

Vortrag 1 Entwurf

Datenbanken

Lukas Wais

Codersbay

Version: 19. April 2023

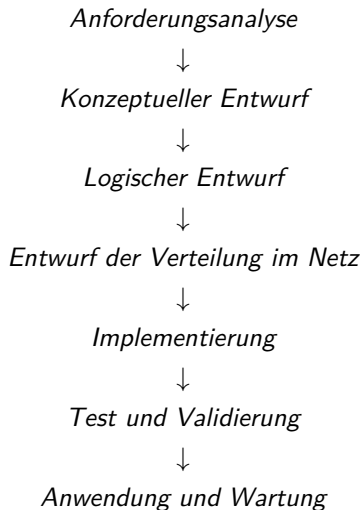
Inhaltsverzeichnis

Datenbank Entwurf

Entity Relationship Diagramme

Normalformen

Datenbank-Lebenszyklus



- ▶ **Anforderungsanalyse:** Die Anforderungen der Benutzer and die neue Datenbank werden gesammelt. Wichtig ist zu wissen welche zu speichern sind und wie die Daten zu bearbeiten sind.
- ▶ **Konzeptueller Entwurf:** Hier werden die Sichten und das konzeptuelle Gesamtschema definiert. Meistens wird zur Darstellung des Gesamtschemas ein Entity-Relationship-Diagramm verwendet. Weiteres wird auch definiert welches DBMS zum Einsatz kommt, um die Datenbank korrekt aufbauen zu können.
- ▶ **Logischer Entwurf:** Jetzt wird die der konzeptuelle Entwurf umgesetzt. Das Datenbankschema wird Normalisiert um z.B. Redundanzen zu beseitigen.
- ▶ **Entwurf der Verteilung im Netz:** Kommt nur bei verteilten Datenbanken vor.

Erklärung II

- ▶ **Implementierung:** Hier wird die Datenbank erstellt und die Abfragen formuliert. Sollte von einem alten System umgestiegen werden, dann folgt hier auch die Datenkonvertierung.
- ▶ **Test und Validierung:** Die DB und erstellte Abfragen werden getestet und auf die Anforderungen geprüft (validiert).
- ▶ **Anwendung und Wartung:** Ständige Wartung, es kann immer wieder zu neuen Anforderung kommen, die oft auch das Datenbankschema betreffen.

Definition Primärschlüssel

Primärschlüssel

Ein Primärschlüssel ist eine Spalte oder eine Spaltengruppe in einer Tabelle, deren Werte eine Zeile in der Tabelle eindeutig kennzeichnen. Eine relationale Datenbank ist darauf ausgelegt, die Eindeutigkeit von Primärschlüsseln zu erzwingen, indem nur eine einzelne Zeile mit einem bestimmten Primärschlüsselwert in einer Tabelle zugelassen wird.

Quelle: ibm.com aufgerufen am 12.10.2022

Definition Fremdschlüssel

Fremdschlüssel

Ein Fremdschlüssel ist eine Spalte oder eine Spaltengruppe in einer Tabelle, deren Werte den Werten des Primärschlüssels in einer anderen Tabelle entsprechen. Um eine Zeile mit einem bestimmten Fremdschlüsselwert hinzufügen zu können, muss in der zugehörigen Tabelle eine Zeile mit demselben Primärschlüsselwert vorhanden sein.

Quelle: ibm.com aufgerufen am 12.10.2022

Das Entity-Relationship-Modell

Das ER-Modell ist eines der beliebtesten grafischen Hilfsmittel, um Ausschnitte der realen Welt modellieren zu können. Es ist unabhängig von einem bestimmten Datenmodell und unterliegt nicht deren Einschränkungen. Üblich wird auf die Chen-Notation gesetzt. Das Entitäten-Beziehungs-Modell ermöglicht es, die konzeptuellen Modelle von Datenbanken auf einfache Art grafisch darzustellen und erlaubt die Anwendung von Abstraktionskonzepten.

Die beiden Hauptbausteine stecken schon im Namen, nämlich **Entitäten** und **Beziehungen**.

