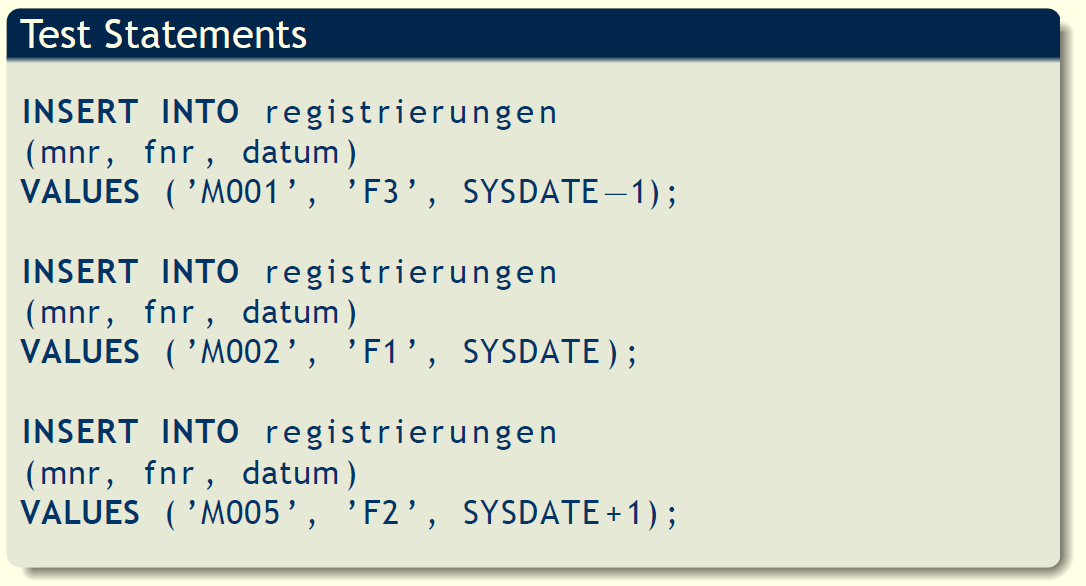


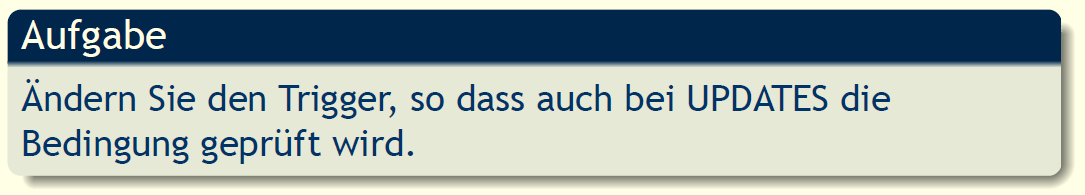




## Konsistent Prüfen

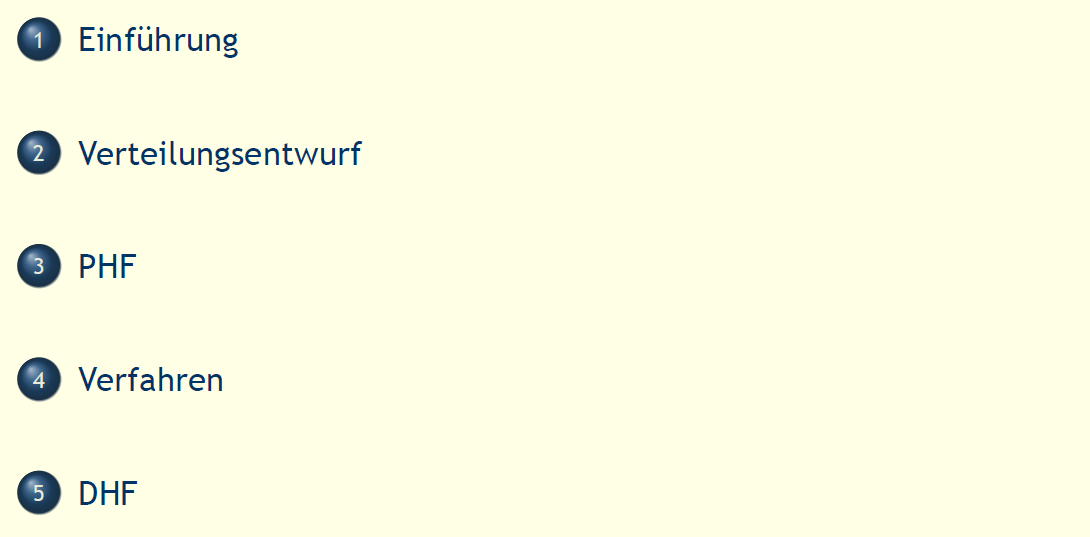






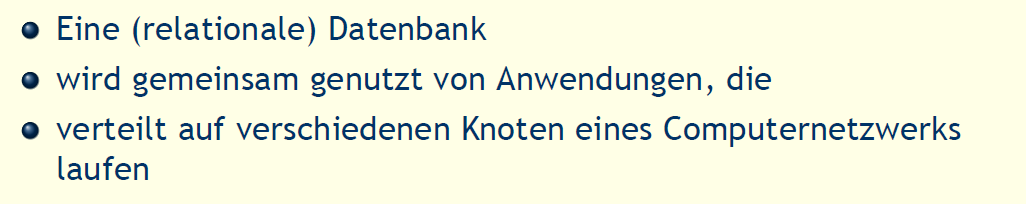
# Woche 2

## Inhalt

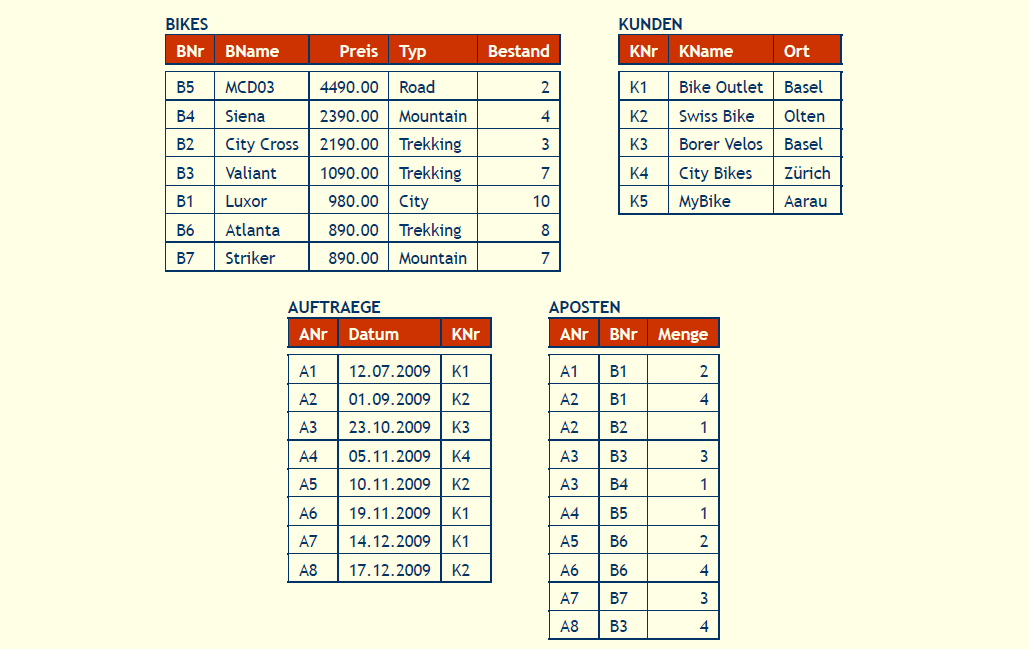


