

UNIVERSITÄT BERN

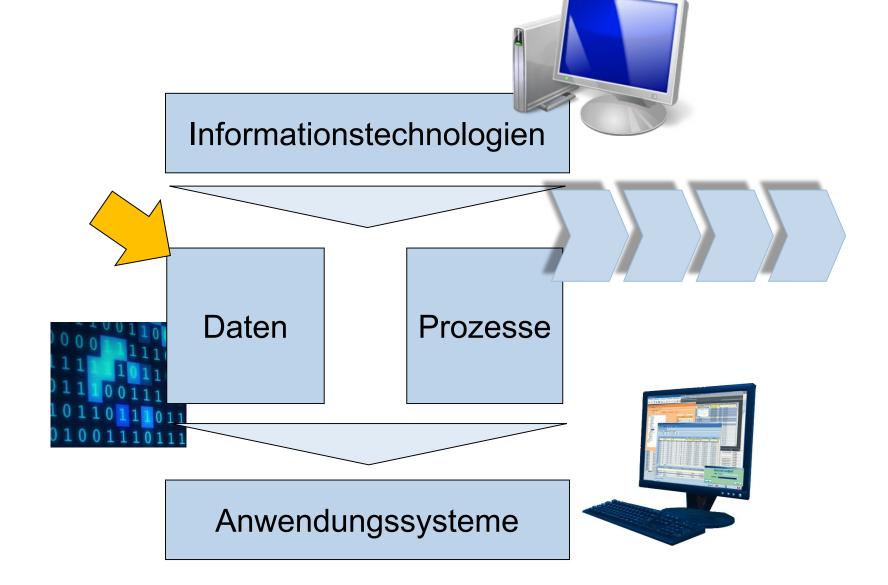
Einführung in die Wirtschaftsinformatik

Dateien und Datenbanksysteme: Grundlagen relationaler Datenbanken

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

Logischer Aufbau





Lernziele



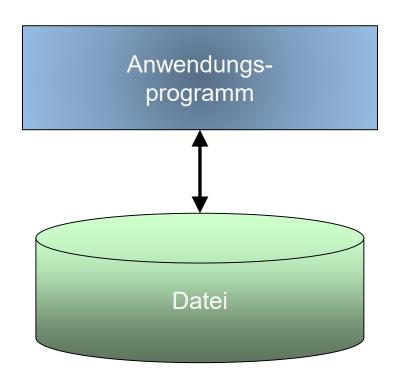
- Sie kennen die Problematik dokumentenorientierter Dateispeicherung.
- Sie kennen das Konzept der logischen und physischen Datenunabhängigkeit.
- Sie wissen, das Datenbanksysteme häufig eine Client-Server-Architektur haben.
- Sie lernen die Eigenschaften des weit verbreiteten Relationalen Datenmodells kennen.
- Sie wissen, was mit elementaren Attributwerten gemeint ist und warum diese für Relationen wichtig sind.
- Sie kennen den Unterschied zwischen Tabellen im Relationalen Datenmodell und in Tabellenkalkulationsprogrammen.
- Sie können wichtige Datenmanipulationsoperationen in SQL formulieren.

Gliederung









- Die Verarbeitung von Daten erfolgt über Anwendungsprogramme.
- Anwendungsprogramme verwalten die von ihnen benötigten Daten herkömmlicherweise selber.
- Die von einem Programm verwalteten Daten entsprechen den Anforderungen des Programms.
- Das können allgemein bestimmte Dateiformate sein oder Dateien mit spezifischen Datenstrukturen.



