

Informatik II

Vorlesung am D-BAUG der ETH Zürich

Dr. Felix Friedrich und Dr. Hermann Lehner

Vorlesung 12, 2018

Datenbanksysteme: Das Entity Relationship (ER) Modell, das Relationale Modell und SQL.

Demo: Java & SQL: Eine GUI Applikation fuer die VBZ Daten

Literatur

Literatur: Kemper, Eickler: Datenbanksysteme: Eine Einführung.
Oldenbourg Verlag, 9. Auflage, 2013.

Quellen: Basismaterial wurde von Prof. Donald Kossmann & Martin Kaufmann freundlicherweise zur Verfügung gestellt. Weitere Quelle: Folien zu *Datenbanksysteme: Eine Einführung*, Lehrstuhl III Datenbanksysteme, Prof. Kemper, TU München

Ziele

Nutzen von Datenbanksystemen verstehen, Modellierungskennntnisse

- ER Modell (Modellierung der Weltsicht)
- Relationales Modell (Modellierung für die DB)

Datenbanksystem anwenden

- Java + SQL

Datenbankverwaltungssysteme

- Ein Datenbankverwaltungssystem (DBMS) ist ein Werkzeug zur Erstellung und Ausführung datenintensiver Anwendungen
- grosse Datenmengen
- grosse Datenströme



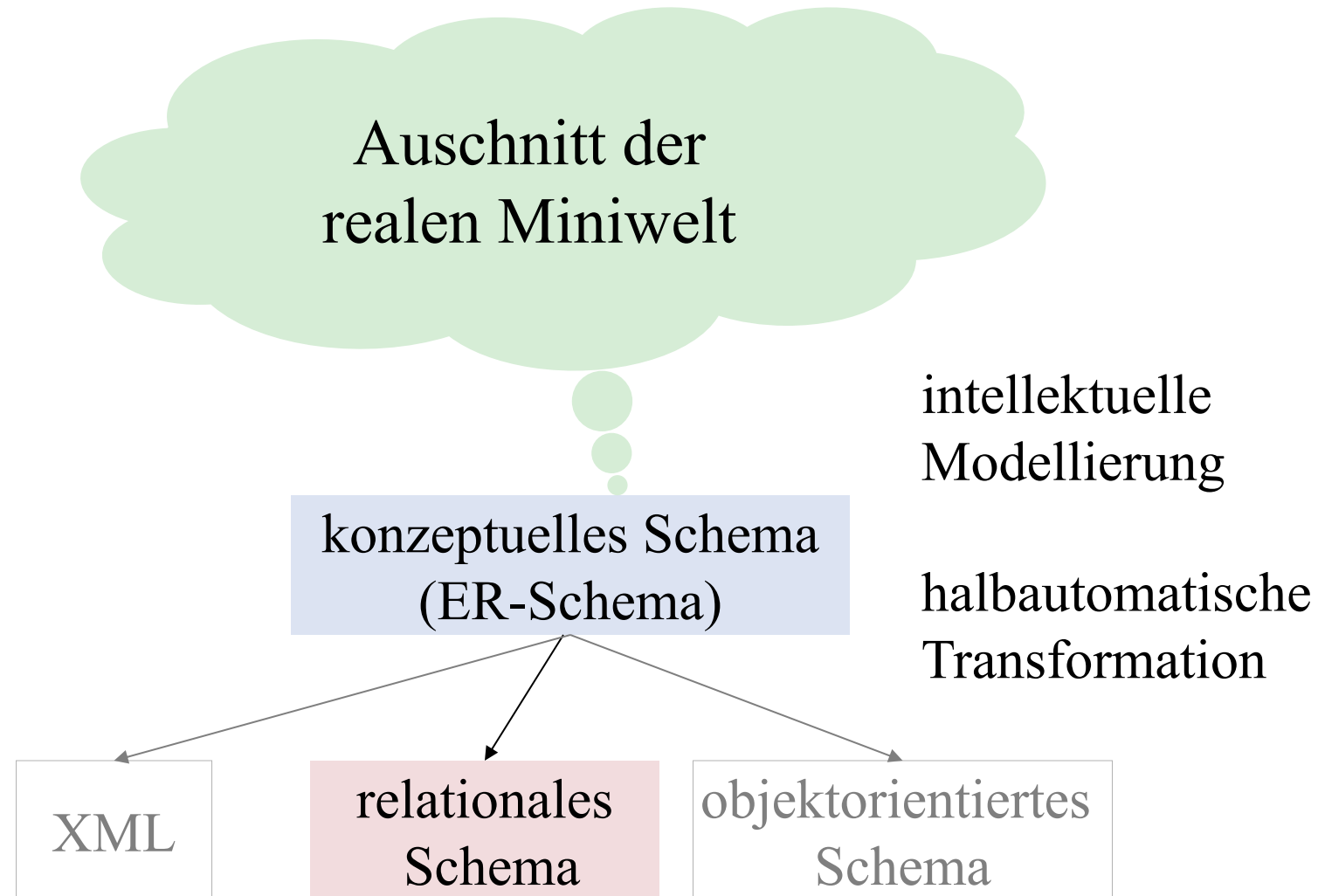
Typische Anwendungen

- **Bank**
z.B. Konten / „Geldtransfer“
- **Bibliothek**
z.B. Bücher / „Ausleihen“
- **Facebook, Twitter, ...**
z.B. Freunde, „Sende Tweet“
- **Geoinformationssysteme**
z.B. Topographische Information, "Erzeuge Karte"
- "Alles, was Sie mit einer Tabelle machen wollen, mit Excel nicht erledigen können und nicht selbst programmieren möchten".

