INFORMATIONSTECHNIKEN

HOCHSCHULE DER MEDIEN

DATENBANKEN

Prof. Dr. Wolf-Fritz Riekert
Hochschule der Medien (HdM) Stuttgart
University of Applied Sciences

mailto:riekert@hdm-stuttgart.de
http://v.hdm-stuttgart.de/~riekert

INHALTSÜBERSICHT

TEIL1: EINFÜHRUNG INHALTSÜBERSICHT

DATEN VERSUS INFORMATIONEN

DATEN – INFORMATIONEN

HOCHSCHULE DER MEDIEN

- **Daten** sind eine Repräsentation von Fakten, Konzepten oder auch Instruktionen in einer formalisierten Art und Weise, die sie für die Kommunikation, Interpretation und Verarbeitung durch **Maschinen** geeignet machen.
- **Informationen** beziehen sich auf die Bedeutung, die **Menschen** durch vereinbarte Konventionen diesen Daten zuordnen. (Beide Zitate: Stallings 1994)
- Informationen sind interpretierte Daten (Neuhold 1977)
- Nutzungsaspekt der Information:
  - ⇨ Informationen bilden Grundlage für menschliche **Entscheidungen**
  - ⇨ Informationen bringen **Nutzen**, besitzen infolgedessen einen monetären Wert

DATA, INFORMATION, KNOWLEDGE

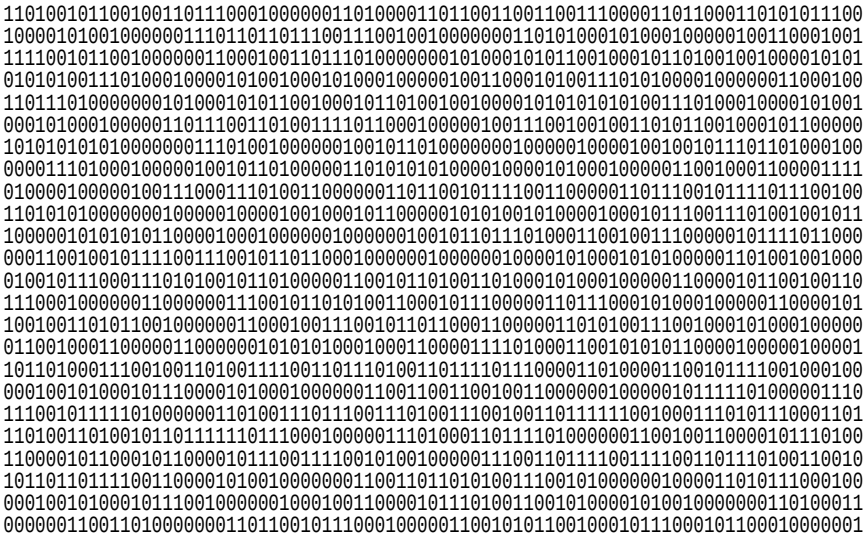
HOCHSCHULE DER MEDIEN

Data	recorded facts, typically maintainable by clerks
Information	data useful to a customer
Knowledge	meta-data, relationships among terms, processing paradigms, ..., needed to transform data to information

Aus: Wiederhold, G.: I³ Vocabulary. In: Hull, R. / King R. (Eds.): Reference Architecture for the Intelligent Integration of Information. 1995. Appendix B. URL: http://infolab.stanford.edu/pub/gio/1994/vocabulary.html.

DATEN: BINÄRE ZUSTÄNDE EINES SPEICHERMEDIUMS

HOCHSCHULE DER MEDIEN



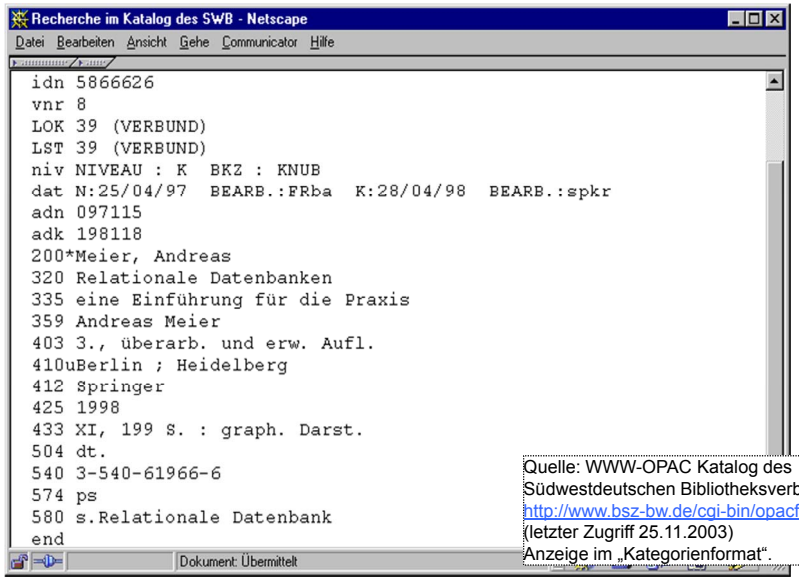
DATEN: ZAHLENWERTE AUF SPEICHERPLÄTZEN

HOCHSCHULE DER MEDIEN

```
0000000: 064544 067040 032066 031470 033065 034412 020166 067162
0000020: 020065 005040 046117 045440 030467 020050 053105 051102
0000040: 052516 042051 005040 046123 052040 030467 020050 053105
0000060: 051102 052516 042051 005040 067151 073040 047111 053105
0000100: 040525 020072 020113 020040 041113 055040 035040 045501
0000120: 052502 005040 062141 072040 047072 030066 027460 033457
0000140: 034465 020040 041105 040522 041056 035113 040525 041040
0000160: 020113 035062 034057 030062 027471 033040 020102 042501
0000200: 051102 027072 045501 045515 005040 060544 067040 030071
0000220: 032461 034067 005040 060544 065440 030471 033060 032471
0000240: 005040 031060 030052 042141 072145 026040 041550 071151
0000260: 071564 067560 064145 071040 045056 005040 031462 030040
0000300: 057501 067137 020151 067164 071157 062165 061564 064557
0000320: 067040 072157 020144 060564 060542 060563 062440 071571
0000340: 071564 062555 071412 020063 032471 020103 027040 045056
0000360: 020104 060564 062412 020064 030063 020066 027040 062544
0000400: 027054 020162 062560 071056 020167 064564 064040 061557
```

ZAHLEN UND ZEICHENFOLGEN: DATEN ODER INFORMATION?

HOCHSCHULE DER MEDIEN



Quelle: WWW-OPAC Katalog des Südwestdeutschen Bibliotheksverbundes
http://www.bsz-bw.de/cgi-bin/opacform.cgi
(letzter Zugriff 25.11.2003)
Anzeige im „Kategorienformat“.

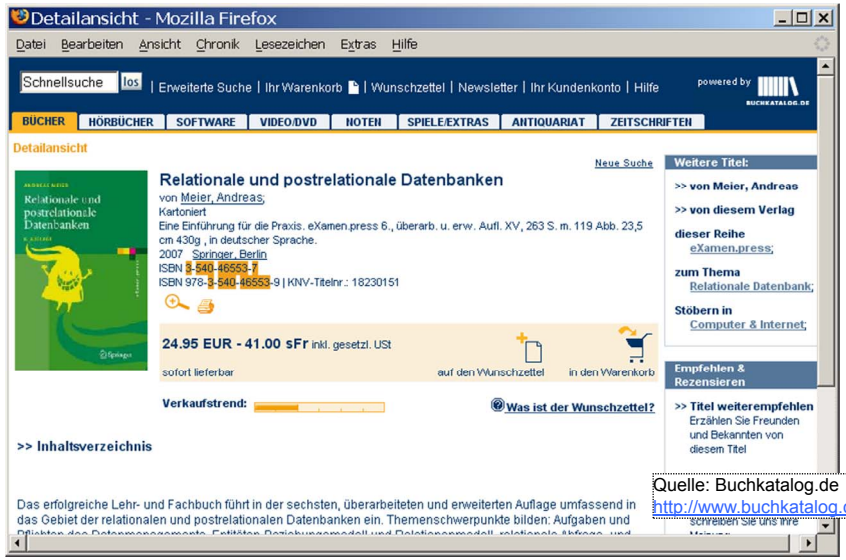
INFORMATIONEN: ZAHLEN UND ZEICHENFOLGEN MIT BEDEUTUNG

HOCHSCHULE DER MEDIEN

Quelle: WWW-OPAC Katalog des Südwestdeutschen Bibliotheksverbundes
http://www.bsz-bw.de/cgi-bin/opacform.cgi
(letzter Zugriff 25.11.2003)

INFORMATIONEN: GRAFIK UND TEXTE

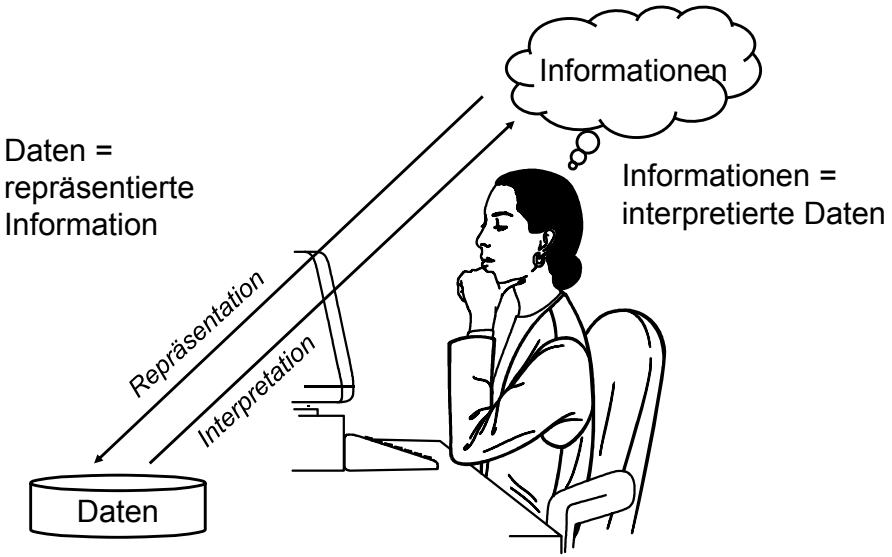
HOCHSCHULE DER MEDIEN



Quelle: Buchkatalog.de
http://www.buchkatalog.de

INTERPRETATION UND REPRÄSENTATION

HOCHSCHULE DER MEDIEN



TRANSIENTE DATEN – FLÜCHTIGE DATEN

HOCHSCHULE DER MEDIEN

- Transiente Daten befinden sich auf einem flüchtigen Speichermedium eines Computers, i.d.R. im Arbeitsspeicher.
- Ein Prozess (ein ablaufendes Programm) belegt Speicherplätze im Arbeitspeicher, legt dort transiente Daten an und weist diesen Werte zu.
- Die Speicherplätze werden spätestens dann freigegeben, wenn der Prozess sich beendet. Die transienten Daten hören dann auf, zu existieren.
- Transiente Daten gehören dem Prozess, der sie erzeugt hat. Andere Prozesse können auf sie nicht zugreifen.
- Beim Neustart eines Programms kann nicht auf die Werte der Daten früherer Programmabläufe zurückgegriffen werden.

