

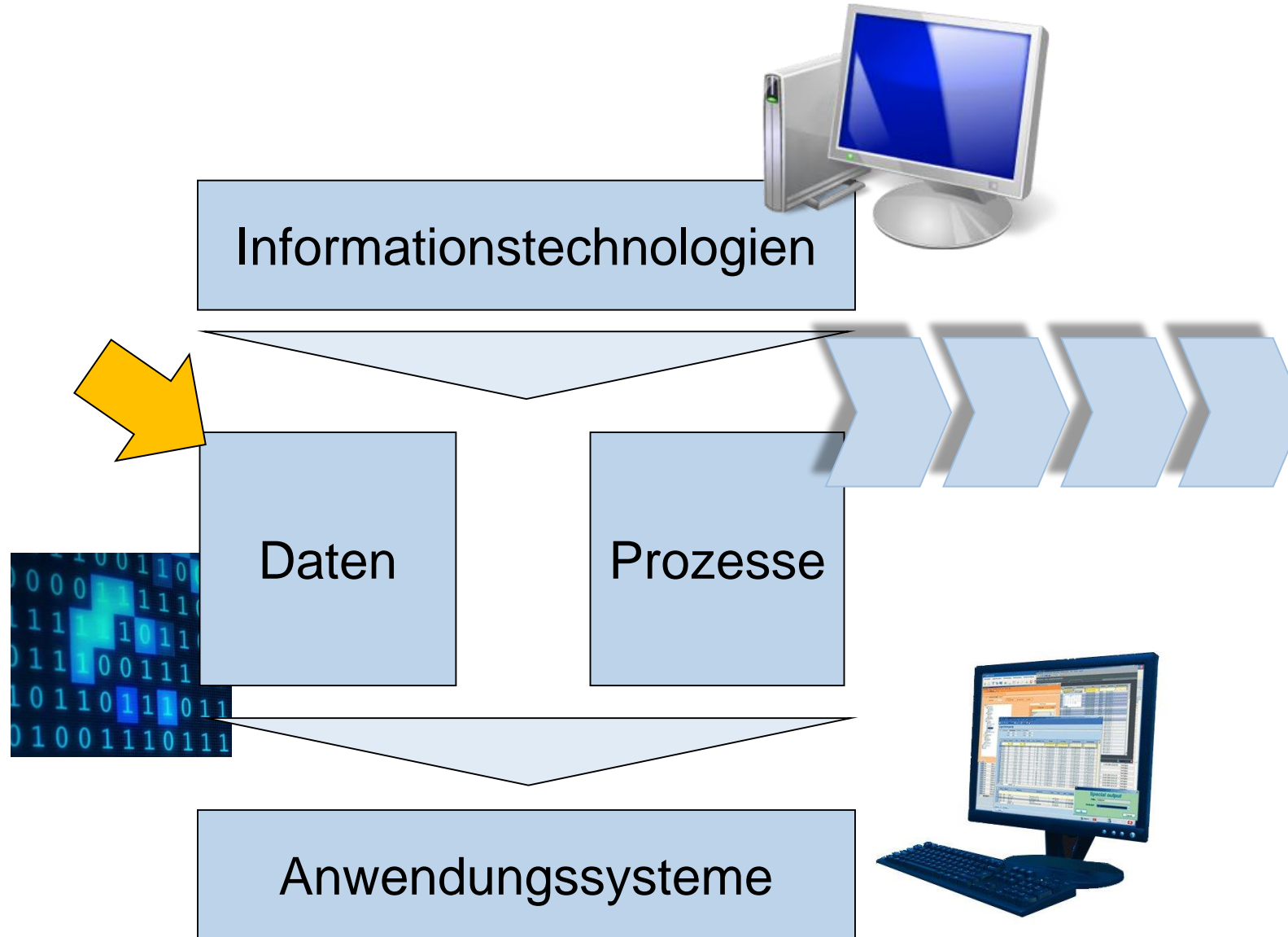
# Einführung in die Wirtschaftsinformatik

## Dateien und Datenbanksysteme: *Theorie relationaler Datenstrukturen*


**Prof. Dr. Thomas Myrach**  
**Universität Bern**  
**Institut für Wirtschaftsinformatik**  
**Abteilung Informationsmanagement**

Bern, 25. März 2020

# Logischer Aufbau



- Sie wissen, was funktionale Abhängigkeiten sind.
- Sie können funktionale Abhängigkeiten ableiten.
- Sie kennen den zentralen Begriff des Schlüssels und können ihn für eine Relation bestimmen.
- Sie wissen, warum Daten nicht immer in einer Relation abgespeichert werden sollten.
- Sie kennen die Anforderungen der ersten, zweiten und dritten Normalformen.
- Sie können Datenstrukturen entsprechend den Anforderungen der Normalformen gestalten.
- Sie kennen den Begriff des Fremdschlüssels und der referentiellen Integrität.