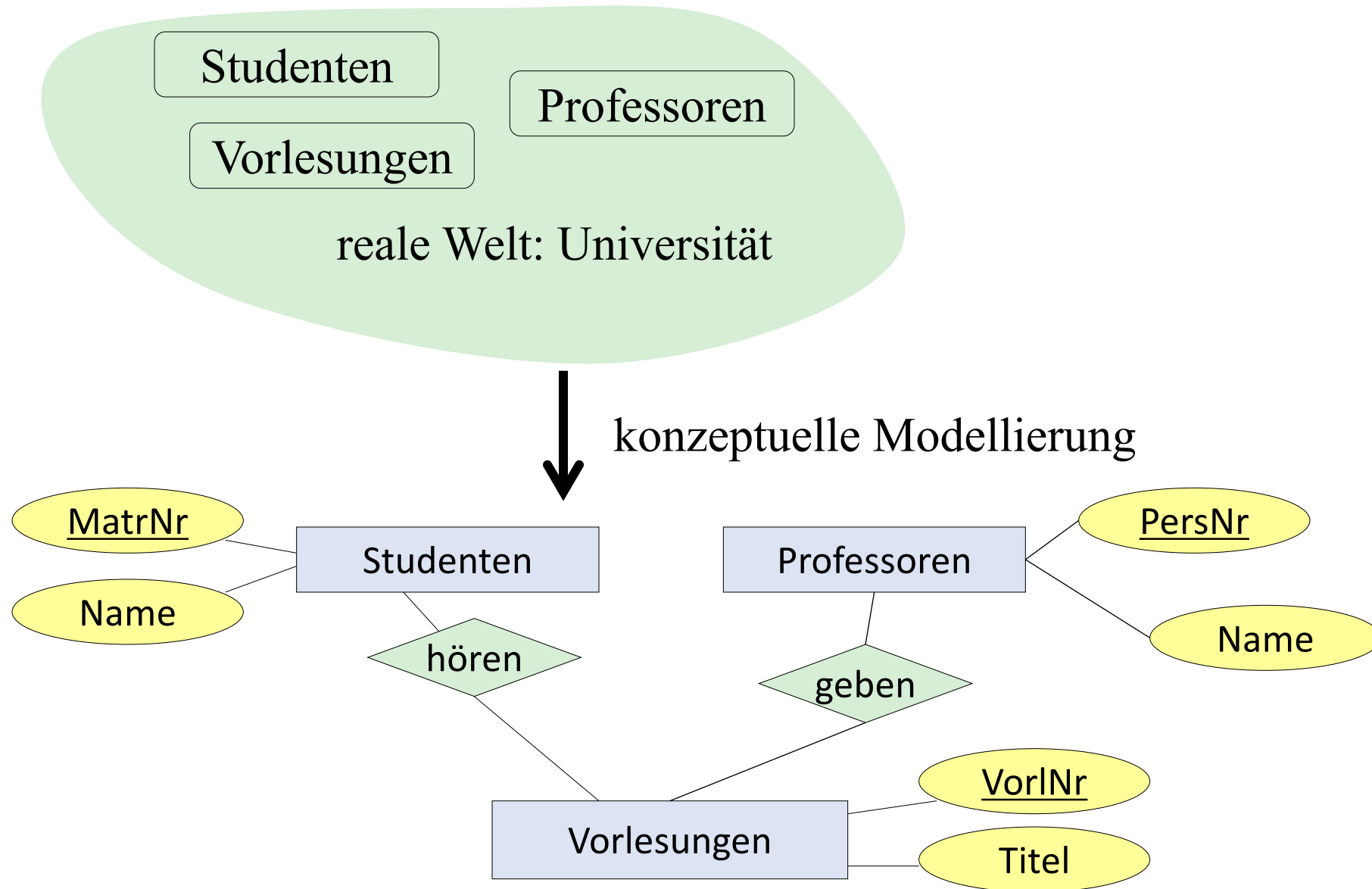
Wozu Datenbanksysteme?

- Vermeide Redundanz und Inkonsistenz
- Deklarativer Zugriff auf die Daten und Unabhängigkeit von der Implementation (physische Datenunabhängigkeit)
- Synchronisiere gleichzeitigen Datenzugriff
- Sicherheit, Vertraulichkeit
- **Minimiere Kosten und Aufwand**
Ähnliche Funktionalität selbst zu implementieren würde Jahre in Anspruch nehmen

Datenmodellierung



1. Teil heutige Vorlesung: Modellierung (ER-Modell)



2. Teil heutige Vorlesung: Relationales Datenmodell

Studenten	
Legi	Name
26120	Fichte
25403	Jonas
...	...

hören	
Legi	VorlNr
25403	5022
26120	5001
...	...

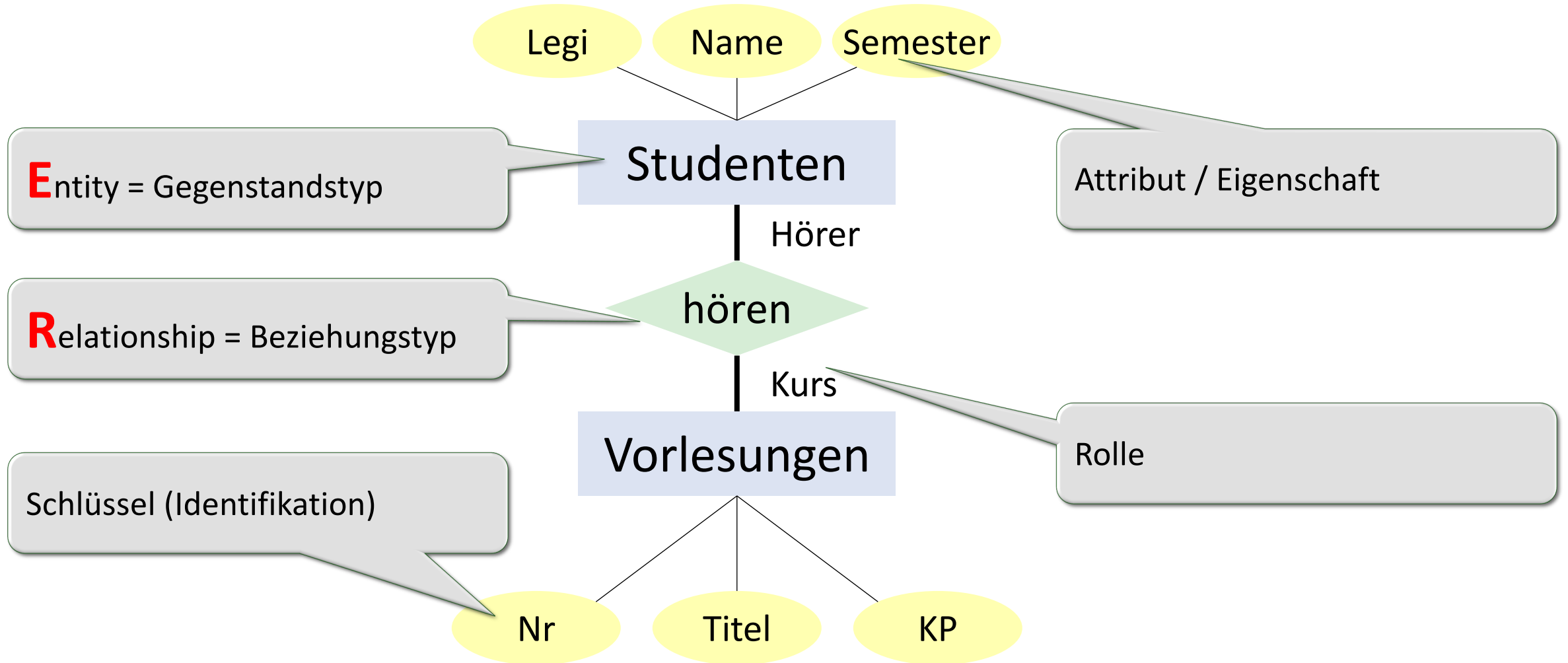
Vorlesungen	
VorlNr	Titel
5001	Grundzüge
5022	Glaube und Wissen
...	...

```
select Name
from Studenten, hören, Vorlesungen
where Studenten.Legi= hören.Legi and
        hören.VorlNr= Vorlesungen.VorlNr and
        Vorlesungen.Titel = `Grundzüge`;
```