PERSISTENTE DATEN (DAUERHAFTE DATEN)

HOCHSCHULE DER MEDIEN

- Persistente Daten befinden sich auf einem nichtflüchtigen Speichermedium eines Computers, i.d.R. auf der Festplatte.
- Wichtigstes Beispiel für persistente Daten: Dateien.
- Persistente Daten können durch Prozesse erzeugt, verändert und gelöscht werden.
- Sie existieren weiter, wenn der Prozess sich beendet, der sie erzeugt hat.
- Sie hören erst auf, zu existieren, wenn ein Prozess sie explizit löscht.
- Andere Prozesse können (prinzipiell) auf die persistenten Daten zugreifen.
- Beim Neustart eines Programms kann auf die persistenten Daten zurückgegriffen werden, die durch früherere Programmabläufe erzeugt wurden.

WAS IST EINE DATENBANK

HOCHSCHULE DER MEDIEN

Erste Definition:

Datenbank = Verwaltungskomponente +
Speicherungskomponente
für persistente Daten,
die einem bestimmten Zweck dienen

Frage:

Warum genügt hierfür nicht das Dateisystem eines Computers?

WAS IST EIN DATENBANKSYSTEM?

HOCHSCHULE DER MEDIEN

Ein Datenbanksystem (auch Datenbankmanagementsystem) ist ein Softwaresystem, das den Aufbau und den Betrieb einer Datenbank unterstützt.

Beispiele für verbreitete Datenbanksysteme:
Access, SQL Server, MySQL, DB2, Oracle, Sybase usw.

Beispiel zum Unterschied Datenbank - Datenbanksystem:

- Eine **Datenbank** mit dem Namen „Adressen“ dient zur Verwaltung von Postadressen und Telefonnummern.
- Die Datenbank „Adressen“ wurde/wird mit Hilfe des **Datenbanksystems** „Access“ aufgebaut und betrieben.

Abkürzungen:
DB = Datenbank, DBMS = Datenbank(management)system

FUNKTIONEN EINES DATENBANKSYSTEMS

HOCHSCHULE DER MEDIEN

- Dauerhafte Speicherung von Daten
- Einrichten einer Datenbank mit Hilfe einer Datendefinitionssprache
  - ⇨ Beispiel: Einrichten einer Adressdatenbank
- Abfragen und Modifikation von Daten mit Hilfe einer Datenmanipulationssprache: Selektieren, Anzeigen, Erzeugen, Verändern, Verknüpfen, Löschen von Daten
  - ⇨ Beispiel: Nachschlagen, Eintragen, Ändern, Löschen von Adressen in einer Adressdatenbank
- Interaktive Benutzungsoberfläche für Datendefinition und Datenmanipulation (ohne Programmierkenntnisse)
- Programmschnittstelle, ermöglicht Datendefinition und Datenmanipulation aus einem Programm heraus

BESONDERE SYSTEMMERKMALE EINES DATENBANKSYSTEMS

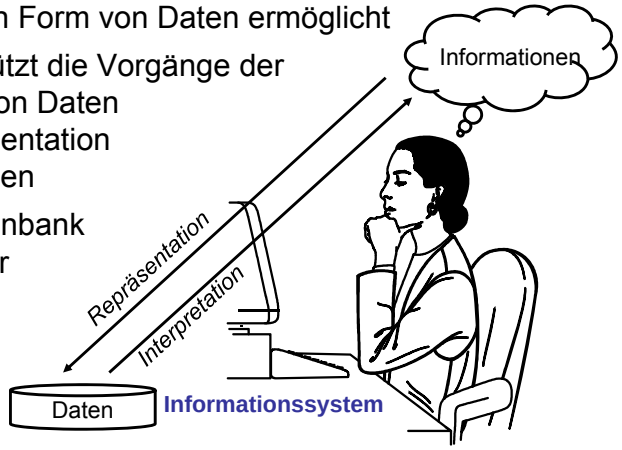
HOCHSCHULE DER MEDIEN

- **Datenunabhängigkeit**: Das Datenbanksystem macht die Nutzer unabhängig von den computerinternen Speicherstrukturen zur Datenablage
  - ⇨ Daten können über genormte Programmschnittstelle von verschiedenen Programmen genutzt werden
- **Mehrbenutzerfähigkeit**: Störungsfreier paralleler Zugriff auf eine Datenbank durch mehrere Prozesse (auch über Netze) mit Hilfe eines sogenannten Transaktionskonzepts
- **Konsistenzerhaltung** auch bei Systemabsturz
  - ⇨ Beispiel: Überweisung eines Betrags von einem Konto auf ein anderes Konto. Es darf nicht passieren, dass nur ein Konto verändert wird und das andere nicht.
- Fortlaufende **Datensicherung** und **Wiederanlauf** nach Systemabsturz

WAS IST EIN INFORMATIONSSYSTEM?

HOCHSCHULE DER MEDIEN

- ein Anwendungsprogramm, das die Bereitstellung von Informationen aus Daten und die Ablage von Informationen in Form von Daten ermöglicht
- D.h. es unterstützt die Vorgänge der Interpretation von Daten und der Repräsentation von Informationen
- Kurz: eine Datenbank mit komfortabler Benutzungs-oberfläche



RELATIONALE DATENBANKSYSTEME (RDBMS)

HOCHSCHULE DER MEDIEN

- Unter den Datenbank(management)systemen (DBMS) sind **Relationale Datenbankmanagementsysteme** (**RDBMS**) der derzeit meistverbreitete Typ von DBMS, deshalb Gegenstand dieser Vorlesung
  - ⇨ Fundamentales Konzept: Tabelle (oder „Relation“)
- Ältere Typen von DBMS („Auslaufmodelle“)
  - ⇨ **Hierarchische DBMS**
  - ⇨ **Netzwerk-DBMS**
- Neuere Typen
  - ⇨ **Objektorientierte DBMS** (derzeit nur für Spezialanwendungen in Gebrauch)
  - ⇨ **Objektrelationale DBMS** (um objektorientierte Merkmale erweiterte RDBMS, derzeit aktueller Trend)

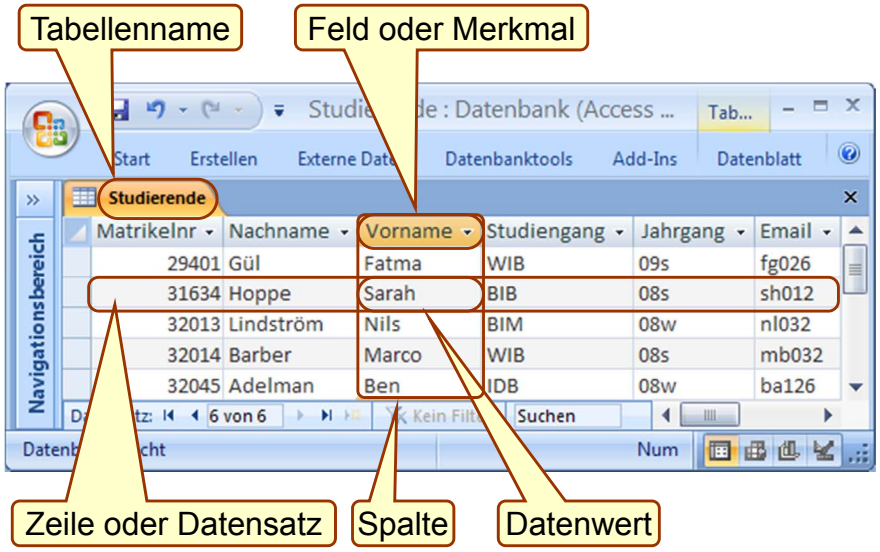
TABELLEN (RELATIONEN)

HOCHSCHULE DER MEDIEN

- Relationale DBMS speichern Daten in **Tabellen** (auch **Relationen** genannt, engl.: *table, relation*)
- Tabellen: intuitiv verständliche, einfach interpretierbare Art der Repräsentation von Informationen
- Die **Zeilen** der Tabelle (auch **Datensätze** genannt) repräsentieren untereinander gleichartige Informationseinheiten
- Datensätze sind gegliedert in **Felder** (auch **Merkmale** oder **Attribute** genannt).
- Die **Spalten** der Tabelle enthalten gleichartige Felder der Datensätze. Sie sind mit den Namen der Felder überschrieben
- In den Spalten stehen **Datenwerte** von gleichartigem **Datentyp** mit vordefiniertem **Wertebereich** (engl. *domain*)

TABELLEN

HOCHSCHULE DER MEDIEN



FELDER UND DATENWERTE

HOCHSCHULE DER MEDIEN

Felder

- besitzen einen **Felddatentyp**, die wichtigsten sind:
  - ⇨ Zahl
  - ⇨ Text mit/ohne festgelegte Maximallänge
  - ⇨ Datum/Uhrzeit
- besitzen einen **Wertebereich** (engl.: domain), z.B. gültige Matrikelnr oder die Menge der Kürzel ESB, WIB, IDB usw.

Datenwerte

- Müssen Felddatentyp und Wertebereich entsprechen
- Können auch undefiniert bleiben (sog. NULL-Werte)

PRIMÄRSCHLÜSSEL: EINFÜHRUNG

HOCHSCHULE DER MEDIEN

- **Primärschlüssel** = Feld oder Kombination von Feldern, deren Werte einen Datensatz eindeutig identifizieren
- Beispiele für Primärschlüssel:
  - ⇨ Die Matrikelnr der Tabelle Studierende
  - ⇨ Die Kombination Landeskennzahl, Ortskennzahl, Rufnummer für eine Tabelle Telefonanschluss
- Primärschlüssel (die aus einer Kombination von Feldern bestehen) sollen minimal sein, d.h. kein Feld aus der Kombination kann gestrichen werden, ohne dass Eindeutigkeit verlorengeht.
- Oft werden **„künstliche“** Primärschlüssel, z.B. laufende Nummern, eingeführt, um Datensätze eindeutig identifizieren zu können.

TABELLEN: ZUSAMMENFASSUNG

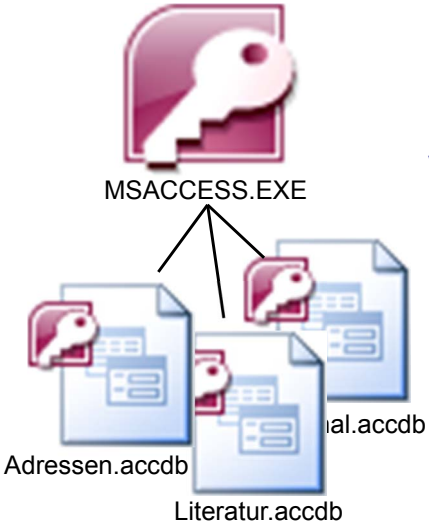
HOCHSCHULE DER MEDIEN

Tabelle = Menge von Datensätzen, die in Felder gegliedert sind, mit folgenden Eigenschaften:

- Die Tabelle hat einen eindeutigen Tabellennamen.
- Innerhalb der Tabelle ist jeder Feldnamen eindeutig und bezeichnet eine Spalte mit der betreffenden Eigenschaft.
- Die Reihenfolge der Spalten ist bedeutungslos.
- Die Reihenfolge der Datensätze ist bedeutungslos.
- Es gibt einen Primärschlüssel, d.h. eines der Felder oder eine Feldkombination identifiziert die Datensätze eindeutig.
- Es gibt keine Begrenzung für die Anzahl der Spalten und der Zeilen einer Tabelle (zumindest in der Theorie).