## Einführung

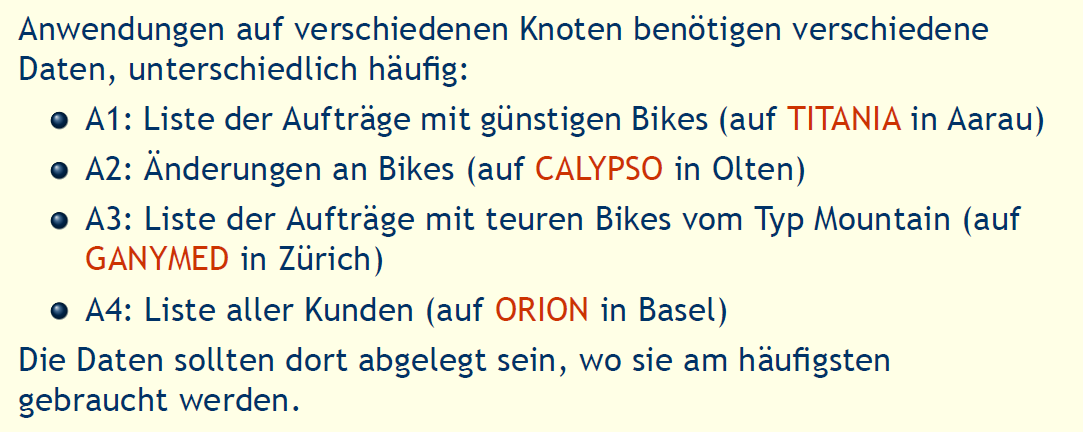
### Ausganglage



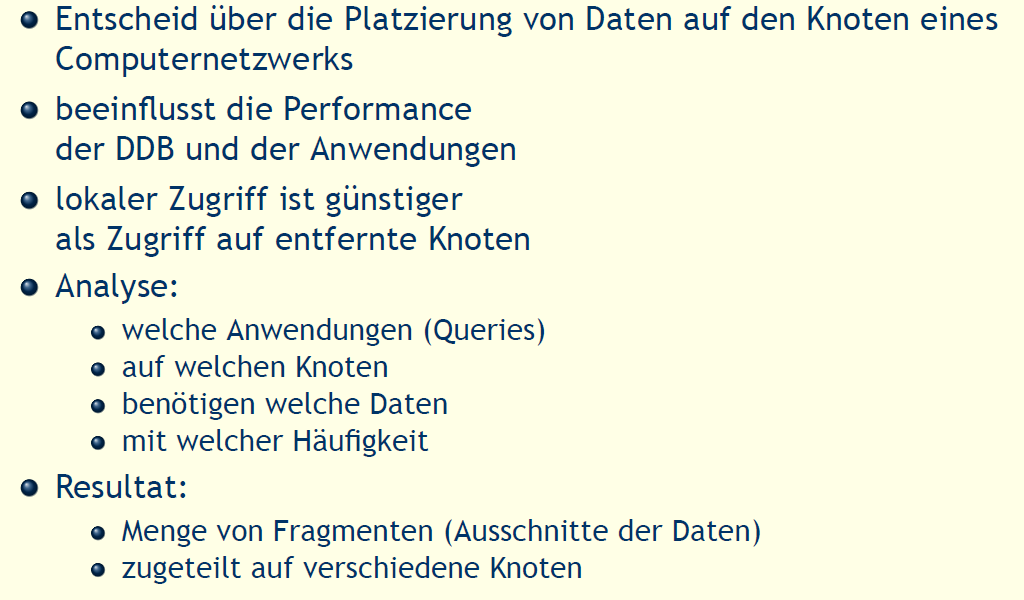
### Datenbank Beispiel



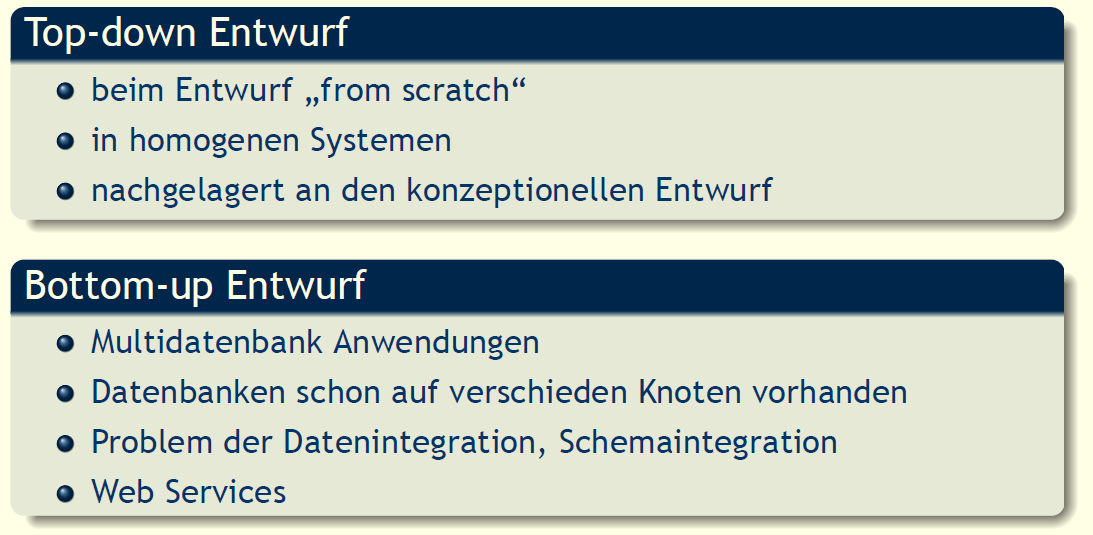