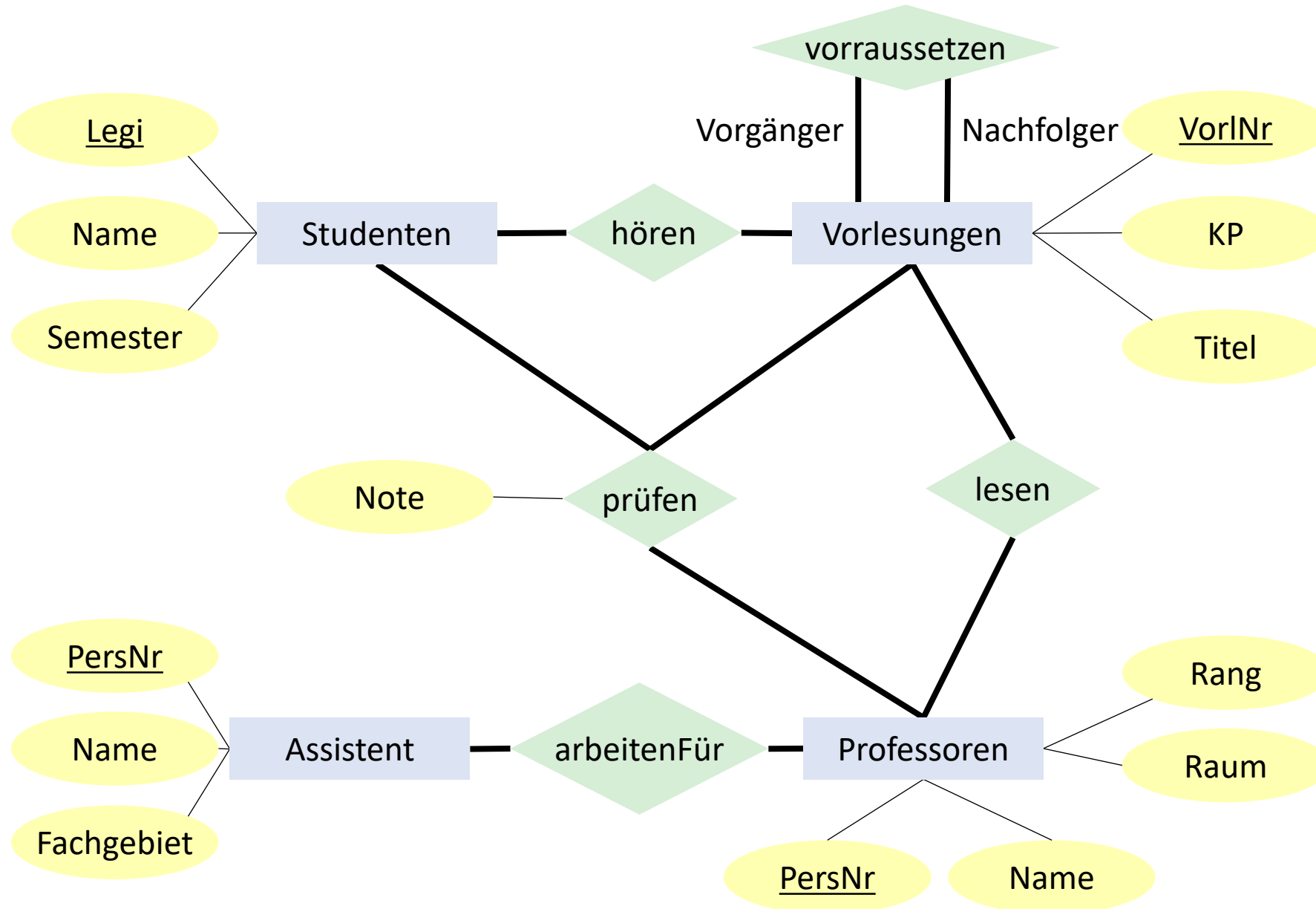
```
update Vorlesungen
set     Title = `Grundzüge der Logik`
where   VorlNr = 5001;
```

Entity Relationship Modell

Entity/Relationship (ER) Modell



Modell einer Universität



... in natürlicher Sprache

- Studenten haben LegiNr, Name und Semester. Die LegiNr identifiziert einen Studenten eindeutig.
- Vorlesungen haben eine VorlNr, Kreditpunkte und einen Titel. VorlNr identifiziert eine Vorlesung eindeutig.
- Professoren haben PersNr, Name, Rang und Raum. PersNr identifiziert einen Professor eindeutig.
- Assistenten haben PersNr, Name und Fachgebiet. PersNr. identifiziert einen Assistenten eindeutig.
- Studenten hören Vorlesungen
- Vorlesungen können Voraussetzung für andere Vorlesungen sein.
- Professoren lesen Vorlesungen.
- Assistenten arbeiten für Professoren
- Studenten werden von Professoren über Vorlesungen geprüft. Studenten erhalten Noten als Teil dieser Prüfungen.
- **Ist das die einzig mögliche Interpretation?**
- **Nein: zu einem ER-Model gehört immer noch Dokumentation und/oder gesunder Menschenverstand**

Warum ER?

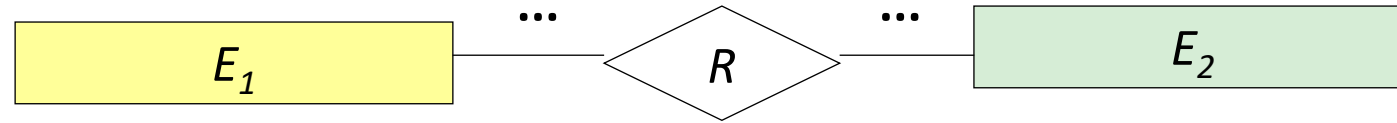
■ Vorteile

- ER Diagramme sind einfach zu erstellen und editieren
- ER Diagramme sind aufgrund der grafischen Darstellung einfach zu verstehen (vom Laien)
- ER Diagramme beschreiben alle Informationsanforderungen

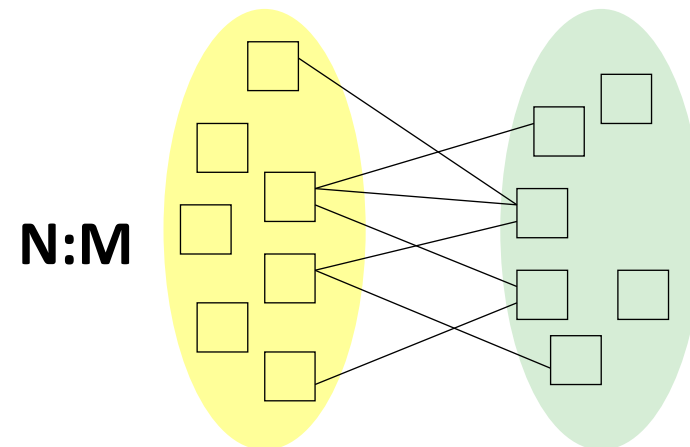
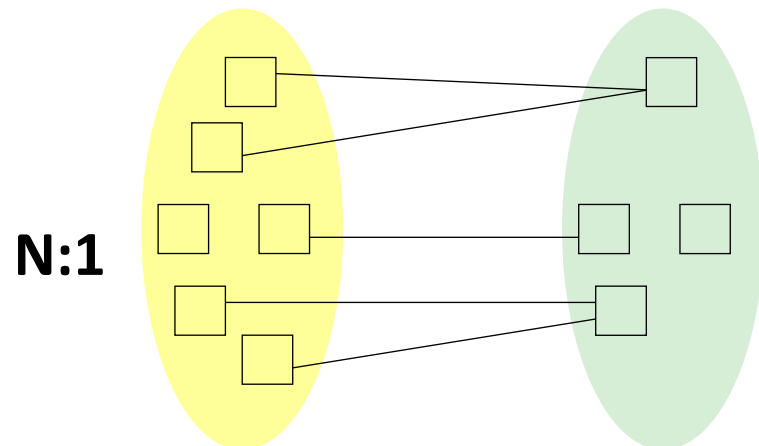
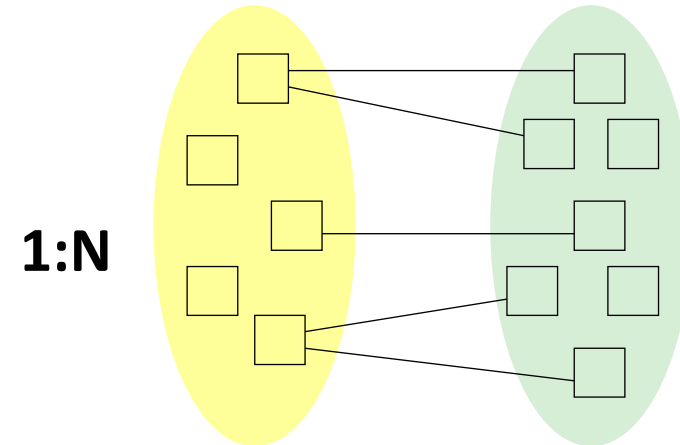
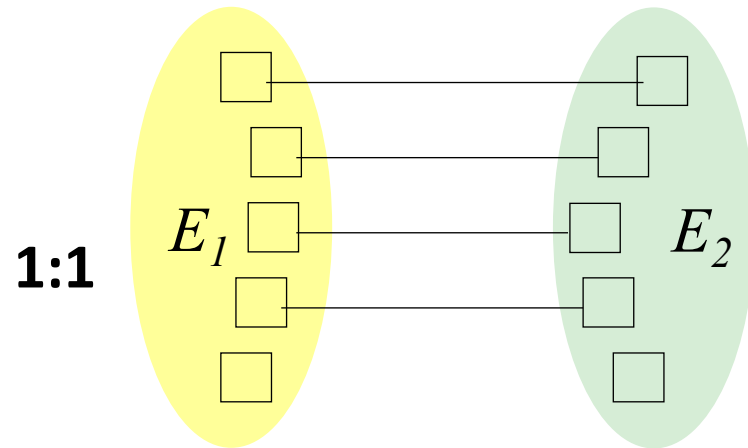
■ Allgemeines