DAS DATENBANKSYSTEM ACCESS

HOCHSCHULE DER MEDIEN



Ein Datenbanksystem:

- Die Anwendung MSACCESS (MSACCESS.EXE)

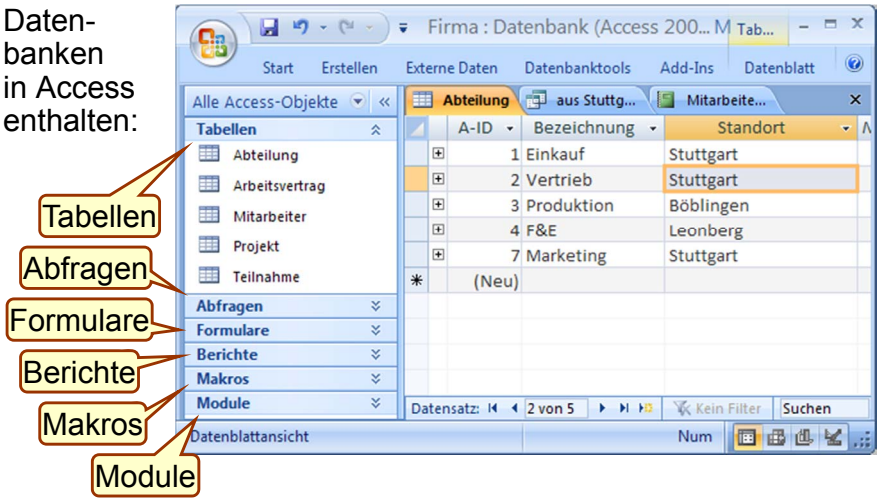
Viele Datenbanken:

- = alle Dateien mit der Endung .accdb (ältere Version: .mdb)
- Beim Anklicken dieser Dateien wird die Anwendung MSACCESS gestartet.
- Neue Datenbanken werden durch Aufruf von MSACCESS erzeugt.

ACCESS-DATENBANKEN: ENTHALTENE OBJEKTE

HOCHSCHULE DER MEDIEN

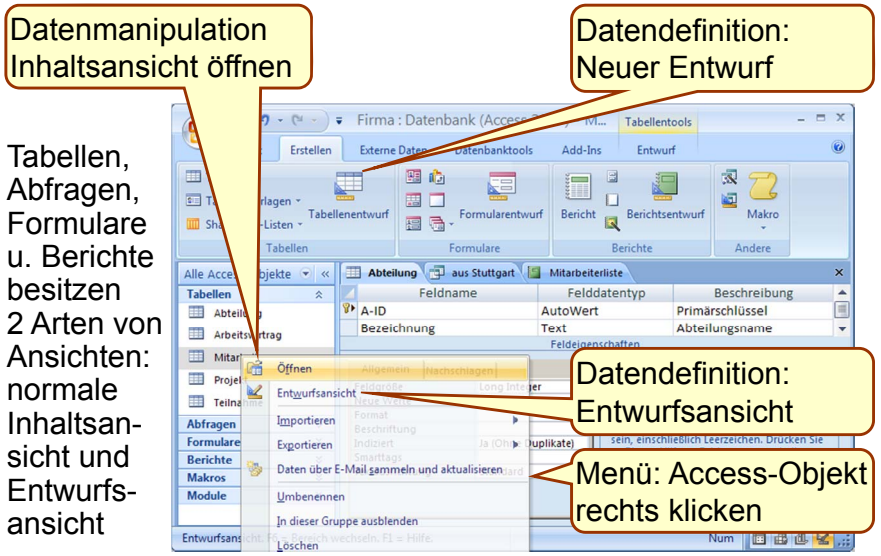
Daten-banken in Access enthalten:



DATENMANIPULATION UND DATENDEFINITION

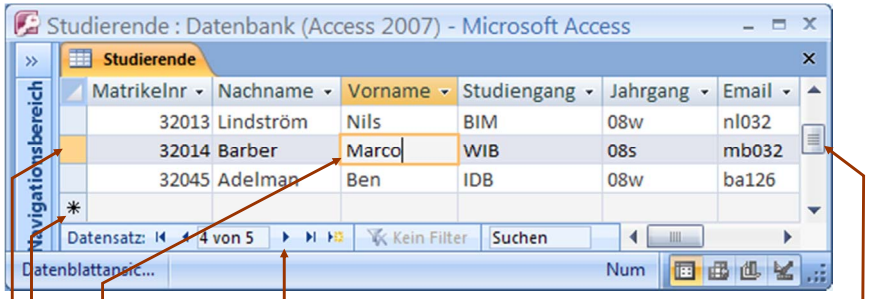
HOCHSCHULE DER MEDIEN

Tabellen, Abfragen, Formulare u. Berichte besitzen 2 Arten von Ansichten: normale Inhaltsan-sicht und Entwurfs-ansicht



DATENMANIPULATION MIT ACCESS: ANZEIGEN UND ÄNDERN VON DATEN

HOCHSCHULE DER MEDIEN



Tabelleninhalte anzeigen durch Blättern, Rollen

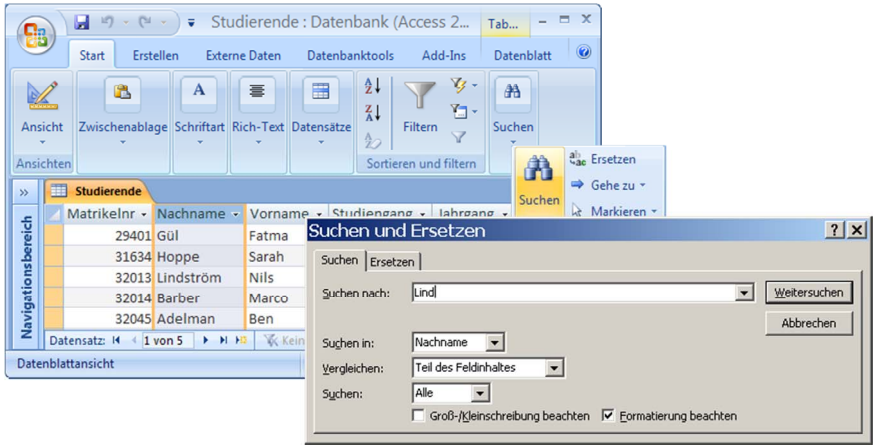
Datensätze ändern durch Überschreiben

Hier neuen Datensatz eintragen

Markierter Datensatz, kann über Menü gelöscht oder kopiert werden

DATENMANIPULATION MIT ACCESS: SUCHEN VON DATENWERTEN

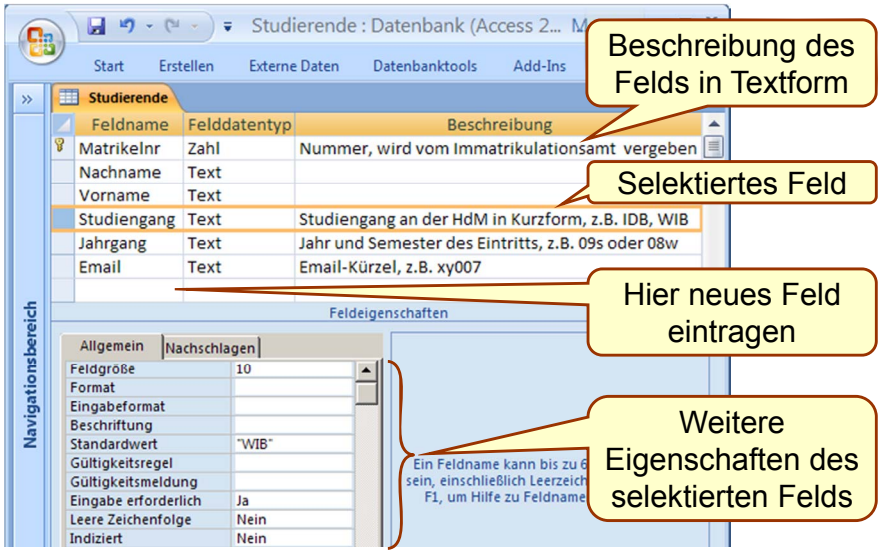
HOCHSCHULE DER MEDIEN



Suchen von Datensätzen mit Menü Bearbeiten – Suchen
(Das ist eine Standard-Office-Funktion, keine Datenbankabfrage!
Zu Datenbankabfragen siehe Teil 3 des Skripts.)

DATENDEFINITION MIT ACCESS: ENTWURFSANSICHT VON TABELLEN

HOCHSCHULE DER MEDIEN



Beschreibung des Felds in Textform

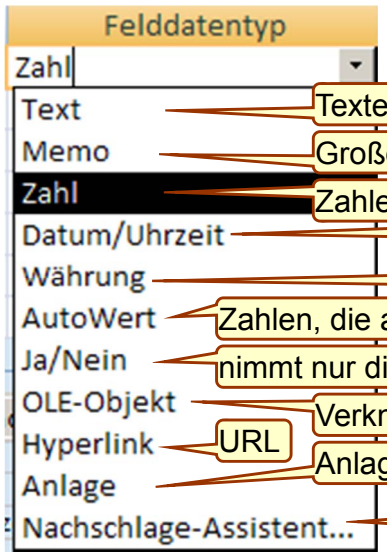
Selektiertes Feld

Hier neues Feld eintragen

Weitere Eigenschaften des selektierten Felds

DATENDEFINITION MIT ACCESS: FELDDATENTYPEN

HOCHSCHULE DER MEDIEN



In Access können Tabellenfelder verschiedene Felddatentypen besitzen

Texte

Große Texte (bis zu 65535 Zeichen)

Zahlen

Kalenderdatum, Uhrzeit

Geldbeträge, z.B. in Euro

Zahlen, die automatisch vergeben werden

nimmt nur die Werte Ja (☑) oder Nein (☐) an

Verknüpfte oder eingebettete Datei

URL

Anlage zum Datensatz, wie bei E-Mail

Erzeugt Auswahlmenü für Werte

DATENDEFINITION MIT ACCESS: EIGENSCHAFTEN DER FELDER

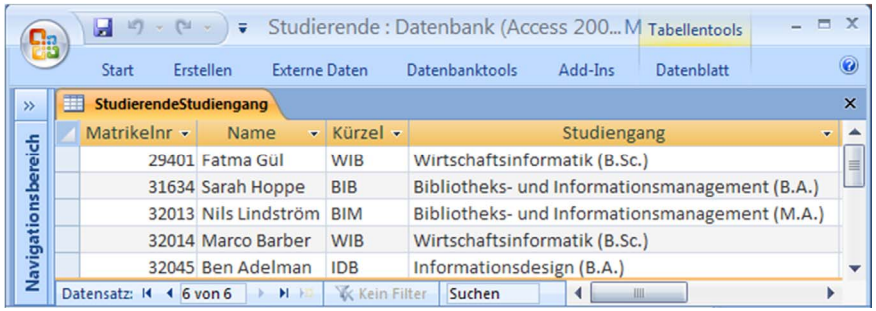
HOCHSCHULE DER MEDIEN

In Access können zu Feldern außer dem Felddatentyp noch weitere Beschreibungselemente festgelegt werden:

- Die **Beschreibung** in Textform dient zur Dokumentation und als Hilfetext für die Benutzungsoberfläche.
- Die **Feldeigenschaften** ermöglichen weitere Festlegungen zum selektierten Feld, so u.a.:
  - ⇨ **Feldgröße**: z.B. maximale Zeichenzahl in Textfeldern
  - ⇨ **Format, Eingabeformat, Dezimalstellenanzeige**: Angaben zur Darstellung der Werte
  - ⇨ **Standardwert**: automatische Vorbelegung mit Werten
- **Nachschlagen**: Der Anfänger nutzt diese Funktion am besten über den Nachschlageassistenten (erreichbar über Felddatentyp). Es kann ein Menü erzeugt werden, aus dem einzugebende Werte ausgewählt werden können.

PROBLEME BEIM DATENBANKENTWURF

HOCHSCHULE DER MEDIEN



Der Vorgang der Datendefinition ist mit Access eigentlich nicht schwer.