### Verteilte Anwendungen



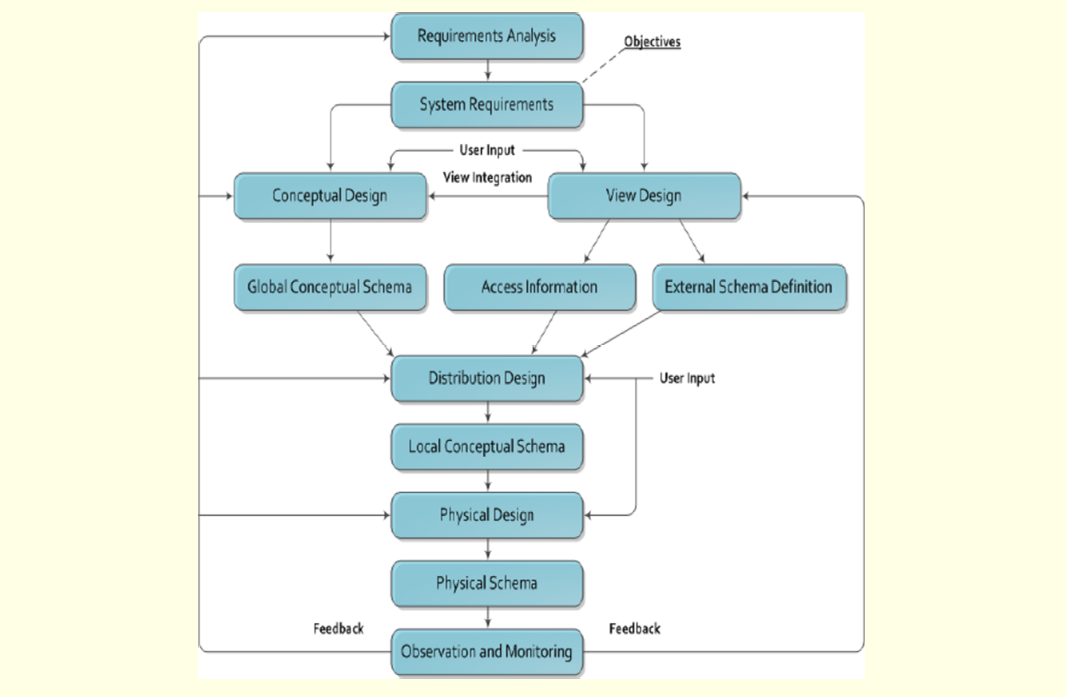
### Entwurfsproblem



### Alternative Entwurfsstrategien

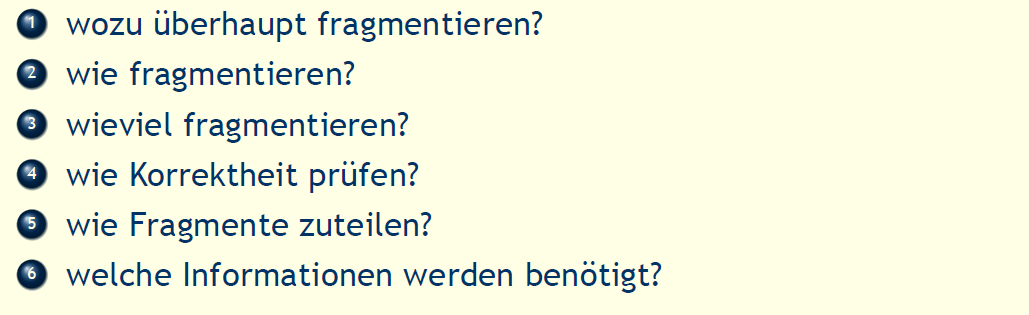


### Top-down Entwurf

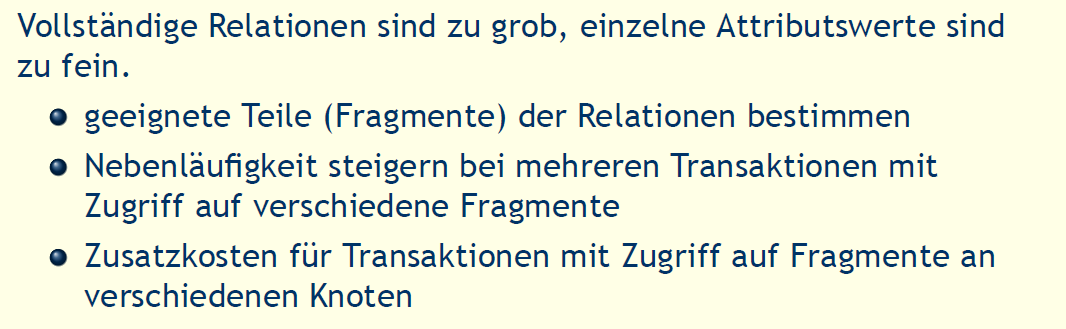