Beispiel: Notenmeldung

Dokument

NOTENMELDUNG

Vorlesnr: 4711

Vorlesung: Digitale Welten

Dozent: Myrach

Semester: FS 2019

ECTS: 4

Datum: 2019-03-20

Matrnr	Name	Vname	Note
8912307	Müller	Jürg	6.0
9056701	Meier	Urs	1.0

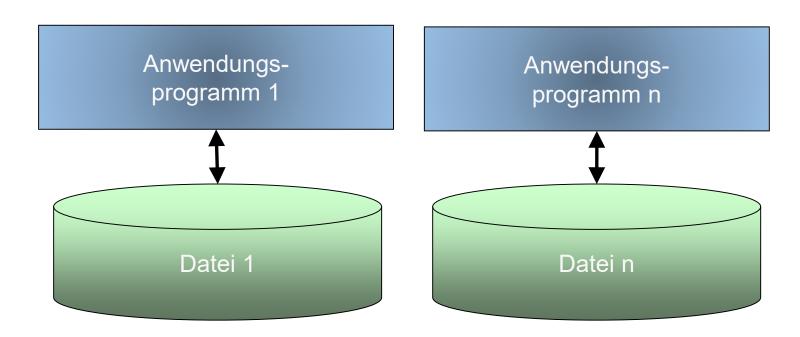
```
Datei (hier: JSON)
"vorlesnr":"4711",
"vorlesung":"Digitale Welten",
"dozent":"Myrach",
"semester": "FS 2019",
"ects":"4",
"datum":"2019-03-20",
"students":[
 {"matrnr":"8912307", "name":"Müller",
  "vname":"Jürg","note":"6.0"},
 {"matrnr":"9056701","name":"Meier",
  "vname":"Urs","note":"1.0"}]
```



UNIVERSITÄT BERN

Verwendung von Daten in verschiedenen Kontexten

- In der Praxis werden Daten über bestimmte Sachverhalte oftmals in verschiedenen Zusammenhängen benötigt.
- Die Ablage von Daten gemäss den Anforderungen der jeweiligen Anwendung kann zu Redundanzen führen.





u^{b}

UNIVERSITÄT BERN

Beispiel: Dokumente mit gleichen Daten

NOTENMELDUNG						
Vorlesnr: 4711						
Vorlesung	•	Dig	gitale Wei	lten		
Dozent:		My	rach			
Semester:		FS	2008			
ECTS: 4						
Datum: 2008-04-26						
Matrnr	Nam	ie	Vname	Note		
8912307	Müll	er	Jürg	6.0		

STUDIENBLATT						
Matrnr: 8	912307					
Name: <i>Müller</i>						
Vname: Jürg						
HFach: E	BWL					
NFach: II	nformatik					
Vorlesnr	Vorlesung	Dozent		Note		
4711	Digitale Well	ten Myrach		6.0		



Beispiel: Dateien mit gleichen Daten

```
"vorlesnr":"4711",
"vorlesung":"Digitale Welten",
"dozent":"Myrach",
"semester":"FS 2019",
"ects":"4",
"datum":"2019-03-20",
"students":[
  "matrnr":"8912307",
  "name":"Müller",
  "vname":"Jürg",
  "note":"6.0"},
```

```
"matrnr":"8912307",
"name":"Müller",
"vname":"Jürg",
"hfach": "BWL",
"nfach":"Informatik",
"vorlesungen":[
   "vorlesnr":"4711",
   "vorlesung": "Digitale Welten",
   "dozent":"Myrach",
   "semester": "FS 2019",
   "ects":"4",
   "note":"6.0"},
```

Zwischenfazit



- Dateien werden durch Programme angelegt, verändert und gelesen.
- Dokumente können dokumentzentriert in Dateiformaten wie etwa XML oder JSON abgebildet werden.
- Sie bilden typischerweise die für ein Programm relevante Sicht auf die (Daten-)Welt ab.
- Bestimmte Sachverhalte k\u00f6nnen in verschiedenen Dateien in unterschiedlichen Kontexten als Daten abgebildet sein.
- Dies führt dann zu einer redundanten Datenhaltung von Daten in den unterschiedlichen Dateien.
- Redundante Datenhaltung stellt ein Problem dar, etwa wenn Daten zu verändern sind.