Trotzdem kann man beim Entwurf einer Datenbank vieles falsch machen.

Wo liegen die Probleme bei dieser Datenbank?

DATENBANDESIGN: PROBLEMPUNKTE

HOCHSCHULE DER MEDIEN

- Datenfelder sollen nur **elementare Werte** enthalten!
  - ⇨ Im vorigen Beispiel: Nachname und Vorname in unterschiedlichen Feldern speichern.
- **Redundanzen vermeiden!** Redundanz bedeutet, dass Daten ohne Informationsverlust weggelassen werden könnten.
  - ⇨ In dieser Tabelle wird der Zusammenhang zwischen Studiengang und Studiengangkürzel redundant (d.h. überflüssigerweise mehrfach) gespeichert.
  - ⇨ Das Problem besteht nicht darin, dass der Wert „WIB“ mehrfach in der Tabelle steht. Es besteht auch nicht darin, dass „WIB“ und „Wirtschaftsinformatik (B.Sc.)“ dasselbe bedeutet.
  - ⇨ **Das Problem besteht darin, dass der Zusammenhang zwischen „WIB“ und „Wirtschaftsinformatik (B.Sc.)“ mehrfach in der Tabelle steht**

ANOMALIEN

HOCHSCHULE DER MEDIEN

Redundanz bewirkt so genannte **Anomalien**

- **Einfügeanomalie:** Man möchte einen neuen Studiengang einführen (z.B. Mobile Services). Das geht nicht, solange es noch keine Studierenden dazu gibt!
- **Löschanomalie:** Löschen von vielen Studierenden kann ungewolltes Löschen eines Studiengangs bewirken (Umkehrung der Einfügeanomalie, hier nicht so relevant).
- **Änderungsanomalie:** Eine Änderung der Studiengangbezeichnung (z.B. Informationswirtschaft statt Informationsmanagement) muss bei allen Studierenden vorgenommen werden, obwohl dies nur einen einzigen Sachverhalt betrifft.

Abhilfe: Verteilung des Sachverhalts auf mehrere Tabellen.
⇨ Das vermeidet Redundanzen und damit die Anomalien.

VERMEIDUNG VON REDUNDANZEN

HOCHSCHULE DER MEDIEN

Redundanz vermeiden!

- **Kompliziertere Informationen in mehreren Tabellen repräsentieren!** Z.B. eine Tabelle für Studierende, eine zweite Tabelle für Studiengänge. (Dies wird im Folgenden noch erklärt.)
- **Tabellenfelder weglassen, die aus anderen berechnet werden können!** Beispiel: In der nachfolgenden Bestelliste ist das Feld Gesamtpreis überflüssig. Wenn benötigt, kann es durch Abfragen oder Formulare berechnet werden.

Ware	Anzahl	Einzelpreis	Gesamtpreis
Cola	2	1,60	3,20
Bier	1	2,80	2,80
Tee	4	1,80	7,20

Redundantes Feld: In Tabelle weglassen!

TEIL 2: AUFBAU VON DATENBANKEN INHALTSÜBERSICHT

HOCHSCHULE DER MEDIEN

AUFBAU VON DATENBANKEN

HOCHSCHULE DER MEDIEN

Der Aufbau von Datenbanken erfolgt in vier Schritten:

1. **Anforderungsanalyse:** Erfassung der zum Aufbau der Datenbank erforderlichen Sachverhalte in der Sprache der künftigen Anwender
2. **Entitäten-Beziehungsmodell**: Entwurf einer graphischen Repräsentation der darzustellenden Objekte („Entitäten“) und ihrer Beziehungen
3. Umsetzung des Entitäten-Beziehungsmodells durch **Definition** von geeigneten **Tabellen**
   - ⇨ **Datendefinition**
4. **Füllen** der Datenbank mit aktuellen **Inhalten**
   - ⇨ **Datenmanipulation**
   - ⇨ *beschrieben in* Teil 3: Nutzung von Datenbanken

AUFBAU VON DATENBANKEN: 1. ANFORDERUNGSANALYSE

HOCHSCHULE DER MEDIEN

Anforderungsanalyse:

Erfassung der in der Datenbank abzubildenden Sachverhalte in der Sprache der künftigen Anwender.

Typischer Ablauf einer „Anforderungsanalyse“:
Der Datenbankentwickler führt ein Gespräch mit den künftigen Anwendern der Datenbank, also mit seinen Kunden. Er erfragt die Anforderungen, d.h. welche Art von Informationen in der Datenbank repräsentiert werden sollen. Das Ergebnis dieses Gesprächs wird in einem kleinen Protokoll festgehalten.

Anmerkung: In einer früheren Version dieses Folienskripts wurde anstelle des Begriffs „Anforderungsanalyse“ der Begriff „Datenanalyse“ (Meier 2001) verwendet. Dieser Begriff hat sich jedoch als missverständlich herausgestellt, da es in dieser Phase gerade noch nicht um Daten, sondern um eine Analyse der Anforderungen der künftigen Anwender geht.

BEISPIEL FÜR DAS ERGEBNIS EINER ANFORDERUNGSANALYSE

HOCHSCHULE DER MEDIEN

In einer Datenbank sollen Sachverhalte der folgenden Art repräsentiert werden können:

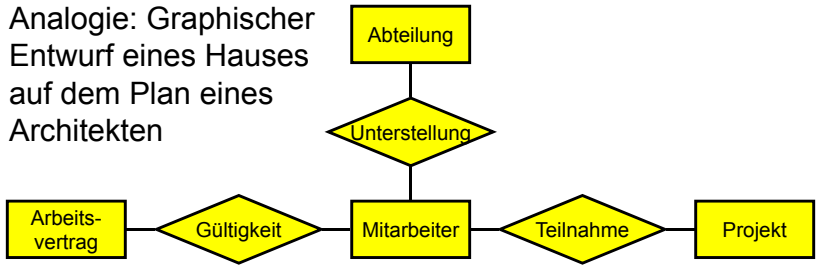
- Jede **Abteilung** besitzt eine Bezeichnung.
- Jeder **Mitarbeiter** ist charakterisiert durch einen Namen sowie durch Straße und Ort, in welchen er wohnt.
- Jeder **Arbeitsvertrag** legt eine Funktion (z.B. „Buchhalter“) eines Mitarbeiters fest sowie das Gehalt, das er verdient.
- Jedes **Projekt** besitzt einen Namen und eine eindeutige Projektnummer.
- Für jeden Mitarbeiter ist ein Arbeitsvertrag **gültig**, jeder Arbeitsvertrag ist genau für einen Mitarbeiter gültig.
- Jeder Mitarbeiter ist einer Abteilung **unterstellt**.
- Mitarbeiter können gleichzeitig an mehreren Projekten **teilnehmen**, wobei die jeweiligen Prozentanteile erfasst werden.

AUFBAU VON DATENBANKEN: 2. ENTITÄTEN-BEZIEHUNGSMODELL

HOCHSCHULE DER MEDIEN

Entitäten-Beziehungsmodell (*entity relationship model - ERM*)

- Graphischer Entwurf der Datenbank
- Alle wesentlichen Elemente werden graphisch dargestellt
- Umsetzung der Anforderungsanalyse in ein erstes Design
  - ⇨ dient auch zur Kommunikation mit dem Kunden
  - ⇨ und zur Verifizierung der Anforderungsanalyse
- Analogie: Graphischer Entwurf eines Hauses auf dem Plan eines Architekten



AUFBAU VON DATENBANKEN: 2. ENTITÄTEN-BEZIEHUNGSMODELL

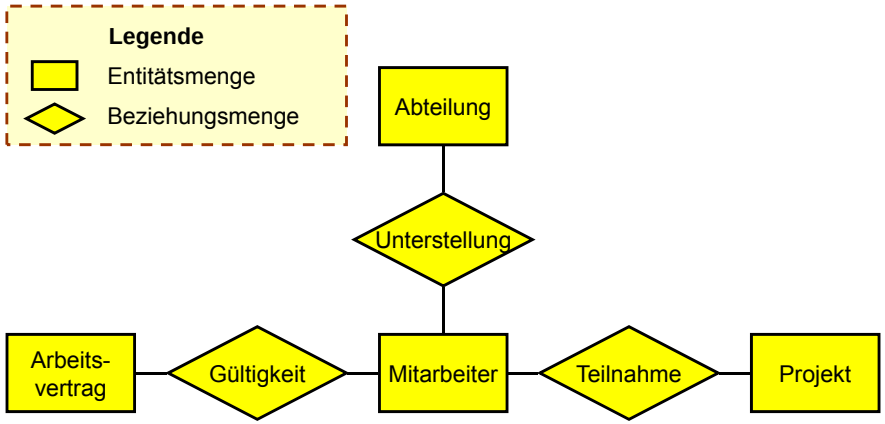
HOCHSCHULE DER MEDIEN

Elemente eines Entitäten-Beziehungsmodells

- **Entität** (*entity*): Wohlunterscheidbares Objekt der realen Welt oder unserer Vorstellung. Beispiele für Entitäten: Individuen, Gegenstände, Begriffe, Ereignisse.
  - ⇨ Entitäten gleichen Typs bilden sog. **Entitätsmengen** (alternativer Name: Entitätstypen, engl.: *entity types*) und besitzen bestimmte Merkmale.
  - ⇨ Entitäten besitzen eindeutigen Identifikationsschlüssel.
- **Beziehung** (*relationship*):
  - ⇨ stellt eine Beziehung zwischen Entitäten dar.
  - ⇨ Beziehungen gleichen Typs bilden **Beziehungsmengen** (auch: Beziehungstypen, engl.: *relationship types*) und können zusätzlich bestimmte Merkmale tragen.

ENTITÄTEN-BEZIEHUNGSMODELL: GRAPHISCHE DARSTELLUNG

HOCHSCHULE DER MEDIEN



BEISPIEL: ENTITÄT MAIER UND ENTITÄTSMENGE MITARBEITER

HOCHSCHULE DER MEDIEN

Entität: **Mitarbeiter Maier**, wohnhaft in der **Türlenstraße** in **Stuttgart**

Entitätsmenge: Menge aller **Mitarbeiter** mit den Merkmalen **Name**, **Straße** und **Ort**

Identifikationsschlüssel: **Mitarbeiternummer** als künstlicher Schlüssel

Darstellung im Entitäten-Beziehungsmodell:

BEISPIEL: BEZIEHUNG UND BEZIEHUNGSMENGE

HOCHSCHULE DER MEDIEN

Beziehung: Mitarbeiter **Maier** **arbeitet** zu **70%** im Projekt Nr. **17**

Beziehungsmenge: Menge aller **Mitarbeiter-Projekt-Teilnahmen** mit den Merkmalen **Mitarbeiternummer**, **Projektnummer** und **Prozentanteil**

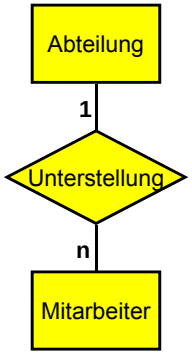
Identifikationsschlüssel: Zusammengesetzter Schlüssel aus **Mitarbeiter- und Projektnummer**

Darstellung im Entitäten-Beziehungsmodell:

MÄCHTIGKEIT VON BEZIEHUNGSMENGEN (1)

HOCHSCHULE DER MEDIEN

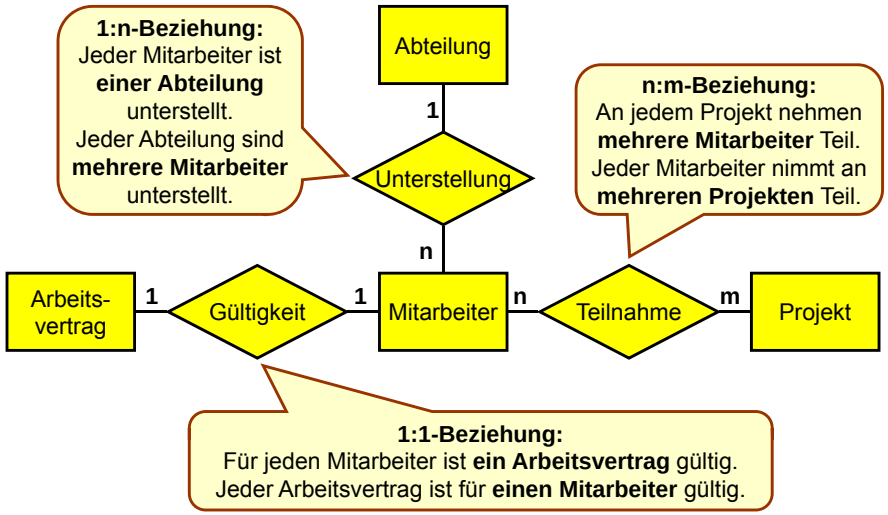
Beziehungsmengen können sich darin unterscheiden, wie „viele“ Entitäten sie einander zuordnen können.