## Verteilungsentwurf

### Kernpunkte des Verteilungsentwurfs



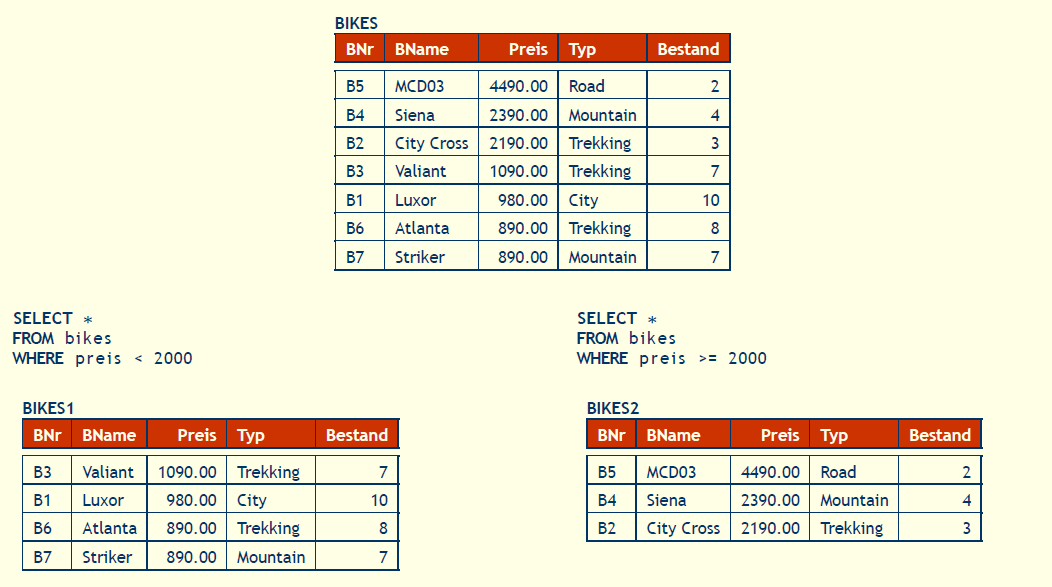
### Grad der Fragmentierung



### Arten der Fragmentierung



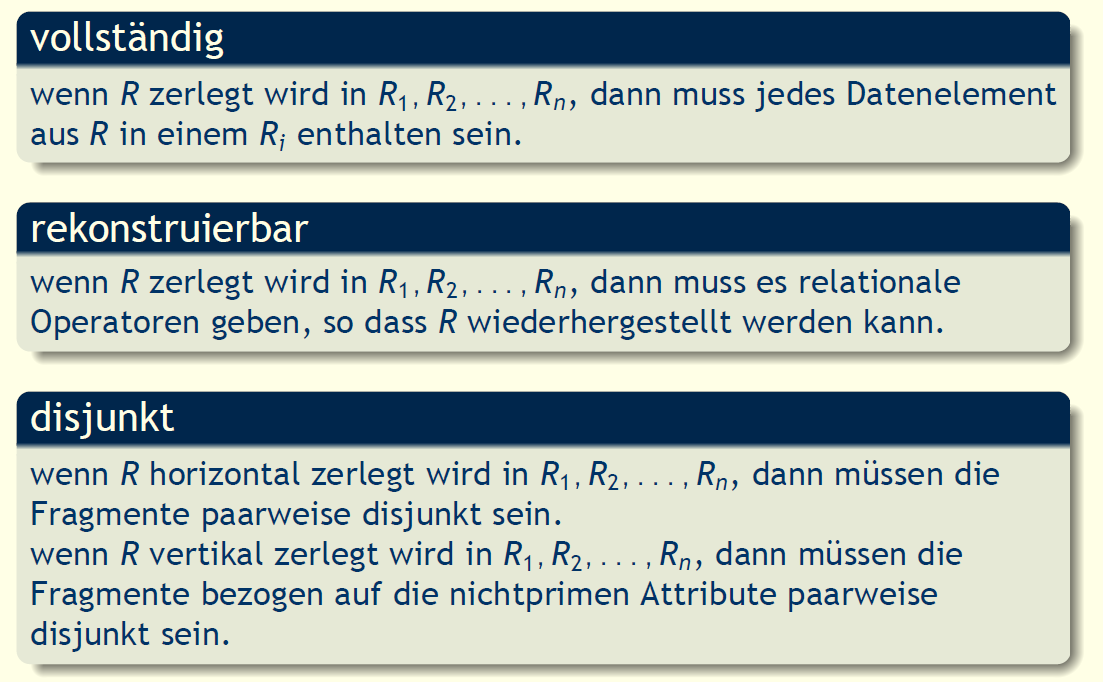
### Möglichkeit der Fragmentierung – Horizontal