Gliederung





Datenbanksysteme



UNIVERSITÄT BERN

- Ein Datenbankmanagementsystem (DBMS) ist eine systemnahe Software für die Verwaltung von Datenbanken.
- Durch DBMS können Daten unabhängig von Anwendungsprogrammen verwaltet werden.
- Komplexe Funktionen der Datenverwaltung brauchen nicht durch Anwendungsprogramme abgedeckt werden.

Anwendungssystem

Datenbanksystem

Betriebssystem

Server-Rechner

Speichersysteme

Netzwerk

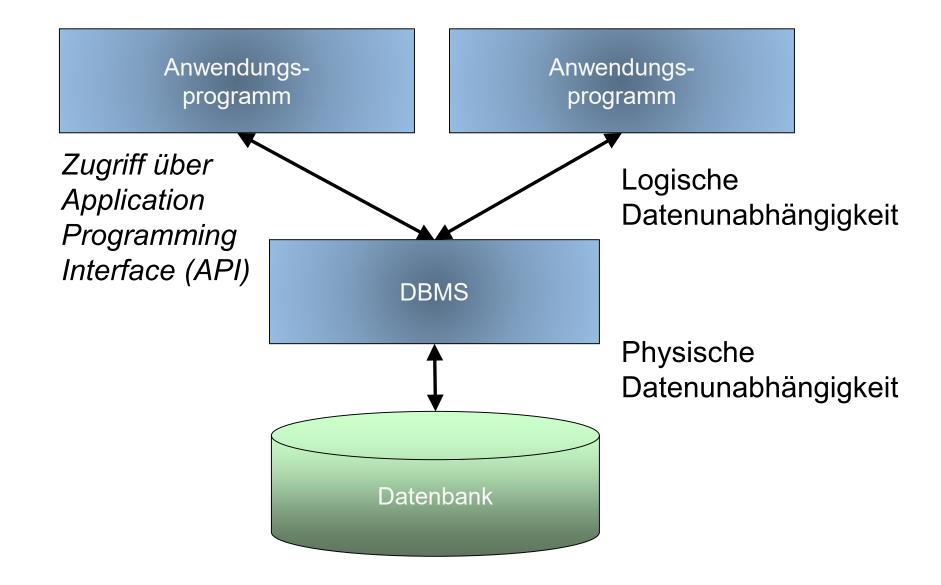
Schöne neue (Datenbank-) Welt



- Idealtypische Annahme, dass alle relevanten Daten in einer einzigen Datenbank gehalten werden.
- Grundlage ist ein einheitliches Datenschema, welches unabhängig von verschiedenen Anwendungen ist.
- Daten über ein bestimmtes Objekt werden nur einmal gehalten.
- Verschiedene Datenbedürfnisse über ein Objekt werden durch Sichten ausgedrückt.
- Programme greifen über die Sichten auf die von ihnen benötigten Daten (und nur diese) zu.

Architektur eines DBMS





Datenunabhängigkeit



- Logische Datenunabhängigkeit
 - Die Verwaltung der Daten erfolgt unabhängig von den Anwendungsprogrammen, die auf sie zugreifen.
 - Eine Änderung der Datenstruktur erfordert nur dann eine Anpassung der Anwendungsprogramme, wenn die Änderung die unmittelbar von ihnen benötigten Daten betrifft.
- Physische Datenunabhängigkeit
 - Die Verwaltung der Daten erfolgt unabhängig davon, wie sie physisch auf einem Speichermedium abgelegt sind.
 - Eine Änderung der Organisation der Daten auf den Speichermedien erfordert keine Anpassung des (logischen) Datenschemas.