- Viele Tools verfügbar
- Kontroverse, ob ER/UML in der Praxis von Nutzen ist
- Keine Kontroverse, dass jeder ER/UML lernen sollte

Funktionalitätsangaben

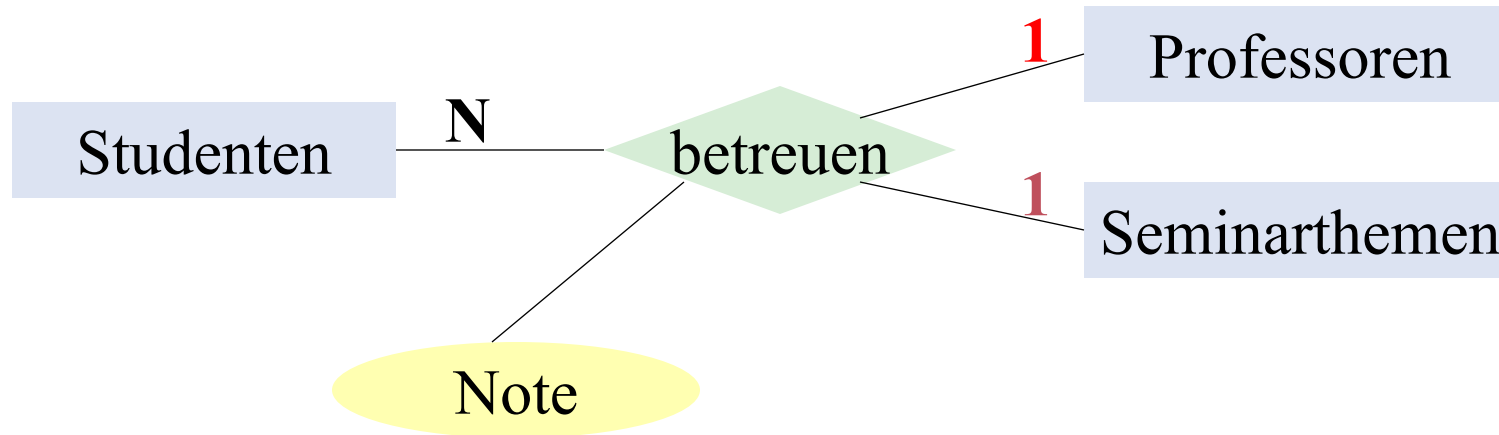


$$R \subseteq E_1 \times E_2$$



Beispiel: Seminar

Um das zu verstehen, denken Sie am besten in Abbildungen (Funktionen): eine Abbildung $f(x,y)$ hat pro Parameter x,y nur einen Wert.