Entitäten

- ▶ Als Entität werden unterscheidbare Dinge aus der realen Welt bezeichnet. Wie etwa Personen, oder Unternehmen.
- ▶ Eine Entitätsmenge bezeichnet eine Sammlung gleichartiger Entitäten.
- ▶ Weiters gibt es auch einen Entity-Typ der gleichartige Entitäten kategorisiert. In der mathematischen Mengenschreibweise schreibt man für eine Entität e eines Entity-Typs E $e \in E$.

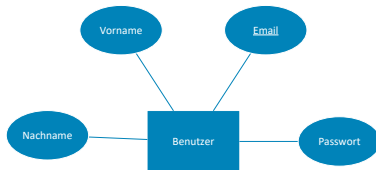
In der folgenden Folie werden die unterschiedlichen Entitäten gezeigt.



Eine Entität, die nicht allein durch ihre Attribute eindeutig identifiziert werden kann. Die Existenz einer schwachen Entität ist abhängig von einer anderen Entität, der Eigentümer-Entität. Der Identifikator der schwachen Entität ist eine Kombination aus dem Identifikator der Eigentümerentität und dem Teilschlüssel der schwachen Entität.



Eine Entität, die in einer Many-to-many-Beziehung verwendet wird (stellt eine zusätzliche Tabelle dar). Alle Beziehungen für die assoziative Entität sollten viele sein



Beziehungen

Beziehungen drücken Abhängigkeiten der Entitäten aus.

Es gibt folgende Beziehungsgrade:

- ▶ Binär; genau zwei Entitäten sind miteinander verbunden
- ▶ Ternär; drei Entities sind verbunden
- ▶ n-är; n (mehrere) Entitäten sind verbunden
- ▶ Rekursiv unär; eine Entität steht mit sich selbst in Beziehung

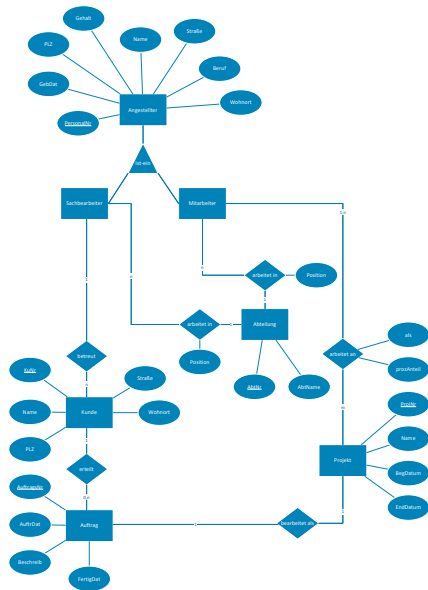
Kardinalitäten

Die Kardinalität legt fest wie viele Entitäten einer Entitätsmenge in Beziehung stehen können. Ein Beispiel wäre: Wie viele Mitarbeiter können an einem Projekt mitarbeiten?

Folgende Beziehungen sind möglich:

- ▶ 1 : 1 eins-zu-eins-Beziehung
- ▶ 1 : n eins-zu- n -Beziehung
- ▶ n : m n -zu- m -Beziehung
- ▶ Es kann auch eine Zahl festgelegt wie 2 :
- ▶ Bereiche sind auch möglich 0, n gibt an, dass 0, 1, 2, ..., n Entitäten infrage kommen. Ist 0 keine Möglichkeit so wird sie einfach weggelassen.

Auf der folgenden Folie ist ein Beispiel eines kompletten ER-Diagramms



Textbeschreibung Entity-Typen

ANGESTELLTER(PersonalNr, Name, GebDat, PLZ, Wohnort, Strasse, Gehalt, Beruf)

MITARBEITER(PersonalNr, Name, GebDat, PLZ, Wohnort, Strasse, Gehalt, Beruf)*

SACHBEARBEITER(PersonalNr, Name, GebDat, PLZ, Wohnort, Strasse, Gehalt, Beruf)*

ABTEILUNG(AbtNr, AbtName)

KUNDE(KuNr, Name, PLZ, Wohnort, Strasse)

AUFTRAG(AuftragsNr, AuftrDat, Beschreib, FertigDat)

PROJEKT(ProjNr, Name, BegDatum, EndDatum)

*Die Attribute erben diese Entity-Typen von dem Entity-Typ ANGESTELLTER.