### Möglichkeit der Fragmentierung – Vertikal

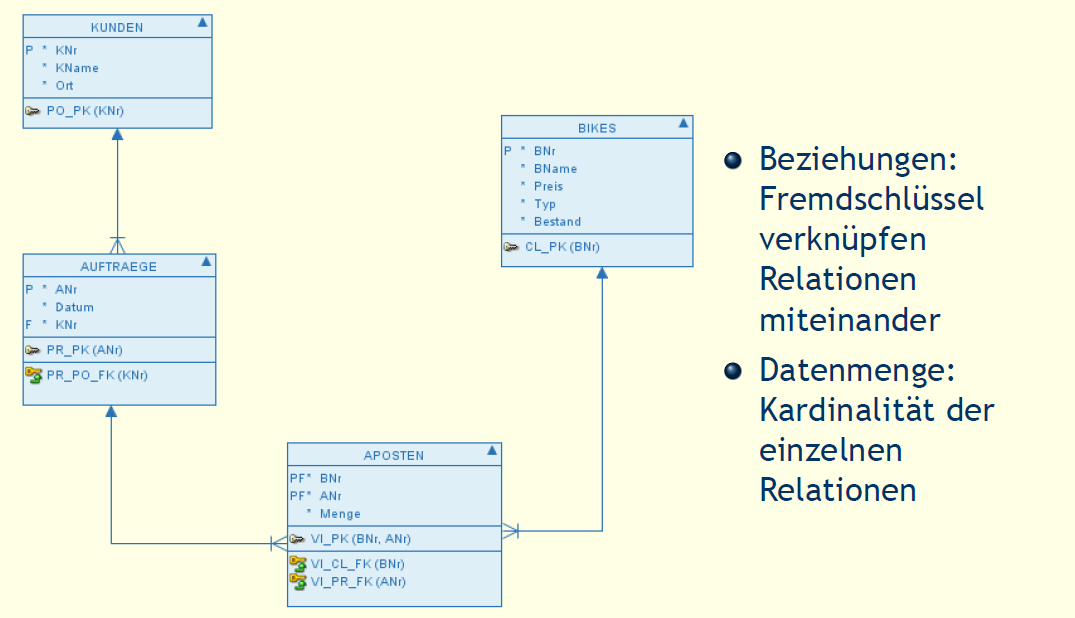


### Korrektheit der Fragmentierung



## PHF

### PHF – Information über Daten



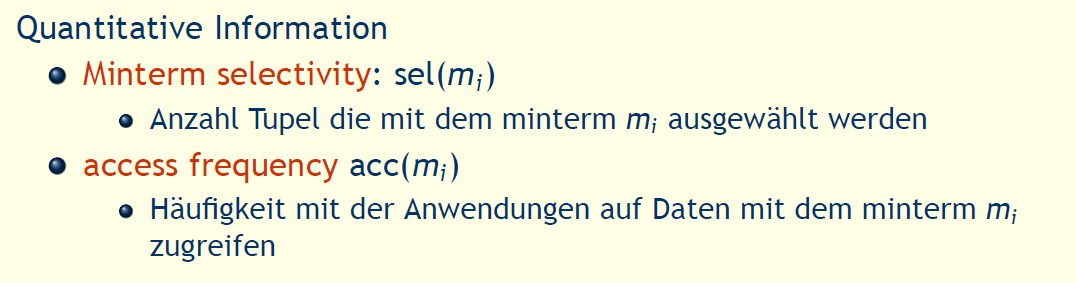
### PHF – Information über Anwendungen



### Minterm Beispiele

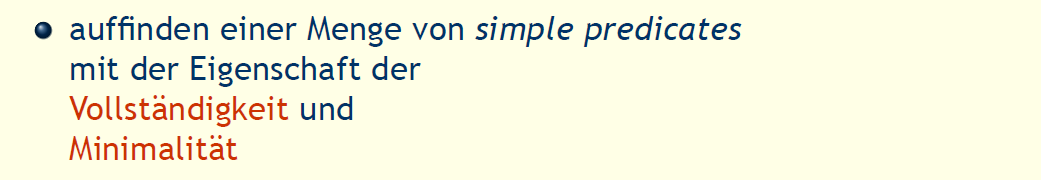