betreuen: Professoren x Studenten \rightarrow Seminarthemen

betreuen: Seminarthemen x Studenten \rightarrow Professoren

Konsistenzbedingungen des Seminar

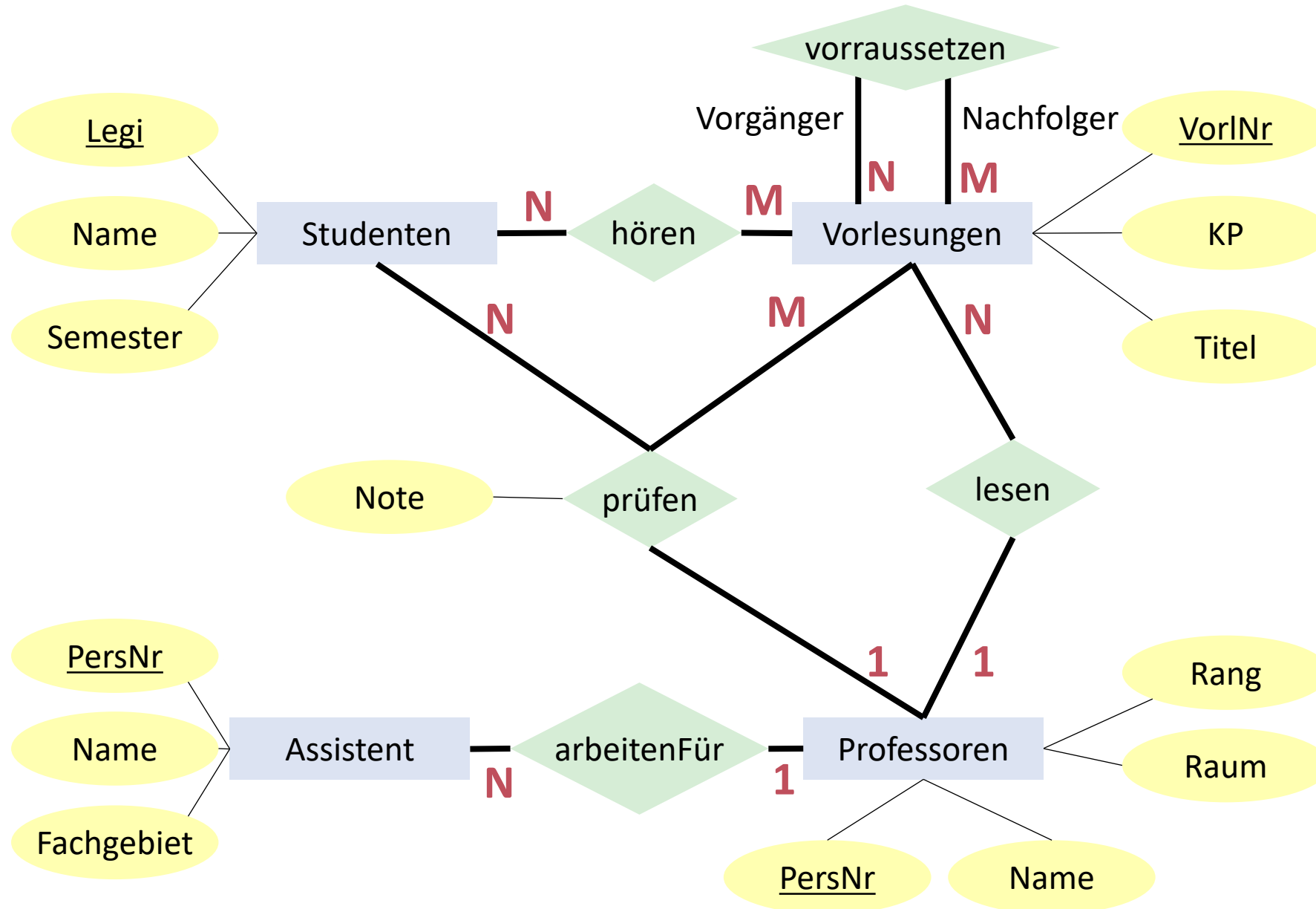
Einschränkungen

1. Studenten dürfen bei einem Professor nur ein Seminarthema bearbeiten
2. Studenten können dasselbe Seminarthema nur einmal bearbeiten

Möglichkeiten

1. Professoren können das Seminarthema für andere Studenten wiederverwenden
2. Dasselbe Thema kann von verschiedenen Professoren verwendet werden

Universität mit Funktionalitäten



Daumenregeln

Wann Attribut, wann Entität?

- Entität, wenn das Konzept mehr als eine Beziehung hat
- Attribut, wenn das Konzept nur eine 1:1 Beziehung hat

Partitionierung von ER-Modellen

- Realistische Modelle sind grösser als eine Seite
- Nach Bereichen / Organisationseinheiten partitionieren
- Kein gutes automatisches Graphenpartitionierungstool bekannt

Tipps

- Keine Redundanz modellieren
- Je weniger Entitäten desto besser

ER Modellierung: Zusammenfassung

ER beschreibt eine Miniwelt

- Das "was" und die Regeln
- ER ist statisch. Es beschreibt keine Übergänge

Nützlich zum Erstellen von Software zur Beantwortung von (An)fragen über die Miniwelt

- es folgt nun: ER-Modell → relationales Modell

Ähnliche Modellierungsmöglichkeiten bietet **UML** (mehr auf OOP zugeschnitten)

Auch andere graphische Darstellungen des ER Modells gebräuchlich, z.B.

"Krähentrittnotation" optisch näher bei UML

Relationales Modell

Die Welt in Tabellen

Relationales Modell, Formalismus

- **Relation** R

- $R \subseteq D_1 \times \dots \times D_n$
- D_1, D_2, \dots, D_n sind Domänen

Beispiel: Telefonbuch \subseteq string \times string \times integer

- **Tupel:** $t \in R$

Beispiel: $t = (\text{„Mickey Mouse“}, \text{„Main Street“}, 4711)$

- **Relationenschemata** werden wie folgt beschrieben

Telefonbuch: {[Name: string, Strasse: string, Telefon#:integer]}

{[...] } deuten an, dass ein Schema eine Menge von Tupeln [] ist

Name des Attributes

Typ des Attributes

Relationales Modell

Telefonbuch		
Name	Strasse	<u>Telefon#</u>
Mickey Mouse	Main Street	4711
Minnie Mouse	Broadway	94725
Donald Duck	Broadway	95672
...

Ausprägung: Zustand der Datenbank

Schlüssel: minimale Menge von Attributen, welche ein Tupel eindeutig identifizieren
z.B. {Telefon#} oder {Name, Geburtstag}

Primärschlüssel (durch Unterstreichung hervorgehoben): Ausgewählter Schlüssel, welcher üblicherweise zur Identifikation eines Tupels in einer Relation verwendet wird.

Regel #1: Darstellung von Entities

Studenten:

{[Legi:integer], *Name:string*, *Semester: integer*}

Vorlesungen:

{[VorlNr:integer], *Titel: string*, *KP: integer*}

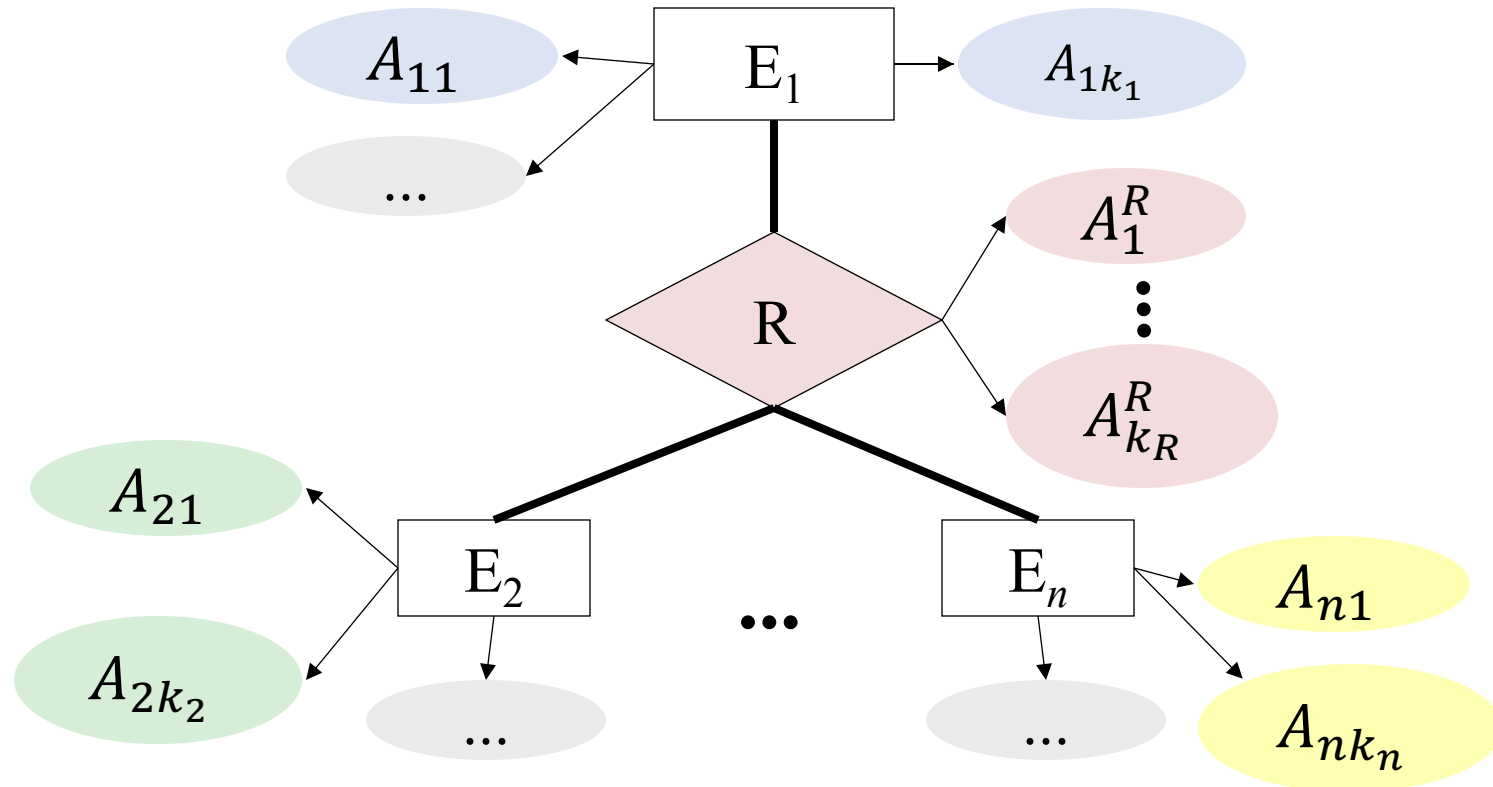
Professoren:

{[PersNr:integer], *Name: string*, *Rang: string*, *Raum: integer*}

Assistenten:

{[PersNr:integer], *Name: string*, *Fachgebiet: string*}

Regel #2: Darstellung von Beziehungen



$$R: \left\{ \underbrace{[A_{11}, \dots, A_{1k_1}]}_{\text{Schlüssel } E_1}, \underbrace{[A_{21}, \dots, A_{2k_2}]}_{\text{Schlüssel } E_2}, \dots, \underbrace{[A_{n1}, \dots, A_{nk_n}]}_{\text{Schlüssel } E_n}, \underbrace{[A_1^R, \dots, A_{k_R}^R]}_{\text{Attribute von } R} \right\}$$

Darstellung von Beziehungen

hören:

{[Legi: integer, VorlNr: integer]}

lesen :

{[PersNr: integer, VorlNr: integer]}

Fremdschlüssel, identifizieren
Tupel aus anderen Entitäten

arbeitenFür :

{[AssiPersNr: integer, ProfPersNr: integer]}

voraussetzen:

{[Vorgänger: integer, Nachfolger: integer]}

prüfen :

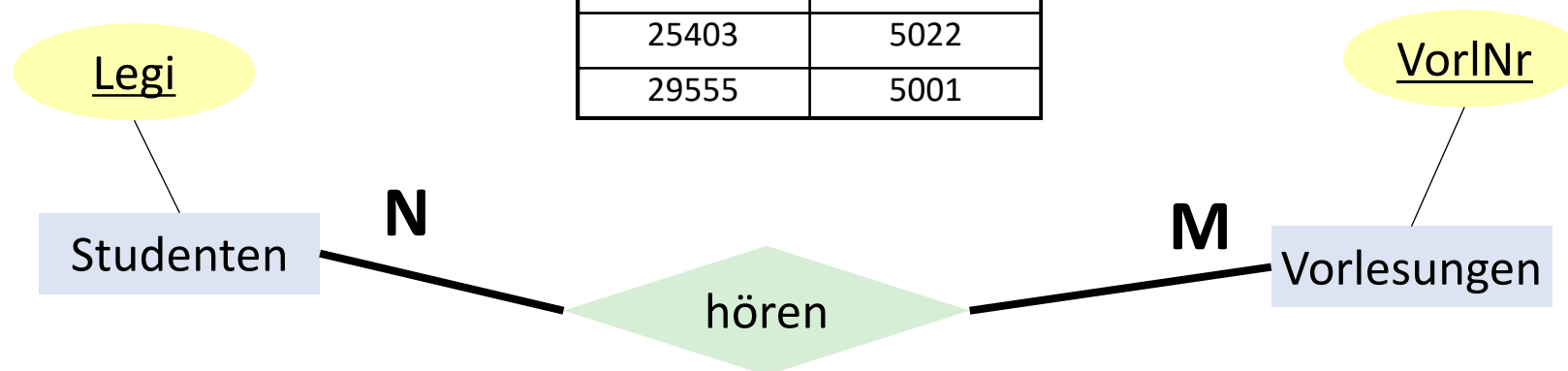
{[Legi: integer, VorlNr: integer, PersNr: integer, Note: decimal]}

Ausprägung von *hören*

Studenten	
<i>Legi</i>	...
26120	...
27550	...
...	...

hören	
<i>Legi</i>	<i>VorlNr</i>
26120	5001
27550	5001
27550	4052
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5022
25403	5022
29555	5001

Vorlesungen	
<i>VorlNr</i>	...
5001	...
4052	...
...	...

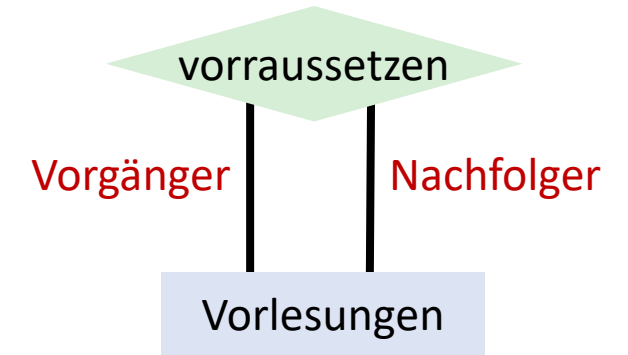


Zur Regel #2: Benennung der Attribute?

Spezifiziert das ER-Modell **Rollen**, dann

- nimm den Namen der jeweiligen Rolle

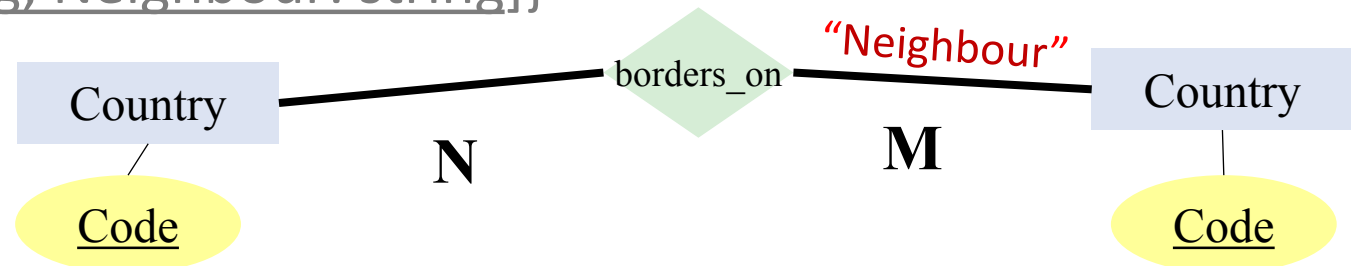
`voraussetzen: {[Vorgänger:string, Nachfolger:string]}`



andernfalls

- benutze die Namen der Schlüsselattribute der Entitäten
- bei Mehrdeutigkeit erfinde aussagekräftigen Namen

`borders_on: {[Code: string, Neighbour: string]}`



Regel #3: Zusammenfassung von Relationen

Vorlesungen : {[VorlNr, Title, CP]}
Professoren: {[PersNr, Name, Level, Room]}
lesen: {[VorlNr, PersNr]}

Fasse (nur) Relationen mit **gleichem Schlüssel** zusammen
(also auch nur (N:1), (1:N) oder (1:1) Beziehungen)

Zusammenfassen:

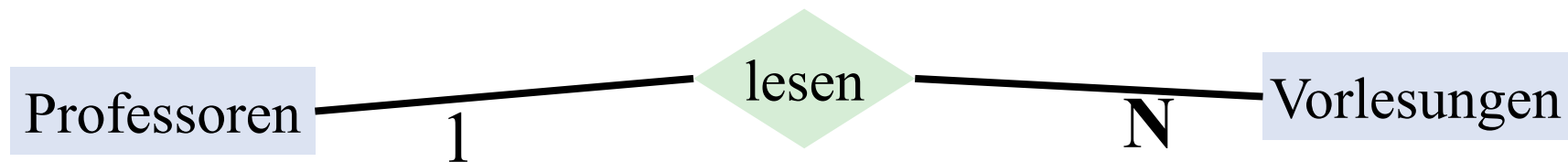
Vorlesungen : {[VorlNr, Title, CP, **gelesenVon**]}
Professoren : {[PersNr, Name, Level, Room]}



Ausprägung von *Professoren* und Vorlesungen

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	FP	226
2126	Russel	FP	232
2127	Kopernikus	AP	310
2133	Popper	AP	52
2134	Augustinus	AP	309
2136	Curie	FP	36
2137	Kant	FP	7

Vorlesungen			
VorlNr	Titel	KP	gelesenVon
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
4630	Die 3 Kritiken	4	2137



Das funktioniert NICHT

Professoren				
PersNr	Name	Rang	Raum	liest
2125	Sokrates	FP	226	5041
2125	Sokrates	FP	226	5049
2125	Sokrates	FP	226	4052
...
2134	Augustinus	AP	309	5022
2136				

Vorlesungen		
VorlNr	Titel	KP
5001	Grundzüge	4
5041	Ethik	4
5043	Erkenntnistheorie	3
5049	Mäeutik	2
4052	Logik	4
		3
		2
		2
		2
		4

Problem: Redundanz und Anomalien
 PersNr ist kein gültiger Schlüssel für
 Professoren mehr



Relationales Modell der Uni-DB

Professoren			
PersNr	Name	Rang	Raum
2125	Sokrates	FP	226
2126	Russel	FP	232
2127	Kopernikus	AP	310
2133	Popper	AP	52
2134	Augustinus	AP	309
2136	Curie	FP	36
2137	Kant	FP	7

Studenten		
Legi	Name	Semester
24002	Xenokrates	18
25403	Jonas	12
26120	Fichte	10
26830	Aristoxenos	8
27550	Schopenhauer	6
28106	Carnap	3
29120	Theophrastos	2
29555	Feuerbach	2

Vorlesungen			
VorlNr	Titel	KP	gelesenVon
5001	Grundzüge	4	2137
5041	Ethik	4	2125
5043	Erkenntnistheorie	3	2126
5049	Mäeutik	2	2125
4052	Logik	4	2125
5052	Wissenschaftstheorie	3	2126
5216	Bioethik	2	2126
5259	Der Wiener Kreis	2	2133
5022	Glaube und Wissen	2	2134
	Die 3 Kritiken	4	2137

hören	
Legi	VorlNr
26120	5001
27550	5001
27550	4052
28106	5041
28106	5052
28106	5216
28106	5259
29120	5001
29120	5041
29120	5049
29555	5022

Assistenten			
PerslNr	Name	Fachgebiet	Boss
3002	Platon	Ideenlehre	2125
3003	Aristoteles	Syllogistik	2125
3004	Wittgenstein	Sprachtheorie	2126
3005	Rhetikus	Planetenbewegung	2127
3006	Newton	Keplersche Gesetze	2127
3007	Spinoza	Gott und Natur	2126

voraussetzen	
Vorgänger	Nachfolger
5001	5041
5001	5043
5001	5049
5041	5216
5043	5052
5041	5052
5052	5259

prüfen			
Legi	Nr	PersNr	Note
28106	5001	2126	1
25403	5041	2125	2
27550	4630	2137	2

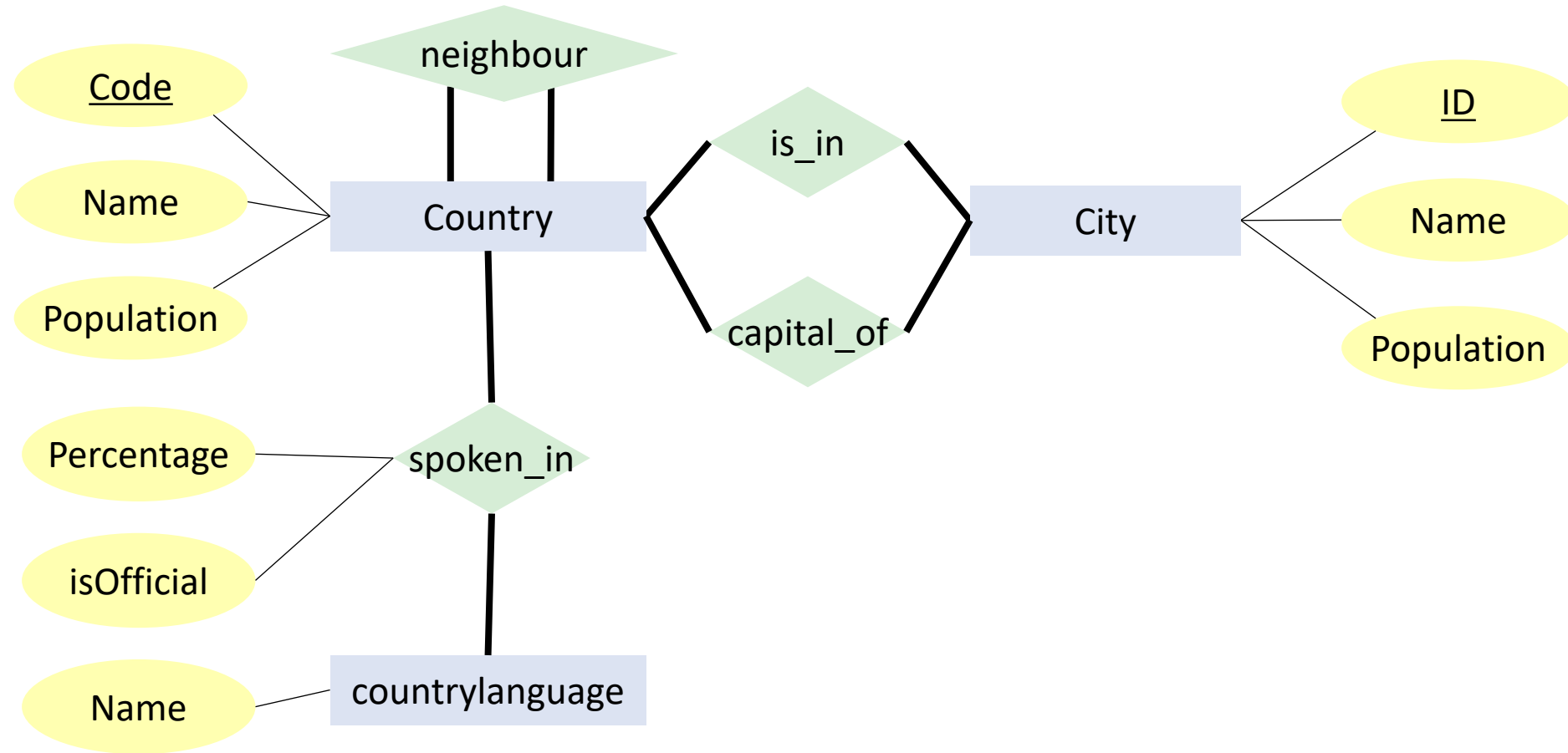


```
public static void loadDatabase(File file, String query) {  
    if (file == null) return;  
    records.clear();  
    maxDelay = 0;  
    Connection connection = null;  
    try {  
        String url = "jdbc:sqlite:" + file.getAbsolutePath();  
        connection = DriverManager.getConnection(url);  
        ResultSet rs = connection.createStatement().executeQuery(query);  
        while (rs.next()) {  
            Record r = new Record(  
                rs.getString("Name"),  
                rs.getInt("Delay"),  
                rs.getDouble("Latitude"),  
                rs.getDouble("Longitude")  
            );  
            maxDelay = Math.max(maxDelay, r.getDelay());  
            records.add(r);  
        }  
    } catch (SQLException e) {  
        Alerts.showExceptionInfo("Ein Problem ist aufgetreten bei Abfragen der Datenbank", e);  
    } finally {  
        try {  
            if (connection != null) {  
                connection.close();  
            }  
        } catch (SQLException e) {  
            Alerts.showExceptionInfo("Ein Problem ist aufgetreten beim Schliessen der Datenbankverbindung", e);  
        }  
    }  
}
```