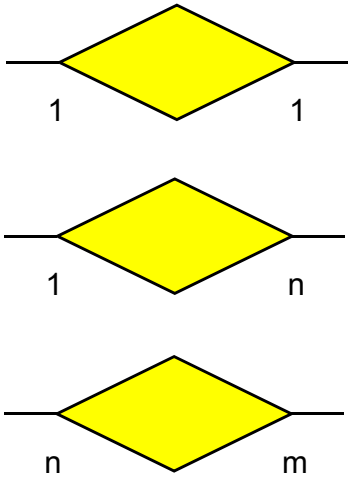
- Beispielsweise kann für die Beziehungsmenge „Unterstellung“ festgelegt werden, dass jeder Mitarbeiter **einer** Abteilung zugeordnet ist, jede Abteilung jedoch **mehrere** Mitarbeiter besitzt.
- Diese Eigenschaft der Beziehungsmenge wird als **Mächtigkeit** (alternative Namen: **Kardinalität**, **Assoziationstyp**) bezeichnet.
- Die Mächtigkeit wird mit Hilfe von Zahlen (**1** für „ein(e)“) und Symbolen (**n**, **m** oder ∗ für „mehrere“) notiert.

MÄCHTIGKEIT VON BEZIEHUNGSMENGEN (2)

HOCHSCHULE DER MEDIEN



MÄCHTIGKEIT VON BEZIEHUNGSMENGEN (3)



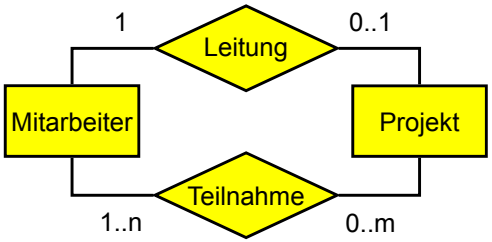
1:1-Beziehung
einfach-einfache Beziehung
(engl.: *one-to-one relationship*)

1:n-Beziehung
einfach-komplexe Beziehung
(engl.: *one-to-many relationship*)

n:m-Beziehung
komplex-komplexe Beziehung
(*many-to-many relationship*)

INTERVALLDARSTELLUNG FÜR DIE MÄCHTIGKEIT

In der so genannten Intervalldarstellung lässt sich die Mächtigkeit von Beziehungsmengen durch Angabe von Untergrenzen noch präziser festlegen. Beispiel:



- Jeder Mitarbeiter kann ein Projekt leiten, muss aber nicht: Mächtigkeit **0..1**, d.h. zwischen 0 und 1
- Jeder Mitarbeiter kann an mehreren Projekten teilnehmen, muss aber nicht: Mächtigkeit **0..m**, d.h. zwischen 0 und m
- An jedem Projekt nimmt mindestens ein Mitarbeiter teil: Mächtigkeit **1..n**, d.h. zwischen 1 und n

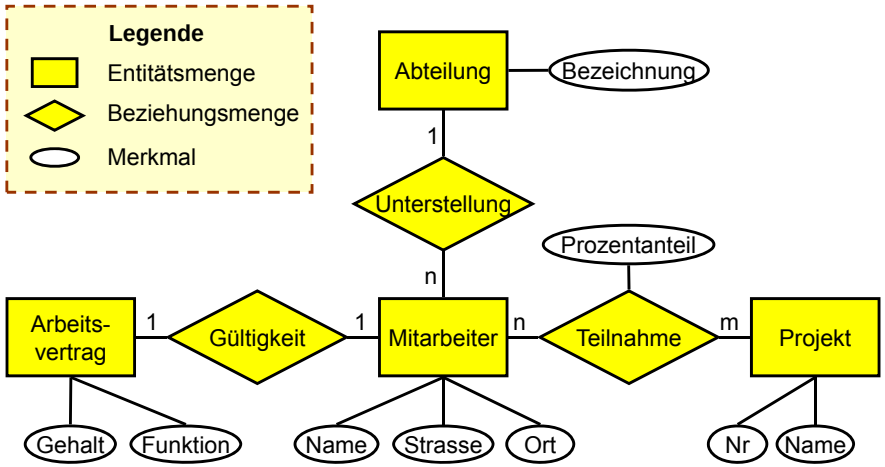
ZUR NOTATION DER MÄCHTIGKEIT VON BEZIEHUNGSMENGEN

Achtung: Die Mächtigkeiten von Beziehungsmengen werden in der Literatur sehr unterschiedlich dargestellt.

Dieses Skript orientiert sich an der Notation, die von der Mehrheit der Wissenschaftler angewandt wird, so auch vom Erfinder der Entitäten-Beziehungsmodellierung (Chen 1976) und vielen Lehrbuchautoren (z.B. Hansen & Neumann 2002).

In einer früheren Version dieses Skripts (bis zum Sommersemester 2003) wurde eine andere Notation angewandt, die auf Andreas Meier zurückgeht, da sich das Skript damals stark an dessen Lehrbuch (Meier 2001) orientierte. Meier nimmt eine spiegelbildliche Darstellung vor. D.h. bei einfach-komplexen Beziehungen sind die Plätze von „1“ und „n“ vertauscht. Dieser Unterschied ist zu beachten, wenn das Buch von Meier für Studienzwecke genutzt wird.

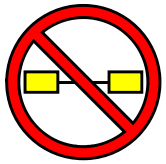
ENTITÄTEN-BEZIEHUNGSMODELL: DARSTELLUNG MIT MERKMALEN



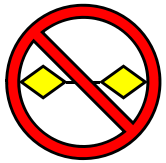
„SYNTAX“ VON ENTITÄTEN-BEZIEHUNGSMODELLEN

HOCHSCHULE DER MEDIEN

Entitäten-Beziehungsmodelle besitzen eine „Syntax“:



- Zwei Entitätsmengen (Rechtecke) dürfen nie direkt nebeneinander liegen; sie dürfen nur über eine Beziehungsmenge (Raute) miteinander verbunden werden.



- Zwei Beziehungsmengen (Rauten) dürfen nie direkt nebeneinander liegen; dazwischen muss immer eine Entitätsmenge (ein Rechteck) liegen.

- Merkmale (Ellipsen) können nur direkt entweder an eine einzige Entitätsmenge (Rechteck) oder eine einzige Beziehungsmenge (Raute) angehängt werden.

AUFBAU VON DATENBANKEN: 3. DATENDEFINITION

HOCHSCHULE DER MEDIEN

Als nächster Schritt muss das Entitäten-Beziehungsmodell in Form von Tabellen umgesetzt werden.

Dies ist der Vorgang der **Datendefinition.** Darunter verstehen wir die Festlegung des sogenannten **Datenbankschemas**, das die **Struktur** einer Datenbank beschreibt:

- Namen der Tabellen und ihrer Felder
- allgemeine Eigenschaften zu den Tabellen und Feldern (z.B. Wertebereiche, Beziehungen, Integritätsbedingungen)

*Der konkrete **Inhalt** der Datenbank ist kein Gegenstand der Datendefinition. Der Inhalt ist Gegenstand des 4. Schritts, der **Datenmanipulation** genannt wird. Die Datenmanipulation wird in einem separaten Vorlesungsabschnitt behandelt.*

PRIMÄRSCHLÜSSEL

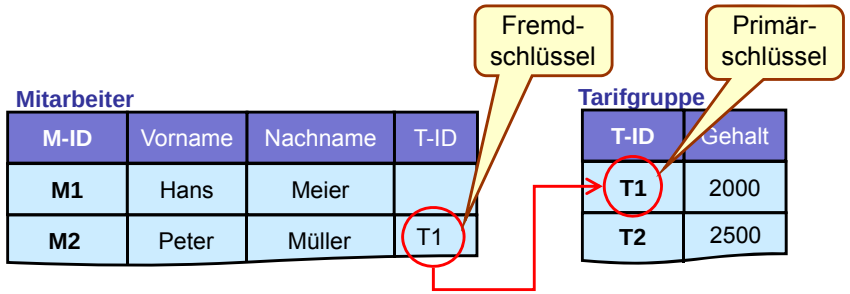
HOCHSCHULE DER MEDIEN

- Ein Feld oder eine Feldkombination in einer Tabelle heißt **eindeutig**, wenn dadurch jeder Datensatz in der Tabelle eindeutig bestimmt ist.
- Eine eindeutige Feldkombination heißt **minimal**, wenn keines der Felder weggelassen werden kann, ohne die Eindeutigkeit aufzugeben.
- Eindeutige Felder und eindeutige, minimale Feldkombinationen heißen **Schlüsselkandidaten**.
- Einer der Schlüsselkandidaten einer Tabelle hat die Rolle des **Primärschlüssels**. Ein solcher dient dazu, die Datensätze in der Tabelle eindeutig zu bezeichnen.