### PHF – Information über Anwendungen

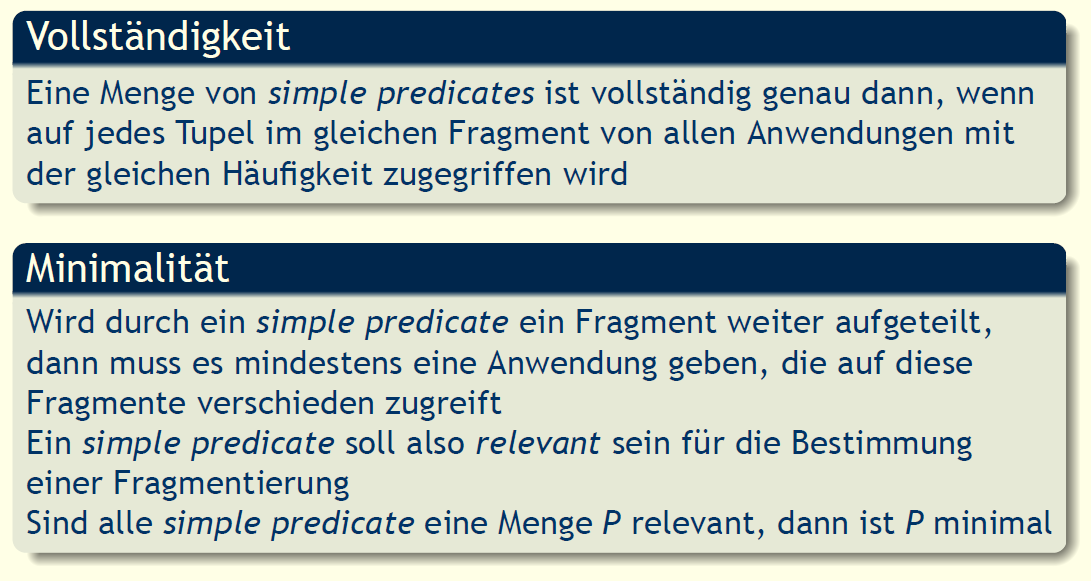


## Verfahren

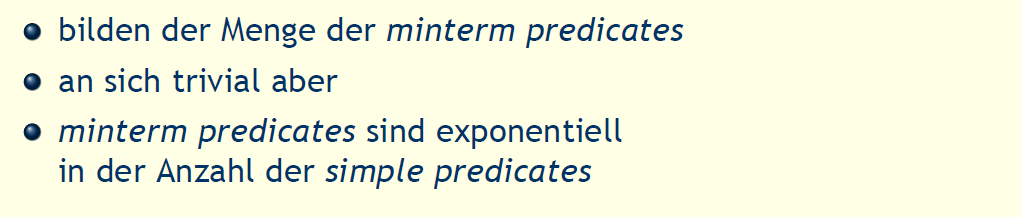
### PHF Verfahren, 1. Schritt



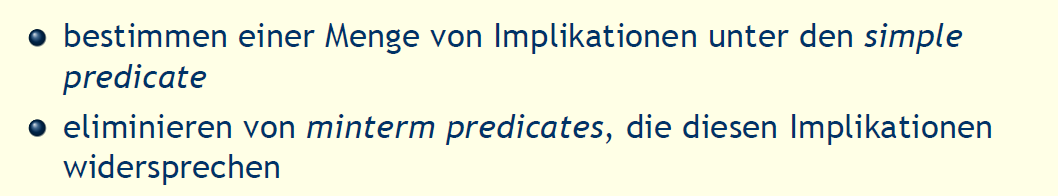
### Vollständigkeit und Minimalität



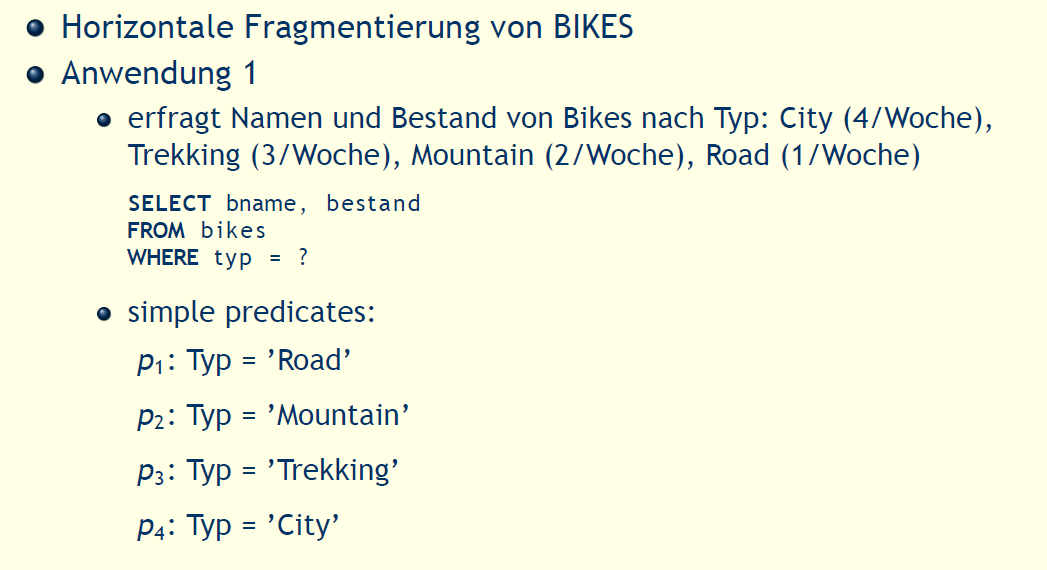