Funktionale Abhängigkeiten

Schlüssel

Normalisierung

Fremdschlüssel

# Funktionale Abhängigkeiten

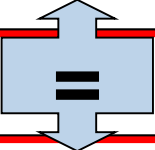
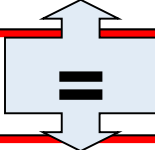
- Funktionale Abhängigkeiten sind ein formales Konstrukt, um bestimmte Zusammenhänge zwischen Attributen auszudrücken.
- Der Zusammenhang wird geschrieben als  $X \rightarrow Y$ 
  - Attribut X determiniert Attribut Y
  - Attribut Y ist funktional abhängig von Attribut X
- Funktionale Abhängigkeiten sind ein notwendiges Hilfsmittel
  - zur Bestimmung von Schlüsseln
  - zur Überprüfung der Zweckmässigkeit von Datenschemata
  - zur Ableitung zweckmässiger Datenschemata

# Definition funktionale Abhängigkeiten

- Eine funktionale Abhängigkeit  $X \rightarrow Y$  besagt:
  - $\forall r_1, r_2 \in R: r_1(X) = r_2(X) \Rightarrow r_1(Y) = r_2(Y)$   
bzw.
  - Für alle Tupel  $r_1, r_2 \in R$  gilt:  
wenn  $r_1(X) = r_2(X)$  dann  $r_1(Y) = r_2(Y)$
- Dies bedeutet:
  - Wenn bei zwei Tupeln der Relation  $R$  die Werte für das Attribut  $X$  gleich sind, so müssen die Werte für das Attribut  $Y$  ebenfalls gleich sein.
  - Für einen bestimmten Wert  $X$  darf es also nur einen Wert  $Y$  geben.

# Bedeutung funktionale Abhängigkeiten

Funktionale Abhängigkeit  $X \rightarrow Y$

	X		Y	
...	...	...	...	...
...	blabla	...	gaga	...
...		...		...
...		...	gaga	...
...	...	...	...	...

Wenn in zwei Tupeln für das Attribut X der Wert "blabla",  
dann in diesen Tupeln für das Attribut Y der gleiche Wert!

## Aufgabe: Funktionale Abhängigkeiten

- Gegeben sei eine Relation  $R(A1, A2, A3, A4)$  mit  
 $\{A1, A2 \rightarrow A4; A2 \rightarrow A3\}$
- Welche Tupel sind nicht zulässig?

A1	A2	A3	A4
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3





# Aufgabe: Bedeutung der funktionalen Abhängigkeiten

- (1) Die Vorlesungsnummer und die Matrikelnummer zusammen bestimmen die Note.
- (2) Die Matrikelnummer bestimmt den Namen.

Vorlesung	Matrikelnr	Name	Note
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3



# Redundanz und Inkonsistenz

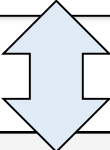
- **Redundanz** liegt vor, wenn identische Daten über ein Realweltobjekt mehrfach abgespeichert werden.
- Funktionale Abhängigkeiten zeigen mögliche Redundanzen an:
  - Wenn  $X \rightarrow Y$  und in zwei Tupeln der Relation sind die Werte für die Attribute X und Y gleich, dann ist die mehrfache Nennung des Wertes von Y redundant!
- **Inkonsistenz** liegt vor, wenn sich Daten über ein Realweltobjekt widersprechen.
- Funktionale Abhängigkeiten können bei Inkonsistenzen verletzt werden:
  - Wenn  $X \rightarrow Y$  und in zwei Tupeln der Relation sind die Werte für das Attribut X gleich und für Y ungleich, dann sind die Werte von Y inkonsistent!

