



Foto: Thomas Josek

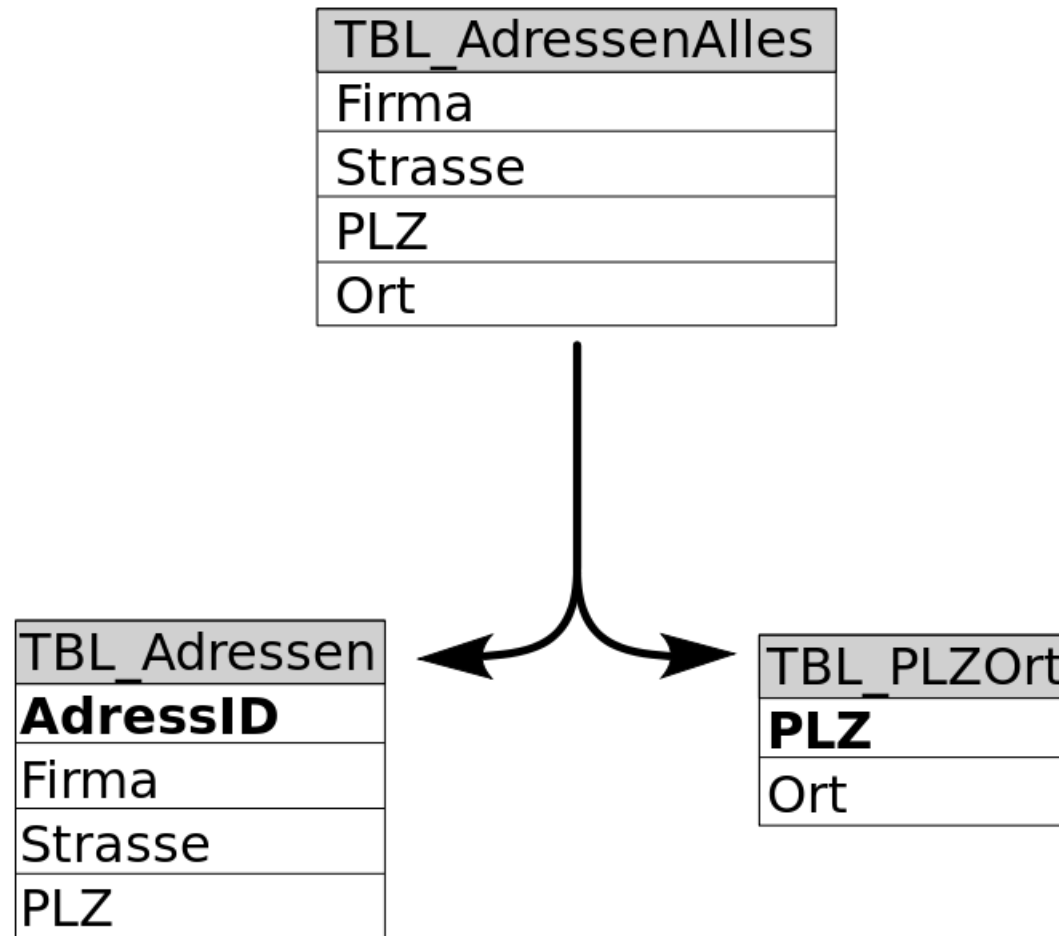
Datenbanktechnologie

SoSe 2022, Woche 4: 2. Mai.

Normalisierung von Datenbanken

- Reduktion von Redundanz
 - Ziel: keine Redundanz
- Änderungen der Tabellenstruktur
 - Relationen zwischen Attributen (Tabellenspalten)
- Konsistenz
 - Anomalien zu vermeiden
 - teilweise Änderungen
 - unterschiedliche Werte für das gleiche Phänomen
 - obsolete Daten
 - widersprechende Daten
 - unnötige Verwendung von Speicherplatz
- In der Phase des Entwurfs durchgeführt

Beispiel



[https://commons.wikimedia.org/wiki/File:Aufspaltung_einer_Tabelle_in_zwei_\(Beispiel\).svg](https://commons.wikimedia.org/wiki/File:Aufspaltung_einer_Tabelle_in_zwei_(Beispiel).svg)

Euku:⇌, Public domain, via Wikimedia Commons

Konsistenz

Zu vermeiden:

- Anomalien
- Teilweise Änderungen
- Unterschiedliche Werte für das gleiche Phänomen
- Obsolete Daten
- Widersprechende Daten
- Unnötige Verwendung von Speicherplatz
- Aktualisierung der Datenbank zu vereinfachen
 - Änderungen lediglich an einer Stelle