### PHF – Verfahren, 2. Schritt

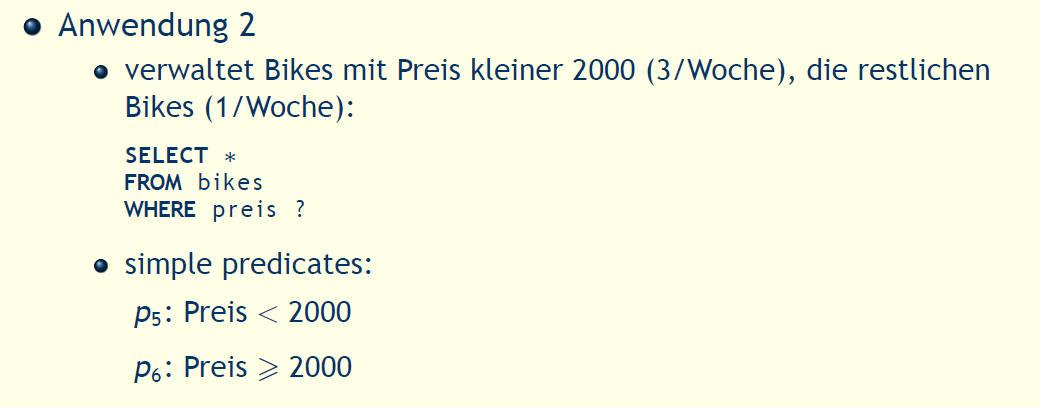


### PHF – Verfahren, 3. Schritt



### PHF – Beispiel



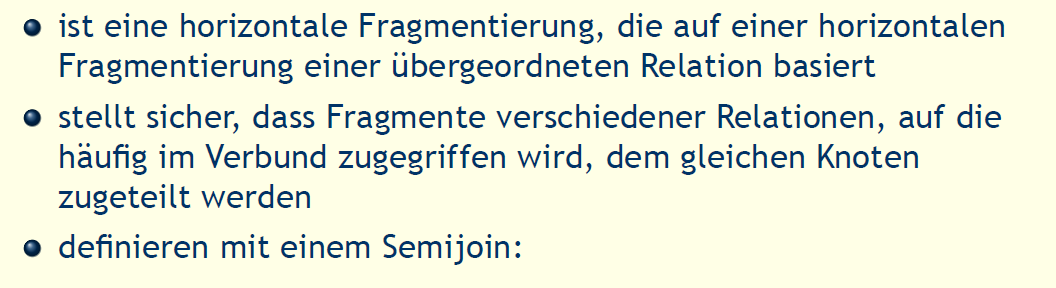




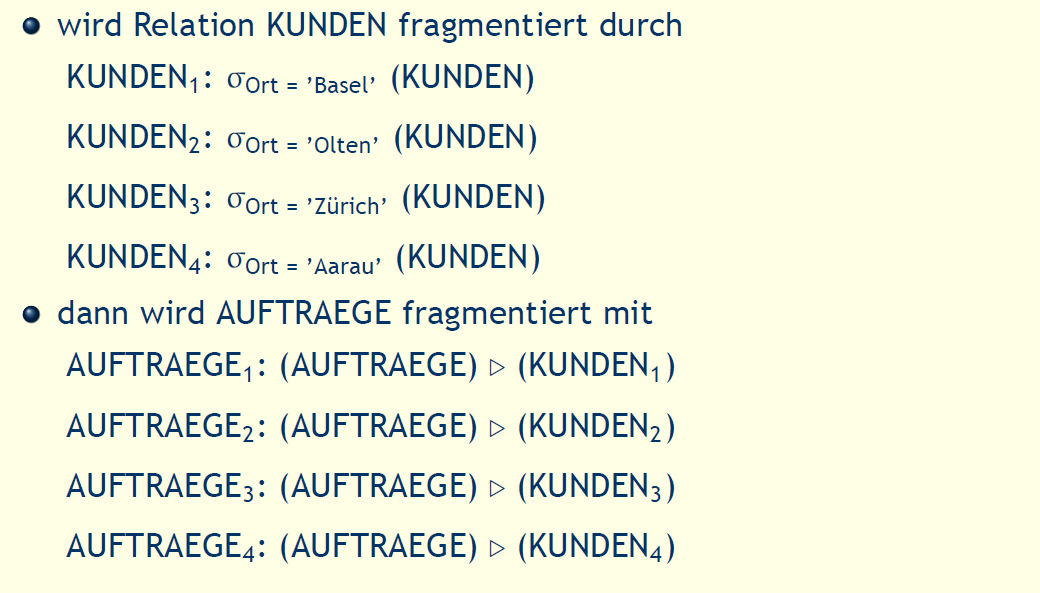


## DHF

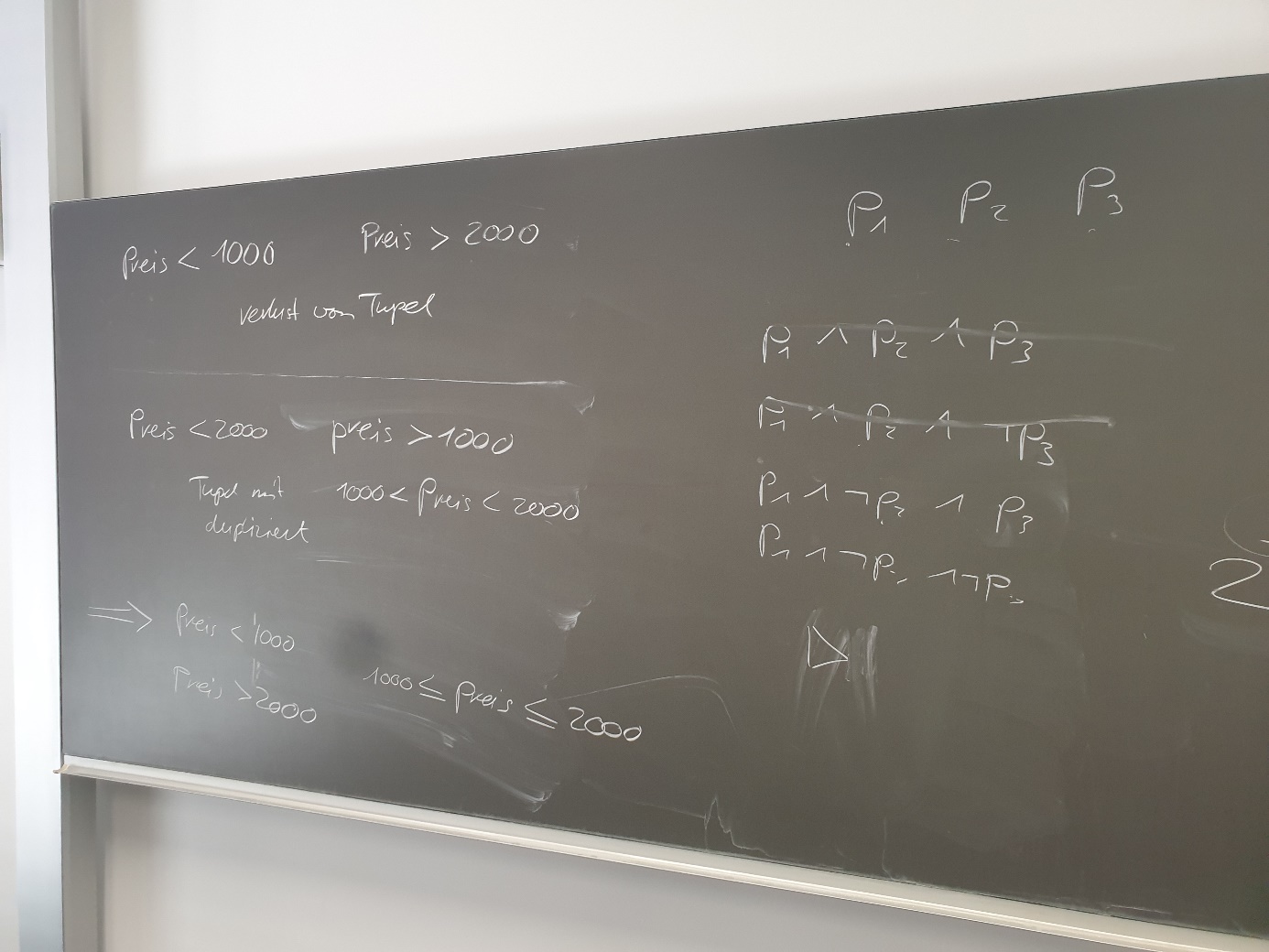
### Abgeleitete horizontale Fragmentierung