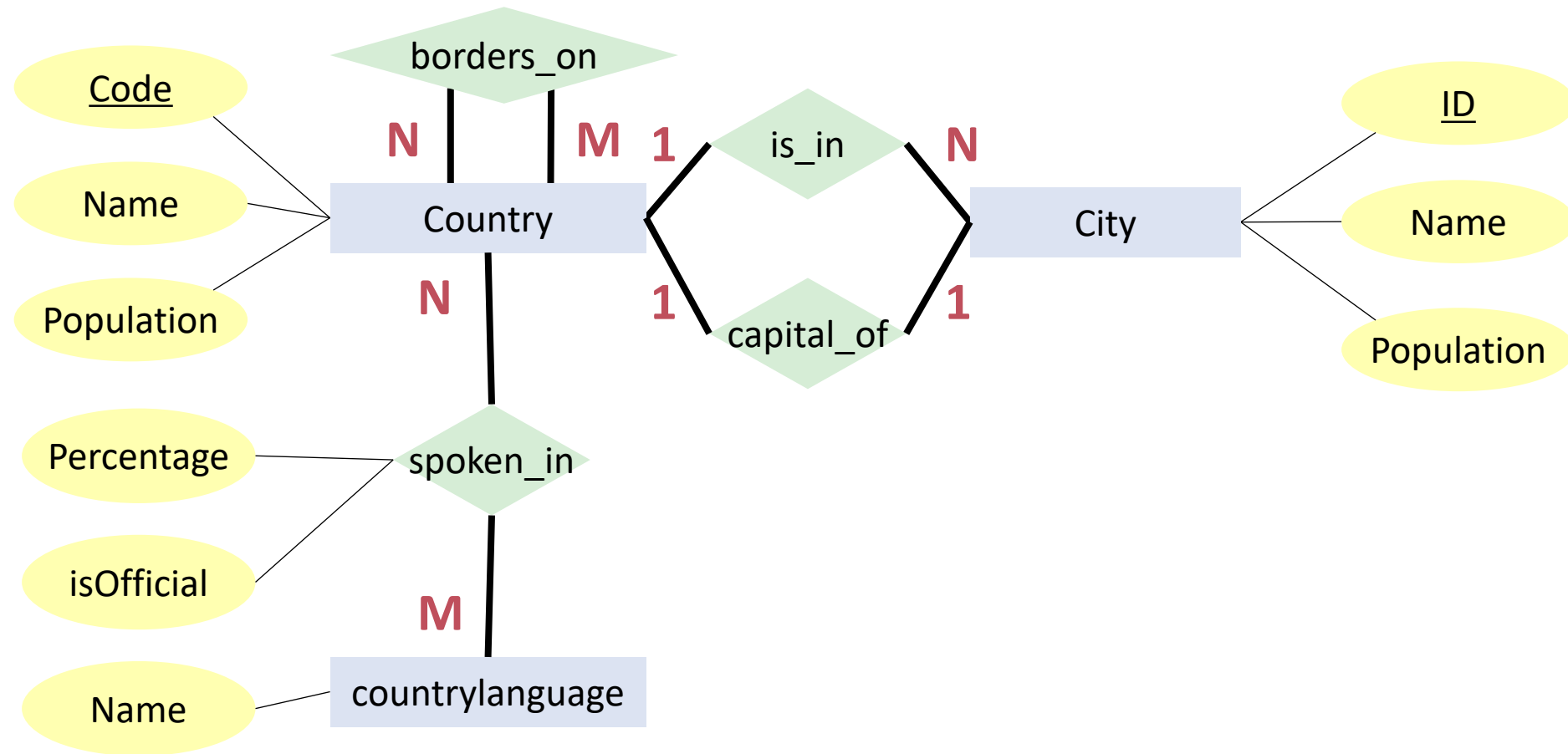
HANDS-ON: ZUGRIFF AUF SQL DATENBANKEN AUS EINER JAVA APPLIKATION

ANHANG: WELT-DATENBANK

[Ein Modell der Welt]



[Modell der Welt mit Funktionalitäten]



[Die Welt in Tabellen]

ID	Name	CountryCode ▲	District	Population
3248	Bern	CHE	Bern	122700
3247	Basel	CHE	Basel-Stadt	166700
3246	Geneve	CHE	Geneve	173500
3245	Zürich	CHE	Zürich	336800
3249	Lausanne	CHE	Vaud	114500
3117	Oldenburg	DEU	Niedersachsen	154125
3130	Remscheid	DEU	Nordrhein-Westfalen	120125
3131	Heilbronn	DEU	Baden-Württemberg	119526
3132	Pforzheim	DEU	Baden-Württemberg	117227
3133	Offenbach am Main	DEU	Hessen	116627
3134	Ulm	DEU	Baden-Württemberg	116103

CountryCode	Language	IsOfficial	Percentage
ABW	Dutch	T	5.3
ABW	English	F	9.5
ABW	Papiamentu	F	76.7
ABW	Spanish	F	7.4
AFG	Balochi	F	0.9
AFG	Dari	T	23.1

Code	Name	Continent	Region	SurfaceArea	IndepYear	Population	LifeExpectancy	GNP	GNPOld	LocalName
ABW	Aruba	North America	Caribbean	193.00	NULL	103000	78.4	828.00	793.00	Aruba
AFG	Afghanistan	Asia	Southern and Central Asia	652090.00	1919	22720000	45.9	5976.00	NULL	Afghanistan/Afq
AGO	Angola	Africa	Central Africa	1246700.00	1975	12878000	38.3	6648.00	7984.00	Angola
AIA	Anguilla	North America	Caribbean	96.00	NULL	8000	76.1	63.20	NULL	Anguilla
ALB	Albania	Europe	Southern Europe	28748.00	1912	3401200	71.6	3205.00	2500.00	Shqipëria
AND	Andorra	Europe	Southern Europe	468.00	1278	78000	83.5	1630.00	NULL	Andorra
ANT	Netherlands Antilles	North America	Caribbean	800.00	NULL	217000	74.7	1941.00	NULL	Nederlandse A
ADE	United Arab	Asia	Middle	83600.00	1971	2441000	74.1	27086.00	26946.00	Al Emarat al 'A

- (borders_on nicht modelliert)