Denormalisierung

Sonderfall

- Performance
- Anfragen zu vereinfachen

Problemverringering

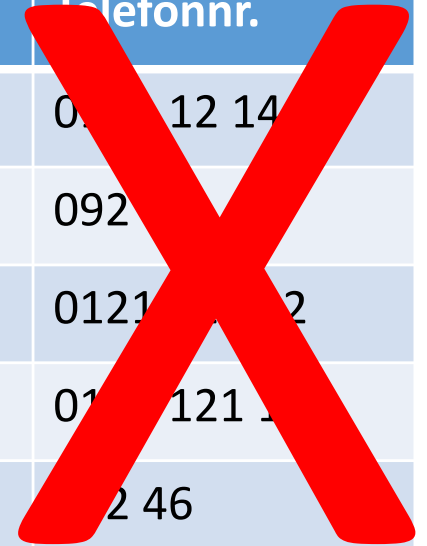
- Automatische Abgleichroutinen
- Read-only
- Exports

Erinnerung: Tiere

ID	Tier	Name	Besitzer	Telefonnr.
1	Fuchs	Foxie	Petra Knoch	0171 12 14
2	Lachs	Salmie	Øyvind Eide	092 46
3	Katze	Mausie	Anna Kurz	0121 121 12

Erinnerung: Tiere

ID	Tier	Name	Besitzer	Telefonnr.
1	Fuchs	Foxie	Petra Knoch	0121 12 14
2	Lachs	Salmie	Øyvind Eide	092
3	Katze	Mausie	Anna Kurz	0121 12 12
4	Rentier	Muzte	Anna Kurz	0121 121 1
5	Maus	Kattie	Øyvind Eide	0121 12 46



Erinnerung: Tiere mit Besitzer

TierID	Tier	Name	BesitzerID
1	Fuchs	Foxie	1
2	Lachs	Salmie	2
3	Katze	Mausie	3
4	Rentier	Muzte	3
5	Maus	Kattie	2

PersonID	Name	Telefonnr.
1	Petra Knoch	0171 12 14
2	Øyvind Eide	075 43 25
3	Anna Kurz	0121 121 12

Grundbegriff: Relation

- Eine Relation besteht aus Tupeln
 - Tupel: Liste von Objekten
 - n-Tupel: (X_1, \dots, X_n)
- Tupeln
 - durch Attribute beschrieben
 - Typ (mögliche Werte) festgelegt
 - Mit einem Namen beschrieben

Grundbegriff: Relation

- Relationen sind als Tabelle ausgedrückt
- Entitäten und Beziehungen
 - Entities and Relationships – ER-Modelle
- Entitäten *und* Beziehungen sind Relationen
- Formal: mathematische Definition
 - Basis der relationalen Algebra (Edgar F. Codd)

Grundbegriff: Funktionale Abhängigkeit

- Eine Relation wird durch Attribute definiert
- Funktionale Abhängigkeit:
 - einige Attribute bestimmen eindeutig die Werte anderer Attribute
- Beispiel: Kundendatenbank
 - die Anschrift und die Telefonnummer eines Kunden sind eindeutig durch seinen Namen zusammen mit seinem Geburtstag bestimmt
 - Anschrift und Telefonnummer sind funktional abhängig von Name und Geburtsdatum

Grundbegriff: Transitive Abhängigkeit

- Transitivität:
 - Eine Relation R : wenn xRy und yRz folgt xRz
 - Beispiel: wenn $x < y$ und $y < z$ folgt $x < z$
 - $3 < 5$ und $5 < 7$ und $3 < 7$
- Transitive Abhängigkeit:
 - Attribut A_2 ist von Schlüsselkandidat P_1 transitiv abhängig wenn:
 - es eine Attributmengende A_1 gibt, sodass $(P_1 \rightarrow A_1)$ und $(A_1 \rightarrow A_2)$

Normalformen

- Für die Normalisierung der Datenbanken
- Fünf Normalformen
- Bestimmte formale Anforderungen
 - an die Tabellenstruktur
 - an die Relationen
- Jede enthält die Kriterien der vorangehenden

Weiterlesen

Es gibt eine Menge Bücher, dieses ist als Hypertext zugänglich:

Relationale Datenbanken : Eine Einführung / Andreas Kelz. <https://www.hdm-stuttgart.de/~riekert/lehre/db-kelz/>

Kap 4: *Der Königsweg: Normalisierung*

Es gibt viele Lehrbücher und Tutorials online, auch auf Deutsch. Wikipedia ist auch für ein Grundverständnis OK.