## Aufgabe: FA (1) Redundanz

Vorlesung, Matrikelnr	→	Note
-----------------------	---	------

Vorlesung	Matrikelnr	Name	Note
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

## Aufgabe: FA (1) Inkonsistenz

Vorlesung, Matrikelnr		→	Note
			<b>Widerspruch!</b>
Vorlesung	Matrikelnr	Name	Note
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

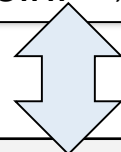
## Aufgabe: FA (2) Redundanz

Matrikelnr → Name

Vorlesung	Matrikelnr	Name	Note
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

## Aufgabe: FA (2) Inkonsistenz

Matrikelnr → Name



**Widerspruch!**

Vorlesung	Matrikelnr	Name	Note
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

- Durch funktionale Abhängigkeiten ist festgelegt, dass in einer Relation bezüglich der abhängigen Attribute Redundanzen auftreten können.
- Treten bei mehreren Tupeln mit gleichen Werten bei den Determinanten für die abhängigen Attribute keine gleichen Werte auf, so liegen Inkonsistenzen vor.
- In diesem Fall können die abhängigen Werte nicht (alle) gültig sein.
- Beim Vorliegen von Inkonsistenzen besteht ein Zweifel darüber, ob und welche Daten gültig sind.
- Inkonsistenzen mindern die Datenqualität.
- Ein gutes Datenbankdesign sollte die Gefahr von Inkonsistenzen nach Möglichkeit ausschalten.



Funktionale Abhängigkeiten

Schlüssel

Normalisierung

Fremdschlüssel



- Auf die Tupel einer Relation kann nur über Attributwerte zugegriffen werden.
- Um auf ein bestimmtes Tupel zugreifen zu können, darf zumindest ein Attributwert bzw. eine Kombination von Attributwerten nur einmal in der Relation vorkommen.
- Diese Eigenschaft wird als Schlüsselintegrität bezeichnet
- Ein Schlüssel garantiert die Schlüsselintegrität und damit den gezielten Zugriff auf ein Tupel.
- Jede Relation muss (mindestens) einen Schlüssel haben.
- Einer der Schlüssel muss als Primärschlüssel definiert sein.
- Schlüssel werden typischerweise bei der Definition eines Datenschemas definiert.

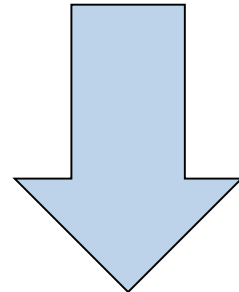
# Eigenschaften eines Schlüssels

- Eindeutigkeit
  - Eine Menge von Attributen ist eindeutig, wenn für sie eine bestimmte Kombination von Attributwerten höchstens einmal in einer Relation vorkommt.
  - Sei  $R$  die Menge aller Attribute und  $S$  eine Untermenge der Relation mit  $S \subseteq R$ , dann muss für Eindeutigkeit die funktionale Abhängigkeit  $S \rightarrow R$  gelten, d.h.
    - für alle  $r_1, r_2 \in R$  gilt: wenn  $r_1(S) = r_2(S)$  dann  $r_1(R) = r_2(R)$
    - die Relation wird durch den Schlüssel determiniert.
- Minimalität
  - Minimalität ist gegeben, wenn von  $S$  kein Attribut entfallen kann, ohne dass die Eindeutigkeit verloren geht.
  - Es existiert kein  $T$  mit  $T \subset S$  für das gilt  $T \rightarrow R$ .

# Ableiten eines Schlüssels

Vorlesung, Matrikelnr  $\rightarrow$  Note

Matrikelnr  $\rightarrow$  Name



Wie komme ich von den  
gegebenen funktionalen  
Abhängigkeiten auf den  
Schlüssel?