FREMDSCHLÜSSEL

HOCHSCHULE DER MEDIEN

Die Werte von Primärschlüsseln können auch in weiteren Feldern auftauchen; solche Felder nennt man **Fremdschlüssel**. Sie werden verwendet, um von einem Datensatz auf einen anderen Datensatz zu verweisen. So können in Tabellen Beziehungen dargestellt werden



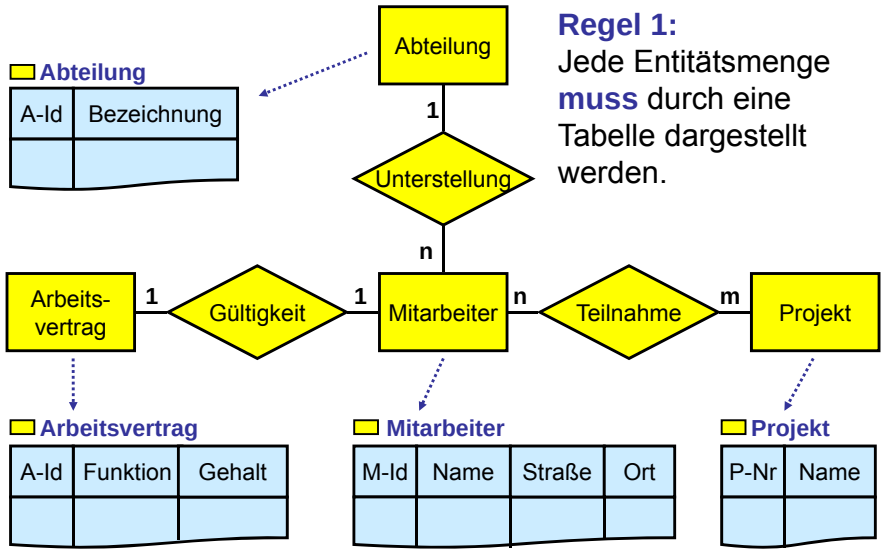
ÜBERFÜHRUNG DES ENTITÄTEN-BEZIEHUNGSMODELS IN TABELLEN

HOCHSCHULE DER MEDIEN

- **Regel 1:** Jede **Entitätsmenge muss** durch eine eigenständige Tabelle dargestellt werden:
  - ⇨ Die Merkmale der Entitätsmenge werden zu Feldern der Tabelle
  - ⇨ Es gibt einen Primärschlüssel, in der Regel ist das der Identifikationsschlüssel der Entitätsmenge
- **Regel 2:** Jede **Beziehungsmenge kann** durch eine eigenständige Tabelle dargestellt werden (manchmal ist das aber nicht unbedingt nötig, siehe weiter hinten):
  - ⇨ Die Merkmale der Beziehungsmenge werden zu Merkmalen der Tabelle
  - ⇨ Primärschlüssel der Beziehungsmenge ist oft die Kombination der Fremdschlüssel der zugehörigen Entitätsmengen oder ein künstlicher Schlüssel (z.B. laufende Nummer)

REGEL 1: MUSS-REGEL FÜR ENTITÄTSMENGEN

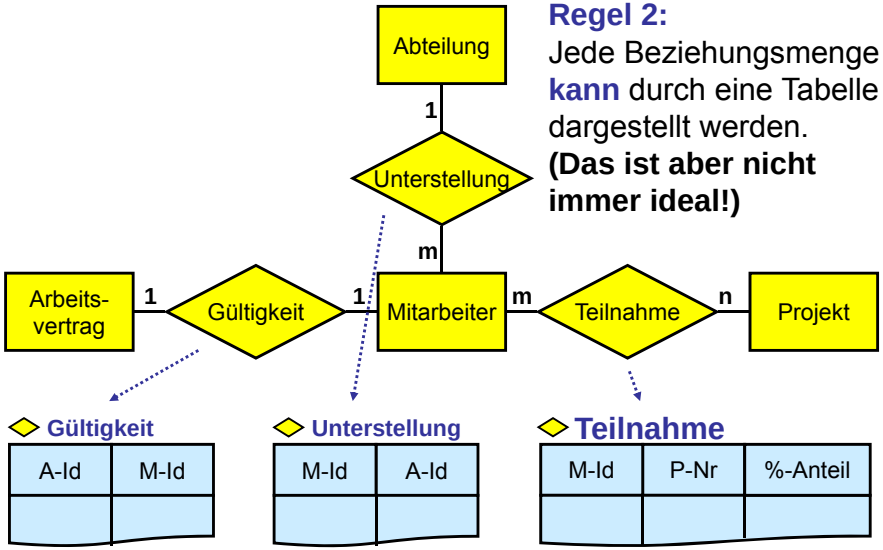
HOCHSCHULE DER MEDIEN



Regel 1:
Jede Entitätsmenge **muss** durch eine Tabelle dargestellt werden.

REGEL 2: KANN-REGEL FÜR BEZIEHUNGSMENGEN

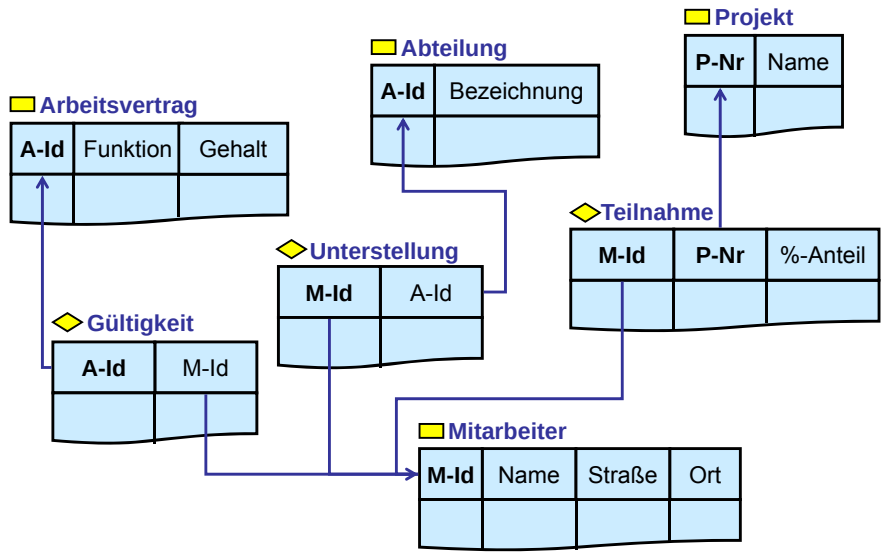
HOCHSCHULE DER MEDIEN



Regel 2:
Jede Beziehungsmenge **kann** durch eine Tabelle dargestellt werden.
(Das ist aber nicht immer ideal!)

ERSTER ANSATZ: UMSETZUNG MIT HILFE DER REGELN 1 UND 2

HOCHSCHULE DER MEDIEN



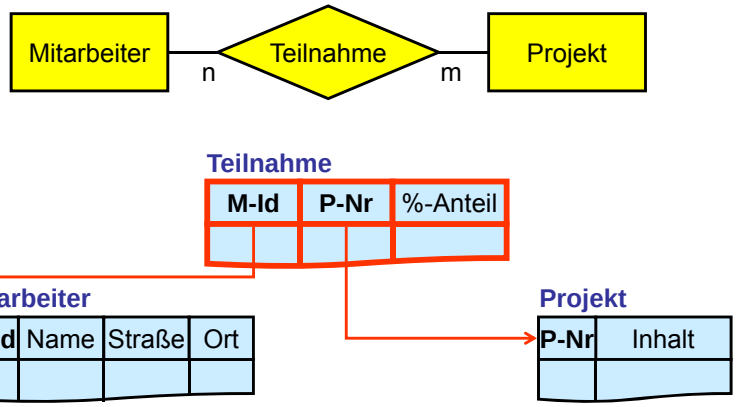
KONSEQUENZEN DES ERSTEN ANSATZES MIT REGEL 1 UND 2

HOCHSCHULE DER MEDIEN

- Wenn man entsprechend Regel 2 alle Beziehungsmengen durch eigenständige Tabellen modelliert, erzeugt man eine große Zahl von Tabellen.
- Dies ist nicht immer notwendig.
- Welche Tabellen könnte man in unserem Beispiel einsparen und wie?
- Die Antwort geben die nachfolgend aufgeführten Regeln 3, 4 und 5.

REGEL 3 FÜR KOMPLEX-KOMPLEXE BEZIEHUNGEN

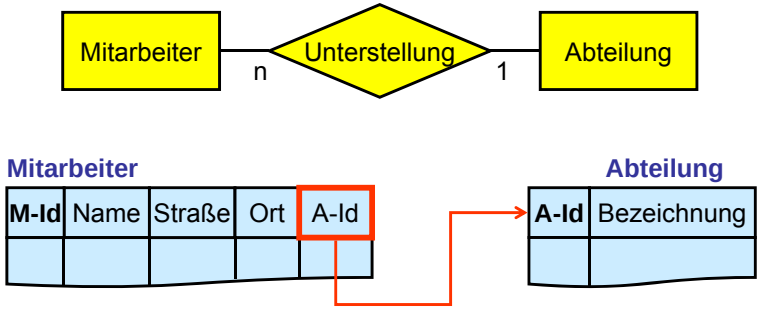
HOCHSCHULE DER MEDIEN



Regel 3: Komplex-komplexe Beziehungsmengen müssen als **eigenständige Tabellen** definiert werden. Primärschlüssel der Beziehungsmengentabelle ist häufig die Kombination der beiden Fremdschlüssel oder ein künstlicher Schlüssel.

REGEL 4 FÜR EINFACH-KOMPLEXE BEZIEHUNGEN

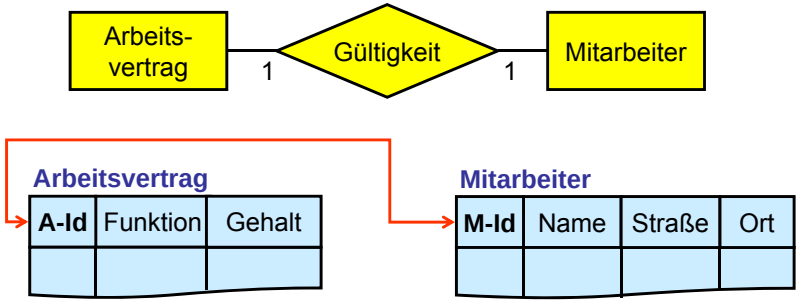
HOCHSCHULE DER MEDIEN



Regel 4: Einfach-komplexe Beziehungsmengen können **ohne eigenständige Beziehungsmengentabelle** definiert werden. Der Entitätentabelle auf der „komplexen“ Seite der Beziehung (markiert durch „n“) wird ein Fremdschlüssel auf die andere Entitätentabelle zusammen mit eventuellen weiteren Merkmalen der Beziehungsmenge hinzugefügt.

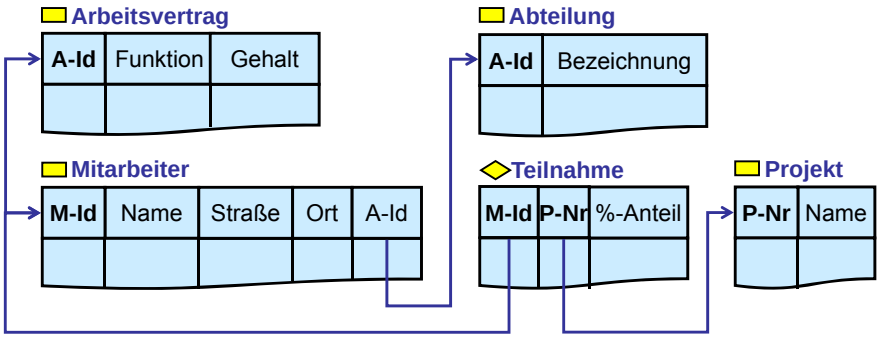
REGEL 5 FÜR EINFACH-EINFACHE BEZIEHUNGEN

HOCHSCHULE DER MEDIEN



Regel 5: Einfach-einfache Beziehungsmengen können **ohne eigenständige Beziehungsmengentabelle** definiert werden. Zusammengehörige Datensätze werden entweder durch identische Primärschlüssel gekennzeichnet (wie hier) oder eine der Tabellen erhält einen Fremdschlüssel, der auf den Primärschlüssel in der anderen Tabelle verweist (wie bei einfach-komplexen Beziehungen, siehe letzte Folie).

ERGEBNIS



Die fettgedruckten Merkmale bzw. Merkmalskombinationen (**A-Id, M-Id, P-Id**) sind Primärschlüssel.