Erste Normalform

- *Jeder Attributwert ist atomar*
 - Jedes Attribut der Relation muss einen atomaren Wertebereich haben, und die Relation muss frei von Wiederholungsgruppen sein
- Atomar: Wertebereiche nicht
 - zusammengesetzt
 - mengenwertig
 - geschachtelt
- Kein Attribut kann in weitere sinnvolle Teilbereiche aufgespaltet sein

Erste Normalform

ID	VerrasserIn	Exemplar1	Exemplar2	Exemplar3	
2345	Alice Walker	Zugvogel	Regal 20-4	Regal 20-5	Regal 20-6

ID	VerfasserIn	Titel	Exemplare
2345	Alice Walker	Zugvögel	Regal 20-4, Regal 20-5, Regal 20-6

BuchID	VerfasserIn	Titel	ExemplarID	Exemplar
2345	Alice Walker	Zugvögel	1367	Regal 20-4
2345	Alice Walker	Zugvögel	1368	Regal 20-5
2345	Alice Walker	Zugvögel	1423	Regal 20-6

Zweite Normalform

- *Jedes Nicht-Schlüsselattribut ist von jedem Schlüsselkandidaten vollständig funktional abhängig*
- Monothematische Relationen
 - jede Relation modelliert nur einen Sachverhalt
- Reduziert
 - Redundanz
 - Inkonsistenz-Gefahr
- Klare Struktur

Zweite Normalform

BuchID	VerfasserIn	Titel	ExemplarID	Exemplar
2345	Alice Walker	Zugvögel	1367	Regal 20-4
2345	Alice Walker	Zugvögel	1368	Regal 20-5
2345	Alice Walker	Zugvögel	1423	Regal 20-6

ExemplarID	BuchID	Exemplare
1367	2345	Regal 20-4
1368	2345	Regal 20-5
1423	2345	Regal 20-6

BuchID	VerfasserIn	Titel
2345	Alice Walker	Zugvögel

Dritte Normalform

- *Jedes Nicht-Schlüssel-Attribut ist von keinem Schlüsselkandidaten transitiv abhängig*
- Ein Nichtschlüsselattribut darf nicht von einer Menge aus Nichtschlüsselattributen abhängig sein.
- Ein Nichtschlüsselattribut darf also nur direkt von einem Primärschlüssel (bzw. einem Schlüsselkandidaten) abhängig sein.

Tierfutter-Rechnungen

Rechnungsnummer	Datum	Name	Vorname	Straße	PLZ	Ort
1	2022-05-07	Eide	Øyvind	Unistr. 1	50949	Köln
2	2022-02-25	Kurz	Anna	Eifelstr. 4	50234	Köln

Rechnungsnummer	Datum	Personalnummer
1	2022-05-07	25
2	2022-02-25	95

PLZ	Ort
50949	Köln
50234	Köln

Personalnummer	Name	Vorname	Straße	PLZ
25	Eide	Øyvind	Unistr. 1	50949
95	Kurz	Anna	Eifelstr. 4	50234

Boyce-Codd-Normalform (BCNF)

- *Eine Relation ist in Boyce-Codd-Normalform, wenn jeder Determinant ein Schlüsselkandidat ist*
- Weiterentwicklung der Dritten Normalform
- Eine Determinante ist eine Attributmengende, von der ein anderes Attribut vollständig funktional abhängig ist

Tierfutter-Rechnungen

Rechnungsnummer	Kostenart	Kostenartnummer	Anzahl
1	Kibble Economy	45	12
2	Kibble Premium	76	5

Rechnungsnummer	Kostenart	Anzahl
1	Kibble Economy	12
2	Kibble Premium	5

Kostenart	Kostenartnummer
Kibble Economy	45
Kibble Premium	76

3 (+1) sind ausreichend... für heute

- Wenn Sie mit Vierter und Fünfter Normalform arbeiten würden, wären Sie tief in Datenbankentwicklung
- Sie müssen dafür diese nicht lernen oder gut verstehen
- Für Vollständigkeit gehen wir aber auch diese kurz durch

Vierte Normalform

- Für jede mehrwertige Abhängigkeit gilt einer Attributmengende Y von einer Attributmengende X :
 - die mehrwertige Abhängigkeit trivial ist oder
 - X ist ein Schlüsselkandidat der Relation
- Eine mehrwertige Abhängigkeit einer Attributmengende Y von einer Attributmengende X ist trivial, wenn Y Teil von X ist oder die Relation nur aus X und Y besteht

Personalnummer	Haustier
1	Katze
2	Hund

Personalnummer	Fahrzeugtyp
1	Toyota
2	Ford

Personalnummer	Haustier	Fahrzeugtyp
1	Katze	Toyota
2	Hund	Ford

Fünfte Normalform

- Für jede Join-Abhängigkeit (R_1, R_2, \dots, R_n) gilt:
 - die Join-Abhängigkeit ist trivial oder
 - jedes R_i aus (R_1, R_2, \dots, R_n) ist Schlüsselkandidat der Relation
- Ähnlich mit Vierter Normalform, geht aber um Joins

Arbeit für nächste Woche

- Machen Sie Ihre erste Datenbankstruktur
- Denken Sie nach, wie sie mit Erster und Zweiter Normalform konsistent sein könnte
- Thema ist frei – gern können Sie Haustiere verwenden
- Erst als Text und Figur, gern können Sie eine Implementation in MySQL versuchen

- Installation von phpMyAdmin zu versuchen
 - Links kommen auf der Webseite