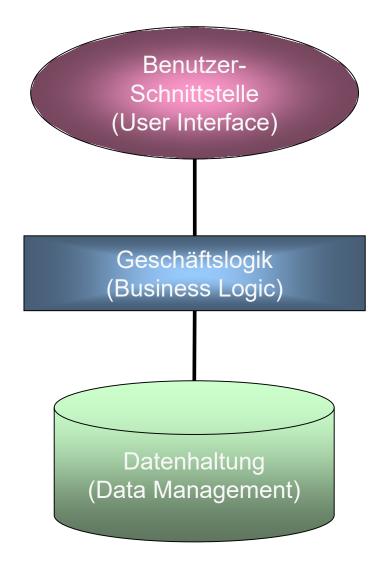
Datenbanksysteme und verteilte DV

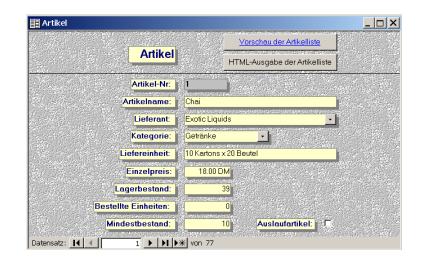


- Datenbanksysteme werden heutzutage typischerweise im Rahmen einer verteilten DV betrieben.
- Dabei wird üblicherweise die Client-Server-Architektur implementiert:
 - Die Datenbank mit dem Datenbankmanagementsystem bilden die Serverkomponente.
 - Programme können als Clients von verschiedenen Orten über das API auf das Datenbankmanagementsystem zugreifen.
- Die örtlich entkoppelte Erreichbarkeit von Datenbanken begünstig eine flexible Nutzung der Daten.
- Im Rahmen des Internets sind Datenbanken weltweit erreichbar.