???  $\rightarrow$  Vorlesung, Matrikelnr, Name, Note

Schlüssel S  $\rightarrow$  Relation R

- Reflexivität:

$$X \rightarrow V \text{ mit } V \subseteq X$$

- Erweiterung:

$$X \rightarrow Y \text{ impliziert } XZ \rightarrow YZ$$

- Transitivität:

$$X \rightarrow Y \text{ und } Y \rightarrow Z \text{ impliziert } X \rightarrow Z$$

- Additivität:

$$X \rightarrow Y \text{ und } X \rightarrow Z \text{ impliziert } X \rightarrow YZ$$

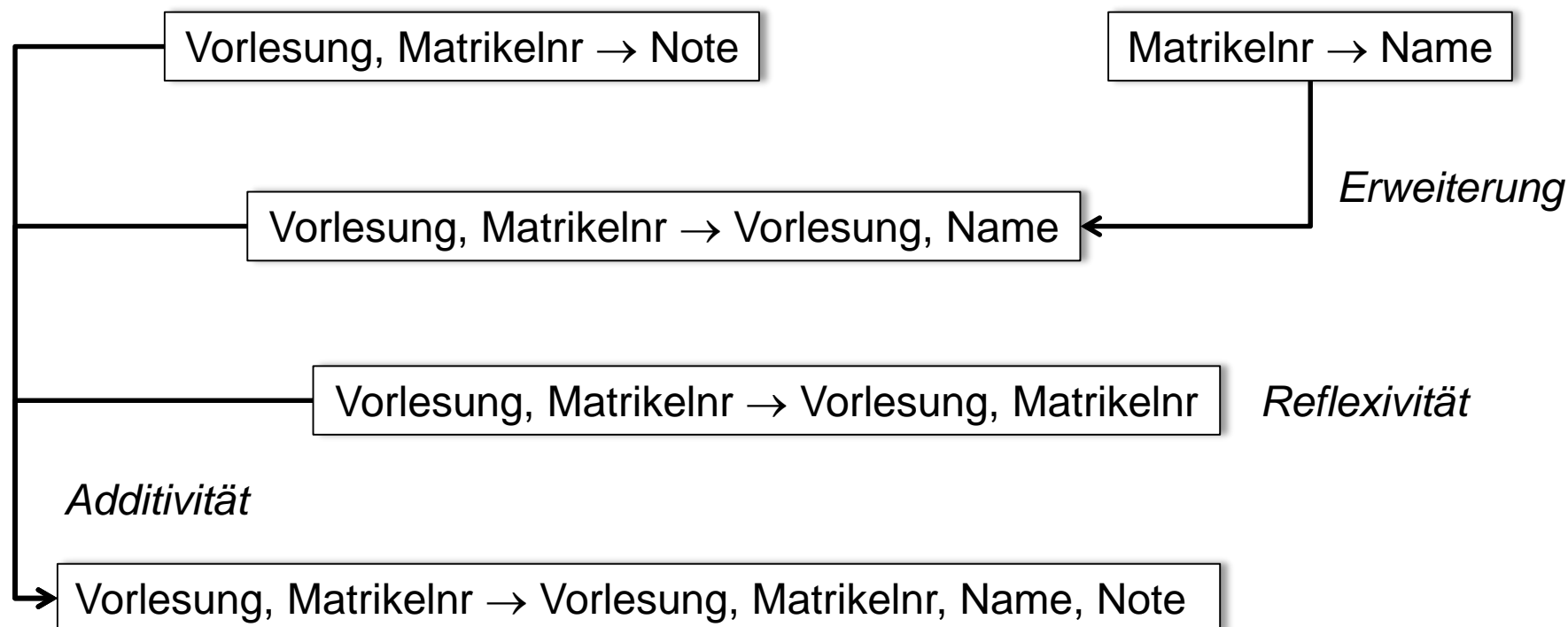
- Projektivität:

$$X \rightarrow YZ \text{ impliziert } X \rightarrow Y \text{ und } X \rightarrow Z$$

- Pseudotransitivität:

$$X \rightarrow Y \text{ und } YZ \rightarrow W \text{ impliziert } XZ \rightarrow W$$

# Ableiten eines Schlüssels



Der Schlüssel setzt sich aus den beiden Attributen  
Vorlesung und Matrikelnr zusammen.