Textbeschreibung Beziehungstypen

is-a(ANGESTELLTER, MITARBEITER, SACHBEARBEITER)

arbeitet_in(ABTEILUNG, SACHBEARBEITER, Position)

arbeitet_in(ABTEILUNG, MITARBEITER, Position)

betreut(SACHBEARBEITER, KUNDE)

erteilt(KUNDE, AUFTRAG)

bearbeitet_als(AUFTRAG, PROJEKT)

arbeitet_an(MITARBEITER, PROJEKT, als, prozAnteil)

betreut(ABTEILUNG, PROJEKT)

Warum Normalisierung des Datenbankschemas

- ▶ Anomalien beheben
- ▶ Redundanzen vermeiden
- ▶ Möglichst einfacher und übersichtlicher Aufbau der Relationen

- ▶ **Einfüge-Anomalie:** Probleme beim einfügen eines neuen Tupels, weil Datenfelder leer sind.
- ▶ **Lösch-Anomalie:** Beim löschen gehen mehr Daten verloren als gewünscht.
- ▶ **Änderungs-Anomalie:** Inkonsistente Relationen beim Updaten. Bsp. änder ein Mitarbeiter seinen Nachnamen muss er überall geändert werden, geschieht dies nicht, dann liegt eine Änderungs-Anomalie vor.

Normalformen

1. Normalform: Alle Attribute enthalten atomare Inhalte, und die Relation hat eine feste Breite.
 2. Normalform: Die 1NF muss vorliegen und jedes nicht-primär-Attribut ist von allen ganzen Schlüsseln abhängig.
 3. Normalform: Die dritte Normalform ist erreicht, wenn sich das Relationenschema in 2NF befindet, und kein Nichtschlüsselattribut von einem anderen Nichtschlüsselattribut funktional abhängig ist.
- Es gibt noch weitere Normalformen, diese finden in der Wissenschaft Bedeutung, kommen in der Praxis aber kaum vor.

0. Normalform

CD_ID	Album	Titelliste	Gründungsjahr
4711	Anastacia - Not That Kind	{1. Not That Kind, 2. I'm Outta Love, 3. Cowboys & Kisses}	1999
4712	Pink Floyd - Wish You Were Here	{1. Shine On You Crazy Diamond}	1965
4713	Anastacia - Freak of Nature	{1. Paid my Dues}	1999

1. Normalform

Jedes Attribut trennen, bis es untrennbar ist.

CD_ID	Album	Interpret	Titel	Track	Gründungsjahr
4711	Not That Kind	Anastacia	Not That Kind	1	1999
4711	Not That Kind	Anastacia	I'm Outta Love	2	1999
4711	Not That Kind	Anastacia	Cowboys & Kisses	3	1999
4712	Wish You Were Here	Pink Floyd	Shine On You Crazy Diamond	1	1965
4713	Freak of Nature	Anastacia	Paid my Dues	1	1999

2. Normalform

Man trennt die Relation in mehrere, in unserem Fall in zwei Relationen auf, sodass jedes Nichtschlüsselattribut voll funktional von seinem Schlüssel abhängig ist.

CD_ID	Album	Interpret	Gründungsjahr
4711	Not That Kind	Anastacia	1999
4712	Wish You Were Here	Pink Floyd	1965
4713	Freak of Nature	Anastacia	1999

CD_ID	Track	Titel
4711	1	Not That Kind
4711	2	I'm Outta Love
4711	3	Cowboys & Kisses
4712	1	Shine On You Crazy Diamond
4713	1	Paid my Dues

3. Normalform

Die Relation CD wird in in zwei weitere Relationen aufgeteilt.

CD_ID	Track	Titel
4711	1	Not That Kind
4711	2	I'm Outta Love
4711	3	Cowboys & Kisses
4712	1	Shine On You Crazy Diamond
4713	1	Paid my Dues

CD_ID	Album
4711	Not That Kind
4712	Wish You Were Here
4713	Freak of Nature

CD_ID	Interpret	Gründungsjahr
4711	Anastacia	1999
4712	Pink Floyd	1965