Schematischer Aufbau eines Anwendungssystems



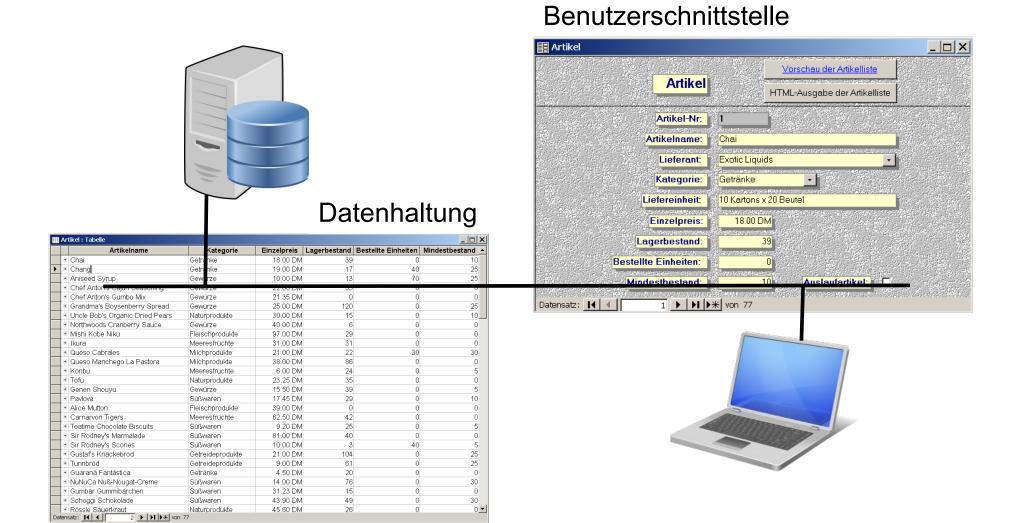




П	Artikelname	Kategorie	Einzelpreis	Lagerbestand	Bestellte Einheiten	Mindestbestand
T	± Chai	Getränke	18.00 DM	39	0	10
1	Chang	Getränke	19.00 DM	17	40	25
ı	Aniseed Syrup	Gewürze	10.00 DM	13	70	25
ı	■ Chef Anton's Cajun Seasoning	Gewürze	22.00 DM	53	0	0
1	Chef Anton's Gumbo Mix	Gewürze	21.35 DM	0	0	0
1	Grandma's Boysenberry Spread	Gewürze	25.00 DM	120	0	25
1	■ Uncle Bob's Organic Dried Pears	Naturprodukte	30.00 DM	15	0	10
ı	Northwoods Cranberry Sauce	Gewürze	40.00 DM	6	0	0
ı	Mishi Kobe Niku	Fleischprodukte	97.00 DM	29	0	0
1	* Ikura	Meeresfrüchte	31.00 DM	31	0	0
1	Queso Cabrales	Milchprodukte	21.00 DM	22	30	30
1	■ Queso Manchego La Pastora	Milchprodukte	38.00 DM	86	0	0
1	■ Konbu	Meeresfrüchte	6.00 DM	24	0	5
I	* Tofu	Naturprodukte	23.25 DM	35	0	0
1	Genen Shouyu	Gewürze	15.50 DM	39	0	5
1		Süßwaren	17.45 DM	29	0	10
1	Alice Mutton	Fleischprodukte	39.00 DM	0	0	0
I	Carnarvon Tigers	Meeresfrüchte	62.50 DM	42	0	0
1	Teatime Chocolate Biscuits	Süßwaren	9.20 DM	25	0	5
1	Sir Rodney's Marmalade	Süßwaren	81.00 DM	40	0	0
	Sir Rodney's Scones	Süßwaren	10.00 DM	3	40	5
ı	■ Gustaf's Knäckebröd	Getreideprodukte	21.00 DM	104	0	25
ı	* Tunnbröd	Getreideprodukte	9.00 DM	61	0	25
1	Guaraná Fantástica	Getränke	4.50 DM	20	0	0
	■ NuNuCa Nuß-Nougat-Creme	Süßwaren	14.00 DM	76	0	30
	■ Gumbär Gummibärchen	Süßwaren	31.23 DM	15	0	0
	Schoggi Schokolade	Süßwaren	43.90 DM	49	0	30
	⊞ Rössle Sauerkraut	Naturprodukte	45.60 DM	26	0	0

Client-Server-Verarbeitung





Gliederung





Relationenmodell: Eigenschaften



UNIVERSITÄT BERN

- Das relationale Datenmodell ist derzeit das populärste in der Praxis verwendete Datenmodell.
- Es zeichnet sich durch folgende Eigenschaften aus:

Vorteile:

- Einfachheit
- Verständlichkeit
- Flexibilität
- Methodische Grundlage auf Basis der Mengentheorie
- Normierung

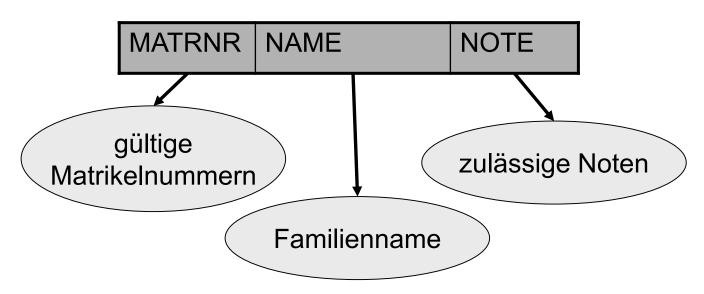
Nachteile:

- Starke Zerlegung bei Einhaltung der Normalformen
- Bei komplexen Strukturen schlecht überschaubar.
- Vergleichsweise ressourcenintensiv.

Relationenschema



- Bevor Daten in eine Relation eingefügt werden können, muss ein Relationenschema erstellt werden.
- Ein Relationenschema ist eine geordnete Liste von Attributen.
- Ein Attribut ist eine Funktion, die auf eine Domäne abbildet.
- Eine Domäne ist eine Menge von Werten, die für ein Attribut zulässig ist.



Relation



- Eine Relation ist eine Menge von Tupeln.
- Es ist die Extension eines Relationenschemas.
- Ein Tupel ist eine geordnete Liste von Attributwerten.
- Für jedes Attribut der Relation hat ein Tupel (normalerweise) genau einen bestimmten Wert.