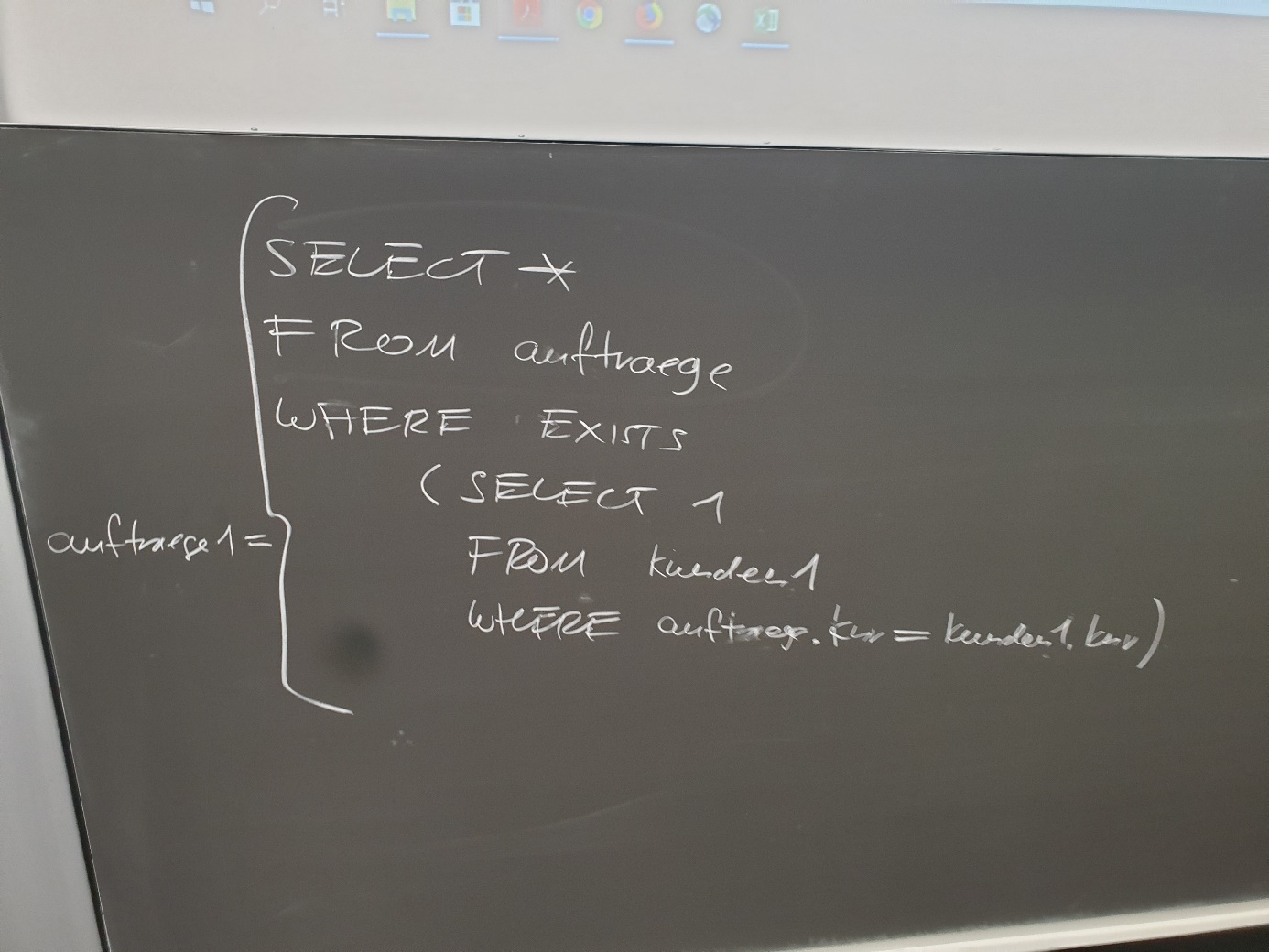


### DHF - Beispiel



## Tafelmitschrift





# Woche 3