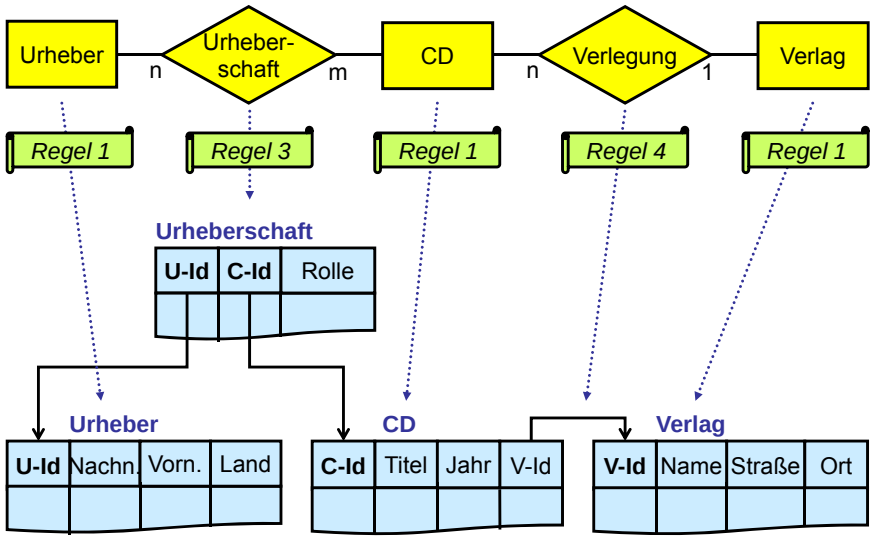
Die Merkmale, von denen Pfeile ausgehen, sind Fremdschlüssel. Die Pfeile enden mit ihren Spitzen bei den Primärschlüsseln. M-Id und P-Id in Teilnahme sind für sich Fremdschlüssel und in der Kombination Primärschlüssel. A-Id in Arbeitsvertrag und M-Id in Mitarbeiter sind zugleich Primär- und Fremdschlüssel.

ÜBUNGSBEISPIEL CD-SAMMLUNG: AUFGABENSTELLUNG

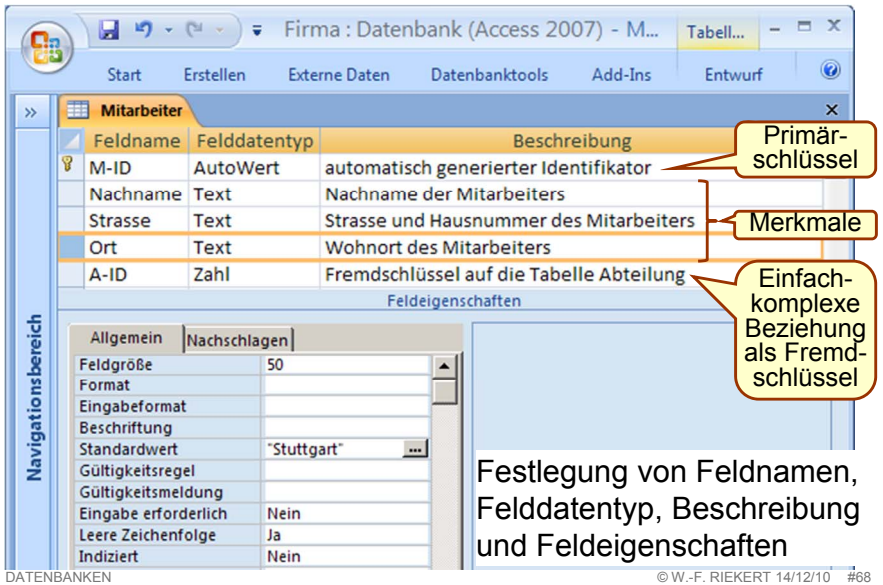
Erstellen Sie für das nachfolgend beschriebene Beispiel einer CD-Sammlung selbst ein Entitäten-Beziehungsmodell und überführen Sie dieses in eine Tabellendarstellung:

- Jede CD besitzt einen eindeutigen Identifikator, einen Titel, sowie ein Erscheinungsjahr
- Es gibt Urheber, die durch Nachnamen, Vornamen und ein Heimatland genauer definiert sind
- CDs haben einen oder mehrere Urheber und umgekehrt. Die Urheberschaft kann durch eine Rolle (z.B. Komponist, Interpret) genauer definiert sein.
- Jeder CD ist genau ein Verlag zugeordnet, der durch einen Namen und die üblichen Adressangaben genauer definiert ist.

ÜBUNGSBEISPIEL CD-SAMMLUNG: LÖSUNG



DEFINITION VON TABELLEN: ENTWURFSANSICHT



Festlegung von Feldnamen, Felddatentyp, Beschreibung und Feldeigenschaften

INDIZIERUNG VON FELDERN

HOCHSCHULE DER MEDIEN

- Ein Feld einer Tabelle kann auf Wunsch **indiziert** werden, d.h. die Datenbank besitzt dann für das Feld einen **Index** = ein sortiertes Verzeichnis der Datenwerte und der zugeordneten Datensätze
- Für die Benutzer ist der Index unsichtbar. Die Benutzung der Datenbank erfolgt mit und ohne Index auf identische Weise.
- Vorteil: Ein Index beschleunigt die Suche nach Datensätzen mit bestimmten Datenwerten im indizierten Feld.
- Nachteil (meist nicht gravierend): Das Einfügen neuer Datensätze in die Tabelle dauert etwas länger und es wird etwas mehr Speicherplatz auf der Festplatte benötigt
- Analogie: Stichwortverzeichnis in einem Buch (oft auch Index genannt).

STRUKTURELLE INTEGRITÄTSBEDINGUNGEN

HOCHSCHULE DER MEDIEN

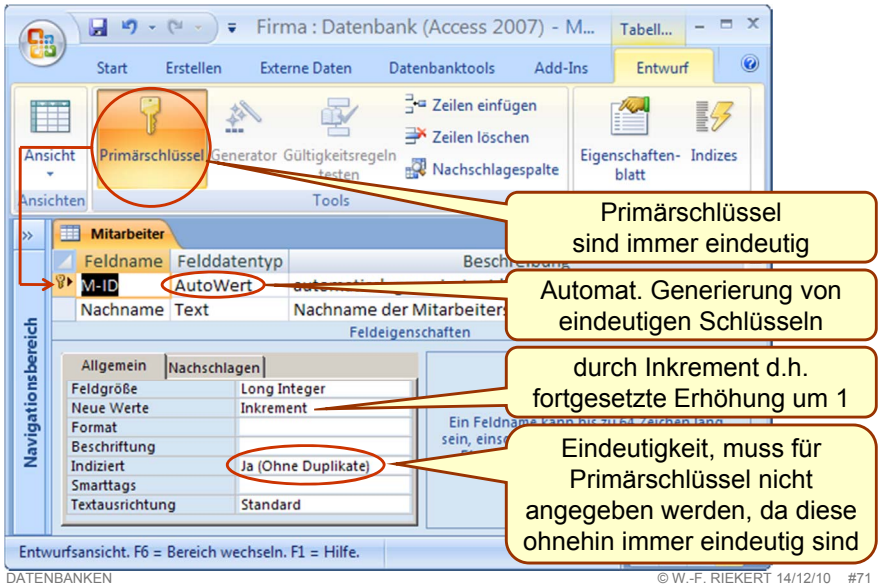
Eine Datenbank ist **integer** oder **konsistent**, falls die zugrundeliegenden Daten fehlerfrei erfasst sind und den gewünschten Informationsgehalt fehlerfrei wiedergeben

Es gibt sogenannte **strukturelle Integritätsbedingungen**, die durch das Datenbanksystem selbst ausgedrückt werden können:

- **Eindeutigkeitsbedingung** (Identifikationsschlüssel muss eindeutig sein)
- **Wertebereichsbedingung** (Merkmale einer Tabelle können nur Datenwerte aus bestimmtem Wertebereich annehmen)
- **Referenzielle Integrität**: Jeder Wert eines Fremdschlüssels muss als Schlüsselwert der referenzierten Tabelle existieren
  (Referenz = Verweis, Fremdschlüssel verweist auf Primärschlüssel)