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6

Attributwerte



- Alle Attributwerte in einem Tupel müssen üblicherweise elementar sein.
 - Nur ein Wert ist im Schnittpunkt einer Zeile und einer Spalte zulässig.
 - Für ein Tupel dürfen nicht mehrere Werte eines Attributes relevant sein oder Wiederholgruppen auftreten.
- Nicht alle Attributwerte in einem Tupel brauchen definiert zu sein.
 - Für die Kennzeichnung undefinierter Attributwerte ist ein spezieller Nullwert vorgesehen.
 - Undefinierte Attributwerte sind möglich, wenn das Attribut für das durch das Tupel abgebildete Objekt nicht relevant oder nicht bekannt ist.

Vom Dokument zur Relation



UNIVERSITÄT BERN

NOTENMELDUNG

Vorlesnr: 4711

Vorlesung: Digitale Welten

Dozent: *Myrach*Semester: *FS 2008*

ECTS: 4

Datum: 2008-04-26

Matrnr	Name	Vname	Note
8912307	Müller	Jürg	6.0
9056701	Meier	Urs	1.0

- Relationen k\u00f6nnen die Daten von Dokumenten enthalten.
- Dafür müssen die Daten des Dokuments strukturell umgeformt werden.
- Grundannahme:
 Alle Daten eines Dokuments sollen in einer Relation enthalten sein.

Elementare Attributwerte



Vorlesnr	Vorlesung	Math	Name	Vname	Note
4711	Digitale Welt	 9123	Müller	Jürg	6.0
		9	Meier	Urs	1.0

Vorlesnr	Vorlesung	 Matrnr	Name	Vname	Note
4711	Digitale Welten	 8912307	Müller	Jürg	6.0
4711	Digitale Welten	 9056701	Meier	Urs	1.0

Exkurs: Daten in Tabellen



UNIVERSITÄ[.] BERN

Tabellenkalkulationsblatt versus Relationen

Tabellenkalkulationsblatt:

- Zeilen und Spalten werden über Indizes adressiert
- Elemente (Zellen) werden über Zeilen- und Spaltenangaben angesprochen
- Zellen können Datenwerte oder Formeln enthalten.
- Zellen sind prinzipiell unabhängig voneinander.

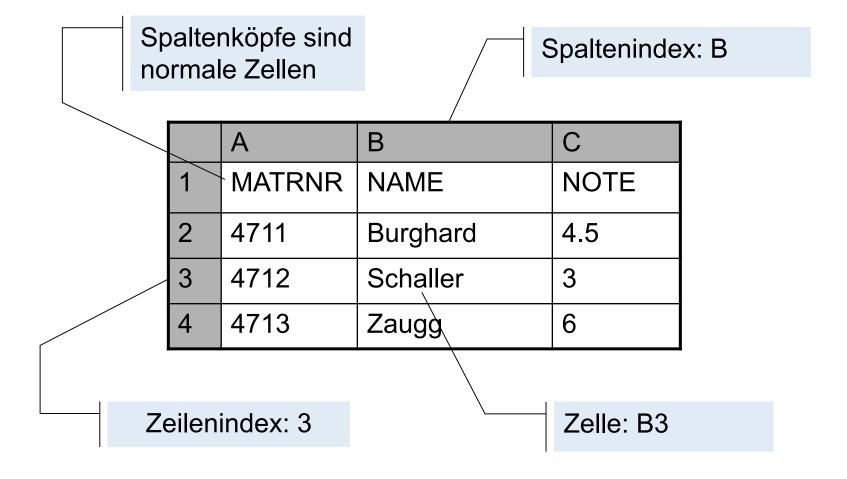
– Relationen:

- Spalten können über Attributnamen angesprochen werden
- Zeilen können nur über bestimmte Attributwerte angesprochen werden
- Elemente enthalten ausschliesslich Datenwerte.
- Alle Elemente einer Zeile gehören zusammen.