# Schlüsseleindeutigkeit (1)

Vorlesung, Matrikelnr → Name, Note

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4711	B	5
W300	4711	A	3
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

## Schlüsseindeutigkeit (2)

Die **Schlüsseindeutigkeit** erlaubt theoretisch, dass Tupel mit gleichen Schlüsselwerte auftreten, wenn diese Tupel die gleichen Werte haben.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4711	A	5
W300	4711	A	5
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

## Schlüsseleindeutigkeit (3)

Die **Schlüsselintegrität** führt in der Praxis dazu, dass pro Schlüsselwert nur ein Tupel auftritt, wodurch die **Schlüsseleindeutigkeit** gewahrt wird.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3



# Schlüssel und Redundanz (1)

Wegen der **Schlüsseleindeutigkeit** kann kein Tupel eingefügt werden, wenn die Schlüsselwerte bereits vorhanden sind.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

insert into R value  ("W300","4711","B",4)

Durch die Schlüsseleindeutigkeit werden gewisse Redundanzen und somit auch Inkonsistenzen ausgeschlossen.

# Schlüssel und Redundanz (2)

Die **Schlüsseleindeutigkeit** schliesst nicht notwendigerweise alle Redundanzen und Inkonsistenzen aus.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

insert into R values ("W303","4711","B",4)



- Mit einem Schlüssel kann jederzeit eindeutig auf ein bestimmtes Tupel einer Relation zugegriffen werden.
- Dies bedeutet:
  - Ein bestimmter Schlüsselwert tritt höchstens einmal in einer Relation auf
  - Würde ein Schlüsselwert mehrfach auftreten, so müsste es sich um identische Tupel handeln.
  - Mehrfach auftretende Tupel sind jedoch redundant und sollten entfallen.
- Für Schlüssel gilt üblicherweise die Eindeutigkeitsrestriktion.
  - Ist ein bestimmter Schlüsselwert in einer Relation bereits vorhanden, so kann dieser Wert nicht erneut eingegeben werden.



Funktionale Abhängigkeiten

Schlüssel

Normalisierung

Fremdschlüssel

**Ich schwöre,  
dass jedes Attribut meiner Relation  
vom Schlüssel abhängt,  
vom ganzen Schlüssel  
und nichts als dem Schlüssel,  
so wahr mir Codd helfe.**

- ... vom Schlüssel abhängt ...: 1. Normalform
- ... vom ganzen Schlüssel ...: 2. Normalform
- ... nichts als dem Schlüssel ...: 3. Normalform

... vom Schlüssel abhängt ...

- Jedes Attribut einer Relation muss vom Schlüssel abhängen.
- Jedes Attribut kann pro Tupel nur je einen Wert haben.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

**Vorlesung, Matrikelnr → Vorlesung, Matrikelnr, Name, Note**

... vom ganzen Schlüssel ...

- Kein Attribut sollte nur vom Teil eines Schlüssels identifiziert werden.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Name	Note
W300	4711	A	5
W300	4712	A	5
W301	4711	A	3
W301	4712	A	3
W302	4712	C	3
W302	4713	A	3

**Vorlesung Matrikelnr → Name, Note**

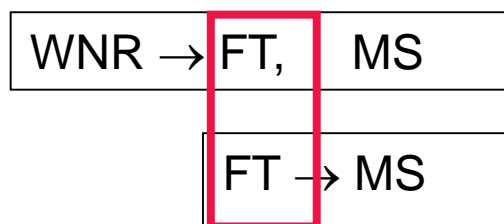
**Matrikelnr → Name**

Name hängt von einem  
Teil des Schlüssels ab!  
2.Normalform ist verletzt.

... nichts als dem Schlüssel ...

- Kein Attribut sollte transitiv vom Schlüssel abhängen.

<u>WNR</u>	FT	MS
W12	Corsa	135
W84	Mondeo	185
W85	Vectra	185
W33	Golf	160
W75	Mondeo	185



Mietsatz (MS) hängt vom Fahrzeugtyp (FT) ab!  
3. Normalform ist verletzt.



# Konsequenzen von Verletzungen der Normalformen

- Werden in einer Relation die Normalformen nicht eingehalten, so können Datenmanipulation problematische Ergebnisse haben.
- Redundanzen und Inkonsistenzen können nicht ausgeschlossen werden.
- Durch eine Zerlegung der Relation in Teilrelationen lässt sich dieses Problem beseitigen.
- Jede der Teilrelationen sollte die Normalformen einhalten.
- Dann kann über die Schlüsselintegrität ausgeschlossen werden, dass Redundanzen und Inkonsistenzen auftreten.