Exkurs: Daten in Tabellen

Tabellenkalkulationsblatt



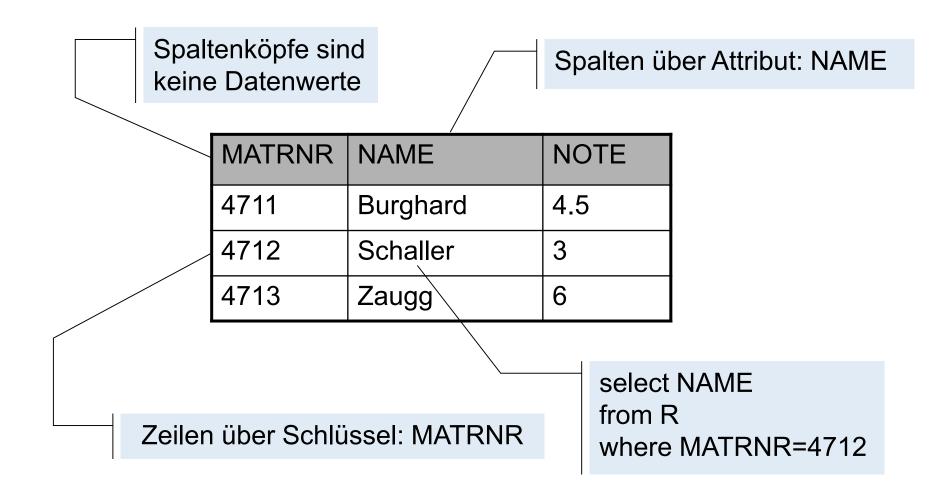


Exkurs: Daten in Tabellen

u^{b}

UNIVERSITÄT BERN

Relation



Zwischenfazit



- Im Relationalen Datenmodell werden die Daten in Tabellenform abgebildet.
- Dafür muss vorgängig ein Datenschema mit den entsprechenden Attributen angelegt werden, welche die Spalten der Tabelle vorgeben.
- Eine Zeile der Datentabelle bildet ein Tupel zusammengehöriger Attributwerte.
- Datenwerte in einer Relation (Tabelle) müssen elementar sein.
- Dies erzwingt unter Umständen die mehrfache Abbildung bestimmter Datenwerte.
- Die Tabellenstruktur der Relationen ähnelt den (Daten-)Tabellen in Tabellenkalkulationsprogrammen.
- Die Manipulation der Daten in einem Relationalen Datenbanksystem und einem Tabellenkalkulationsprogramm unterscheidet sich jedoch grundlegend.

Gliederung





Datenbanksprache SQL



- Die wichtigste Datenbanksprache für relationale Datenbanken.
- Die Sprache basiert auf der relationalen Algebra.
- Wird durch ein gemeinsames Gremium von ISO und IEC standardisiert.
- Durch SQL wird die Unabhängigkeit vom eingesetzten Datenbankmanagementsystem angestrebt.
- Fast alle gängigen Datenbanksysteme unterstützen grundsätzlich SQL.
- Im Umfang und in bestimmten sprachlichen Besonderheiten treten Unterschiede auf.



Datenmanipulationen

 u^{t}

UNIVERSITÄ BERN

Definition von Datenstrukturen

Anlegen neuer Relationen: CREATE TABLE

- ...

Veränderungen von Daten:

– Einfügen von Daten(-sätzen): INSERT

Überschreiben von Daten: UPDATE

Löschen von Daten(-sätzen): DELETE

Abfragen von Daten (SELECT):

Sortieren von Daten

Anzeigen bestimmter Attribute (Spalten): Projektion

Anzeigen bestimmter Tupel (Zeilen): Selektion

Anzeigen von Daten aus verschiedenen Relationen: Verbund

— ...



u^{ϵ}

UNIVERSITÄT BERN

Einfügen von Datensätzen

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6



insert into R values ("4714", "Zampano",5.5)