# Zerlegung der Relation Notenmeldung (1)

Die **Schlüsseleindeutigkeit** verbietet mehrere gleiche Schlüsselwerte in einer Relation.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3
W302	4713	3

**Vorlesung, Matrikelnr → Note**

<u>Matrikelnr</u>	Name
4711	A
4712	A
4712	A
4712	C
4713	A

**Matrikelnr → Name**

Die zerlegten Relationen erfüllen alle Normalformen!

## Zerlegung der Relation Notenmeldung (2)

Durch eine geeignete Zerlegung der Relation können alle Redundanzen und Inkonsistenzen ausgeschlossen werden.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3
W302	4713	3

<u>Matrikelnr</u>	Name
4711	A
4712	C
4713	A

insert into R2  
values ("~~4712~~", "A")

Durch Zerlegung der Relation kann ausgeschlossen werden, dass für einen Studierenden unterschiedliche Namen erscheinen!

## Zerlegung der Relation Mietsatz (1)

Die **Schlüsseindeutigkeit** verbietet mehrere gleiche Schlüsselwerte in einer Relation.

<u>WNR</u>	FT
W12	Corsa
W84	Mondeo
W85	Vectra
W33	Golf
W75	Mondeo

WNR → FT

<u>FT</u>	MS
Corsa	135
Mondeo	185
Vectra	185
Golf	160
Mondeo	185

FT → MS

## Zerlegung der Relation Mietsatz (2)

Durch eine geeignete Zerlegung der Relation können alle Redundanzen und Inkonsistenzen ausgeschlossen werden.

<u>WNR</u>	FT
W12	Corsa
W84	Mondeo
W85	Vectra
W33	Golf
W75	Mondeo

<u>FT</u>	MS
Corsa	135
Mondeo	185
Vectra	185
Golf	160

insert into R2  
values ("Mondeo", "170")

Durch Zerlegung der Relation kann ausgeschlossen werden, dass für einen Fahrzeugtyp unterschiedliche Mietsätze erscheinen!

# Konsequenzen der Zerlegung von Relationen

- Durch eine zweckmässige Datenstrukturierung kann die Qualität von Daten begünstigt werden.
- Ein zentraler Punkt ist dabei die Vermeidung von Redundanzen und damit auch die Gefahr von Inkonsistenzen.
- Dabei spielt die Schlüsseleindeutigkeit eine wichtige Rolle.
- Durch Normalformen lässt sich die Zweckmässigkeit von Relationen prüfen.
- Die Einhaltung der Normalformen erfordert unter Umständen eine Zerlegung von Relationen.
- Die Information in den zerlegten Relationen muss bei Bedarf wieder zusammengeführt werden können.



Funktionale Abhängigkeiten

Schlüssel

Normalisierung

Fremdschlüssel

- Ein Fremdschlüssel referenziert einen Primärschlüssel.
- Ein Fremdschlüssel darf nur Attributwerte aufweisen, die im referenzierten Primärschlüssel enthalten sind.
- Beim Fremdschlüssel handelt es sich quasi um einen Schlüssel einer anderen Relation.
- Ein Fremdschlüssel muss selbst kein Schlüssel sein.
- Fremdschlüssel und Primärschlüssel bildet Zusammenhänge zwischen Relationen ab.
- Diese Zusammenhänge entstehen dadurch, dass zusammengehörige Daten in mehrere Relationen aufgeteilt werden.

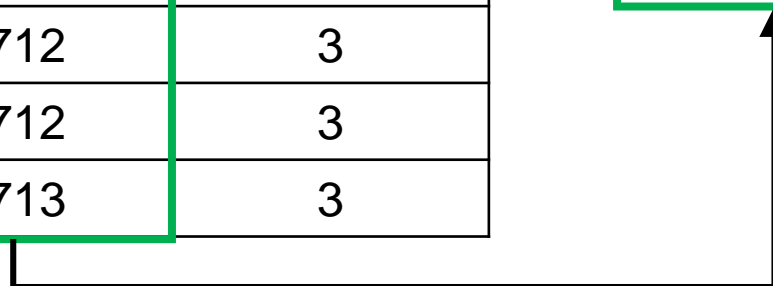


# Zusammenhänge zwischen Relationen

Beziehungen zwischen Relationen werden durch Fremdschlüssel hergestellt.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3
W302	4713	3