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6
4714	Zampano	5.5

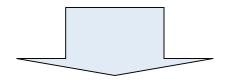
Operationen über einer Relation

$u^{^{\mathsf{b}}}$

UNIVERSITÄT BERN

Verändern von Datensätzen

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6



update R set NOTE=4 where MATRNR=4712

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	4
4713	Zaugg	6

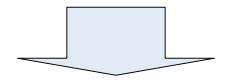
Operationen über einer Relation

u^{ϵ}

UNIVERSITÄT BERN

Löschen von Datensätzen

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6



delete from R where MATRNR="4713"

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3

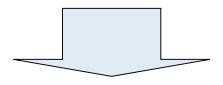
Operationen über einer Relation:

u^{b}

UNIVERSITÄT BERN

Sortieren

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6



select * from R order by NOTE desc

MATRNR	NAME	NOTE
4713	Zaugg	6
4711	Burghard	4.5
4712	Schaller	3

Abfragen einer Relation

u^{b}

UNIVERSITÄT BERN

Projektion

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6



select MATRNR,NOTE from R

MATRNR	NOTE
4711	4.5
4712	3
4713	6

Abfragen einer Relation

Selektion



UNIVERSITÄT BERN

MATRNR	NAME	NOTE
4711	Burghard	4.5
4712	Schaller	3
4713	Zaugg	6



select * from R where NOTE>=4

MATRNR	NAME	NOTE
4711	Burghard	4.5
4713	Zaugg	6





UNIVERSITÄT BERN

Verbundoperation (Join)

 Durch eine Verbundoperation lassen sich die Inhalte aus mehreren Relationen in eine Relation zusammenführen.

- Dies geschieht typischerweise, in dem zwischen Attributen mit gleichen Wertedomänen ein Vergleich angestellt wird.
- Der übliche Vergleich ist der auf Gleichheit (Equi-Join).

 Bei Wertegleichheit werden die Tupel aus den verschiedenen Relationen zu einem Tupel in der Ergebnisrelation zusammengeführt.



Vorlesung	Matrikelnr	Note
W300	4711	5
W300	4712	5
W301	4711	3
W301	4712	3
W302	4712	3

4713

W302

Matrikelnr	Name
4711	Α
4712	С
4713	Α

select * from R1, R2 where R1.MATRIKELNR= R2.MATRIKELNR



R1.Vorlesung	R1.Matrikelnr	R1.Note	R2.Matrikelnr	R2.Name
W300	4711	5	4711	Α
W300	4712	5	4712	С
W301	4711	3	4711	Α
W301	4712	3	4712	С
W302	4712	3	4712	С
W302	4713	3	4713	А

3



Abfragen einer Relation: Verbund (2)



UNIVERSITÄ[.] BERN

select R1.VORLESUNG, R1.MATRIKELNR, R2.NAME, R1.NOTE from R1, R2 where R1.MATRIKELNR=R2.MATRIKELNR



R1.Vorlesung	R1.MatrikeInr	R2.Name	R1.Note
W300	4711	Α	5
W300	4712	С	5
W301	4711	Α	3
W301	4712	С	3
W302	4712	С	3
W302	4713	Α	3

Fazit



- Relationale Datenbanksysteme strukturieren Daten in Form von Tabellen.
- Daten müssen in Form elementarer Attributwerte abgelegt werden.
- Daten in Tabellen k\u00f6nnen anhand einfacher Datenbankoperationen manipuliert werden.
- Die Operationen orientieren sich an mengentheoretischen Konzepten.
- Um nur bestimmte Tupel in einer Tabelle anzusprechen, muss eine Selektion auf bestimmte Attributwerte vorgenommen werden (Selektion).
- Um nur bestimmte Attribute zu sehen, müssen diese jeweils explizit angegeben werden (Projektion).
- Über Verbundoperationen ist es möglich, Daten aus verschiedenen Tabellen in einer Ergebnistabelle zusammenzuführen (Join).