<u>Matrikelnr</u>	Name
4711	A
4712	C
4713	A



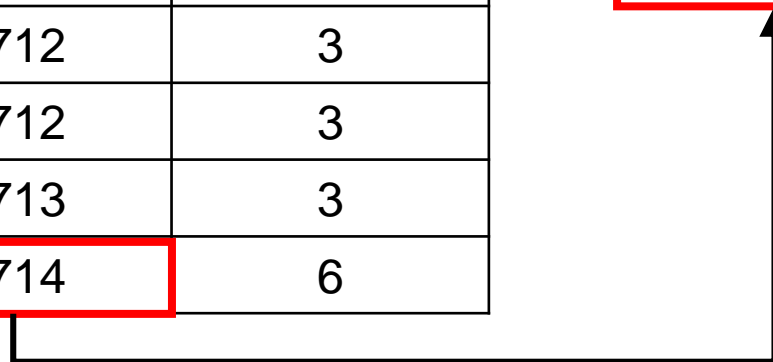
Jede Matrikelnummer in der Notenmeldung (Fremdschlüssel) muss einer Matrikelnummer der Studierendenrelation (Primärschlüssel) entsprechen!

# Zusammenhänge zwischen Relationen

Lassen sich die Zusammenhänge zwischen den Daten der Relationen nicht herstellen, so sind Informationen unvollständig.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3
W302	4713	3
W302	4714	6

<u>Matrikelnr</u>	Name
4711	A
4712	C
4713	A



Welche Namen hat der Student mit der Matrikelnummer 4714?

- Betrifft die Beziehungen zwischen Fremdschlüsseln und Primärschlüsseln.
- Alle Werte eines Fremdschlüssels müssen den Werten eines referenzierten Primärschlüssels entsprechen.
- Durch Datenbankmechanismen lässt sich die referentielle Integrität erzwingen.
- Die Auswirkungen sind zweiseitig:
  - (1) Kein Tupel kann in eine Relation eingefügt werden, wenn der Wert des Fremdschlüssels nicht dem eines referenzierten Primärschlüssels entspricht.
  - (2) Kein Tupel kann aus einer Relation gelöscht werden, wenn der Wert des Primärschlüssels durch Fremdschlüssel referenziert wird.

# Referentielle Integrität (1)

Mechanismen der referentiellen Integrität verhindern, dass durch Einfügung eines Sekundärschlüsselwerts die Beziehungen gestört werden.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3
W302	4713	3

<u>Matrikelnr</u>	Name
4711	A
4712	C
4713	A

insert into R1  
values ("W302", "4714", 6)

## Referentielle Integrität (2)

Mechanismen der referentiellen Integrität verhindern, dass durch Löschung eines Primärschlüsselwerts die Beziehungen gestört werden.

<u>Vorlesung</u>	<u>Matrikelnr</u>	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3
W302	4713	3

<u>Matrikelnr</u>	Name
4711	A
4712	C
4713	A

delete from R2  
where matrikelnr="4713"

# Auswirkungen der referentiellen Integrität

- Durch die referentielle Integrität bleiben die Zusammenhänge zwischen Daten aus verschiedenen Relationen gewahrt.
- Datenbankoperationen, die zu Integritätsfehlern führen, werden abgewiesen.
- Diese Sicherung kann in der praktischen Arbeit auch störende Konsequenzen haben.
- Ohne die genaue Kenntnis der Datenabhängigkeiten kann die Durchführung von Datenoperationen mühsam sein.

- Das Relationale Datenmodell bietet mit der Tabelle eine einfache Abstraktion für die Strukturierung von Daten an.
- Realistische Datenobjekte lassen sich jedoch nicht ohne weiteres mit einer einzigen Tabelle abbilden.
- Die Gestaltung von Relationen nach den Grundsätzen der Normalisierung führt schnell einmal zu einer ganzen Reihe von Tabellen.
- Diese hängen über Primär-/Fremdschlüsselbeziehungen untereinander zusammen.
- Auch wenn diese Zusammenhänge nicht direkt sichtbar sind, müssen sie für die Manipulation der abgelegten Daten genau beachtet werden.
- Dadurch entstehen trotz des einfachen Grundkonstrukts unter Umständen komplexe Datenschemata.