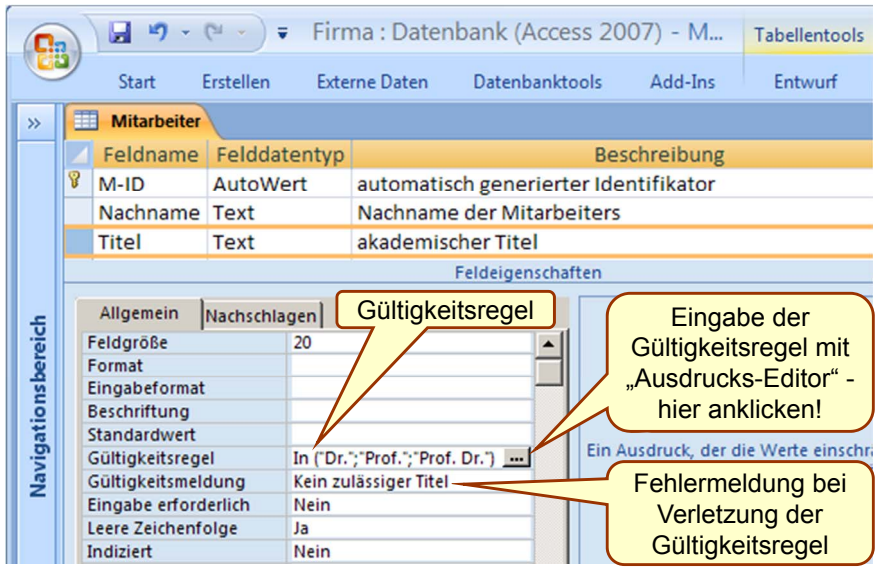
EINDEUTIGKEITSBEDINGUNGEN IN ACCESS

HOCHSCHULE DER MEDIEN



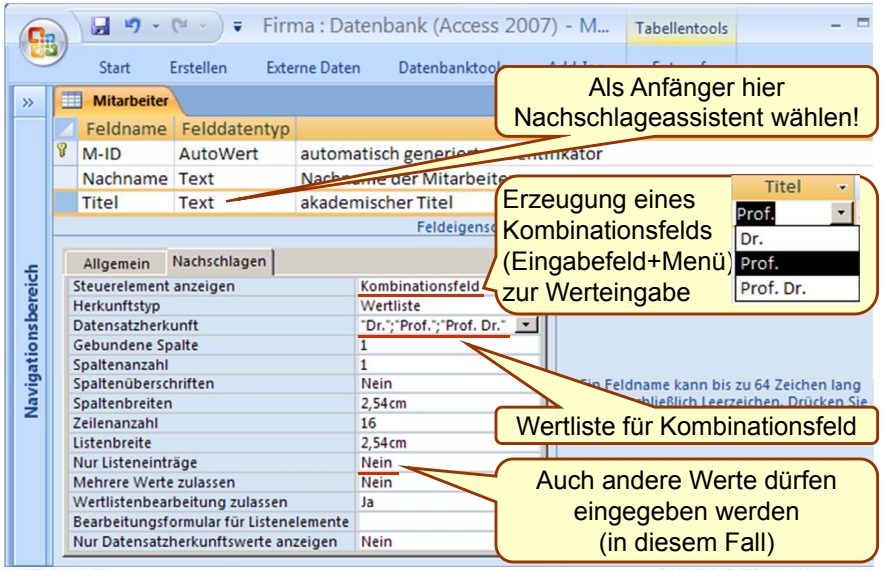
WERTEBEREICHSBEDINGUNGEN IN ACCESS: GÜLTIGKEITSREGELN

HOCHSCHULE DER MEDIEN



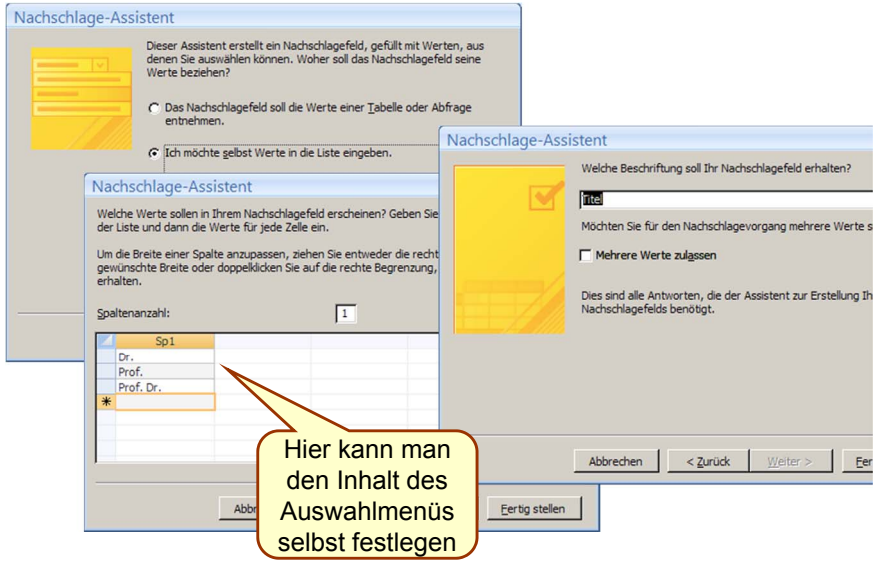
WERTEBEREICHSBEDINGUNGEN IN ACCESS: EINGABEHILFEN

HOCHSCHULE DER MEDIEN



EINGABEHILFE ERSTELLEN MIT NACHSCHLAGEASSISTENT

HOCHSCHULE DER MEDIEN



FREMDSCHLÜSSEL

HOCHSCHULE DER MEDIEN

- **Einfach-einfache** und **einfach-komplexe Beziehungs-mengen** lassen sich mit Hilfe von Fremdschlüsselfeldern in einer der verknüpften Entitätstabellen darstellen.
  - ⇨ Beispiele: M-Id in Arbeitsvertrag, A-Id in Mitarbeiter
- Für **komplex-komplexe Beziehungsmengen** benötigt man eine eigene Tabelle, die zwei Fremdschlüsselfelder enthält.
  - ⇨ Beispiel: Beziehungsmenge Teilnahme als Tabelle mit Fremdschlüsselfeldern M-Id und P-Nr
- Fremdschlüsselfelder haben denselben Felddatentyp wie die zugehörigen Primärschlüsselfelder; bei „Autowerten“ als Primärschlüssel verwendet man den Felddatentyp Zahl.
- Access gibt Feldern vom Datentyp Zahl den Standardwert 0. Fremdschlüssel mit Wert 0 machen meist Probleme, da es dazu in der Regel keinen passenden Primärschlüssel gibt. Deshalb **in Fremdschlüsselfeldern Standardwert löschen**!

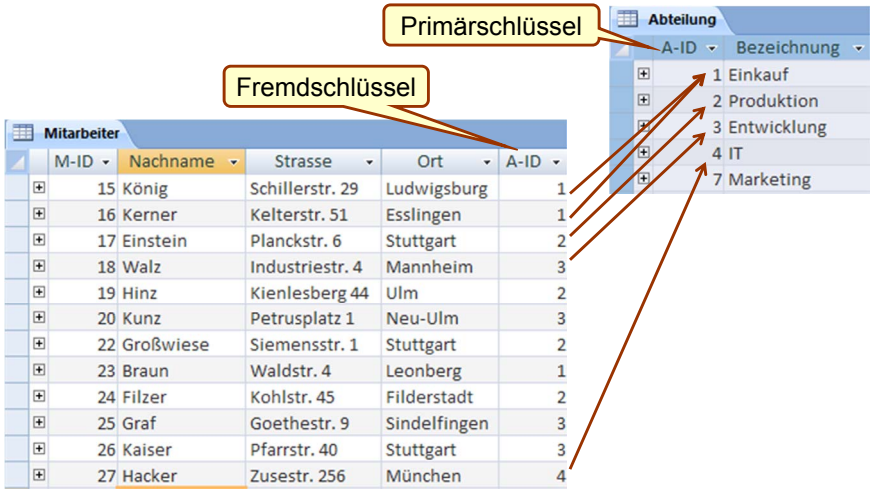
REFERENZIELLE INTEGRITÄT

HOCHSCHULE DER MEDIEN

- Fremdschlüssel dienen dazu, um auf andere Datensätze zu verweisen.
- Die Werte der Fremdschlüssel dienen dabei als Verweise, auch „Referenzen“ genannt.
- Ein Verweis auf einen Datensatz wird hergestellt, indem dessen Primärschlüsselwert in das Fremdschlüsselfeld eingetragen wird.
- Diese Verweise sollen sinnvollerweise nicht ins Leere zeigen, d.h. **jeder Wert eines Fremdschlüssels sollte als Primärschlüsselwert in der referenzierten Tabelle (d.h. der Tabelle, auf die verwiesen wird) vorkommen**.
  - ⇨ Diese Eigenschaft nennt man **referenzielle Integrität.**

REFERENZEN AUF DATENSÄTZE MIT HILFE VON FREMDSCHLÜSSELN

HOCHSCHULE DER MEDIEN



PROBLEMATISCHE DARSTELLUNG VON REFERENZEN IN TABELLEN

HOCHSCHULE DER MEDIEN

Wenn Tabellen in Form von einfachen Datenblattansichten angezeigt werden, ist es schwer, die referenzielle Integrität aufrecht zu erhalten und zu überprüfen.
Dies gilt vor allem, wenn die Fremdschlüssel keine sprechenden Schlüsselwerte (z.B. Abkürzungen) enthalten sondern künstliche Schlüsselwerte (z.B. automatisch generierte Nummern).

- Bei Ansicht der Fremdschlüsselwerte ist schwer nachzuvollziehen, welcher Datensatz gemeint ist.
  - ⇨ Die Einhaltung der referenziellen Integrität lässt sich nur schwer überprüfen.
- Bei der Eingabe von Fremdschlüsselwerten muss man laufend in der referenzierten Tabelle nachschlagen.
  - ⇨ Die referenzielle Integrität lässt sich bei der Dateneingabe nur mit Mühe gewährleisten.

NACHSCHLAGEFUNKTION FÜR FREMDSCHLÜSSEL

HOCHSCHULE DER MEDIEN

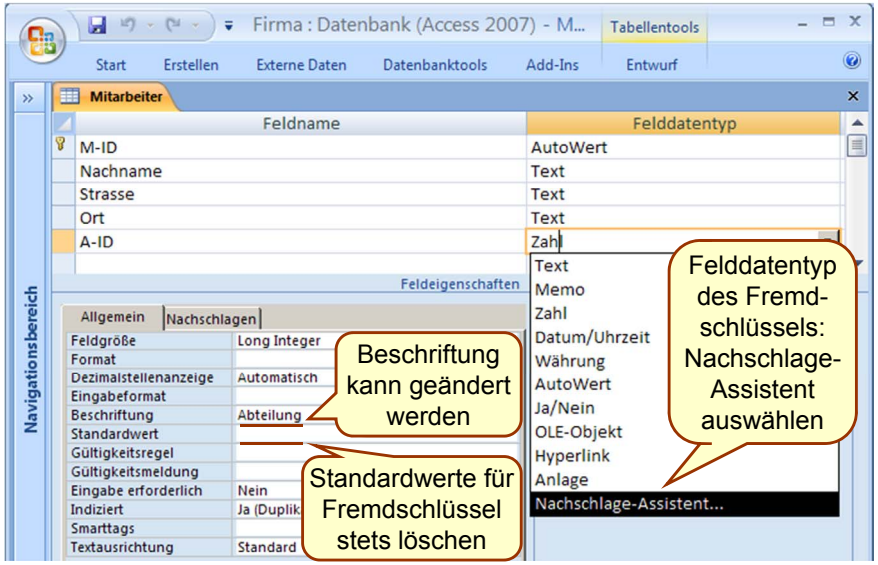
Die Eingabe und Anzeige der Fremdschlüsselwerte und damit die Einhaltung der referenziellen Integrität lässt sich durch Verwendung der Nachschlagefunktion unterstützen:

- Statt des Schlüssels (z.B. *A-Id* aus der Tabelle *Abteilung*) kann ein Merkmal des referenzierten Datensatzes angezeigt werden (z.B. *Bezeichnung* aus *Abteilung*). Achtung: **Intern enthält die Tabelle weiterhin die Schlüsselwerte.**
- Die Beschriftung lässt sich ändern (z.B. zu *Abteilung*)
- Statt den Schlüssel einzugeben, gibt man das betreffende Merkmal aus dem referenzierten Datensatz an (z.B. die Bezeichnung der Abteilung) oder man wählt den entsprechenden Merkmalswert einem Menü (Liste aller Abteilungsbezeichnungen) aus.

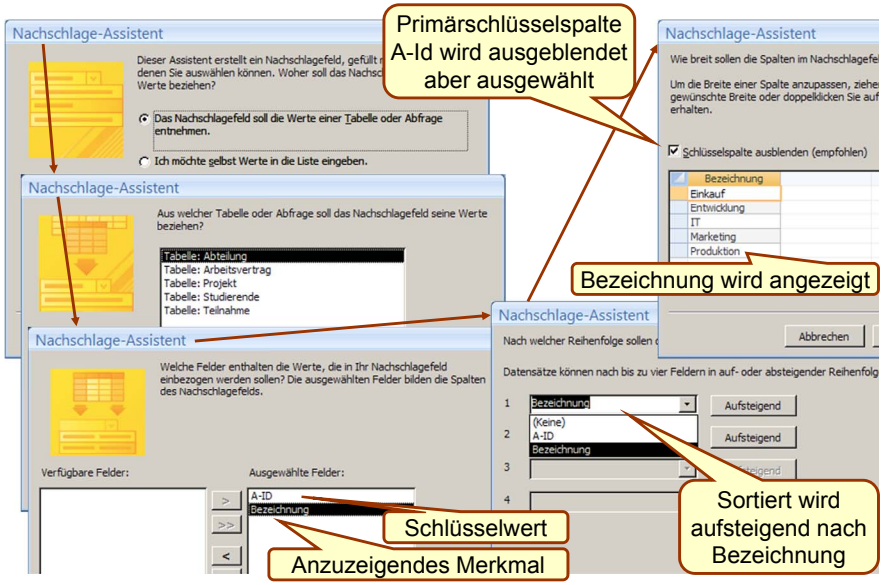
NACHSCHLAGEFUNKTION FÜR FREMDSCHLÜSSEL: BEISPIEL

HOCHSCHULE DER MEDIEN

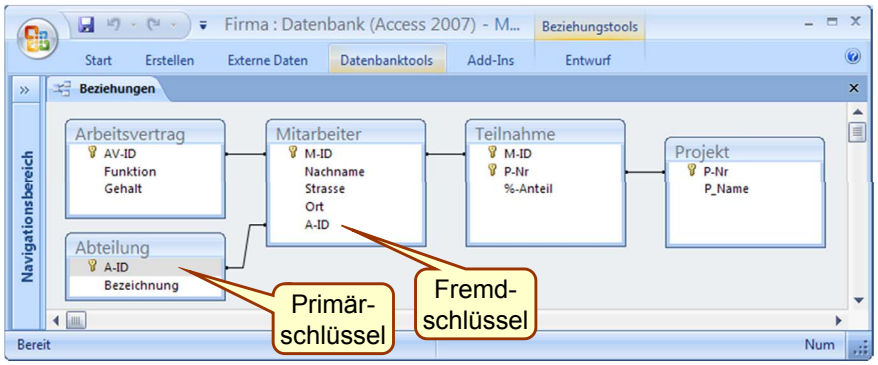
EINRICHTUNG DER NACHSCHLAGEFUNKTION



DER NACHSCHLAGE-ASSISTENT

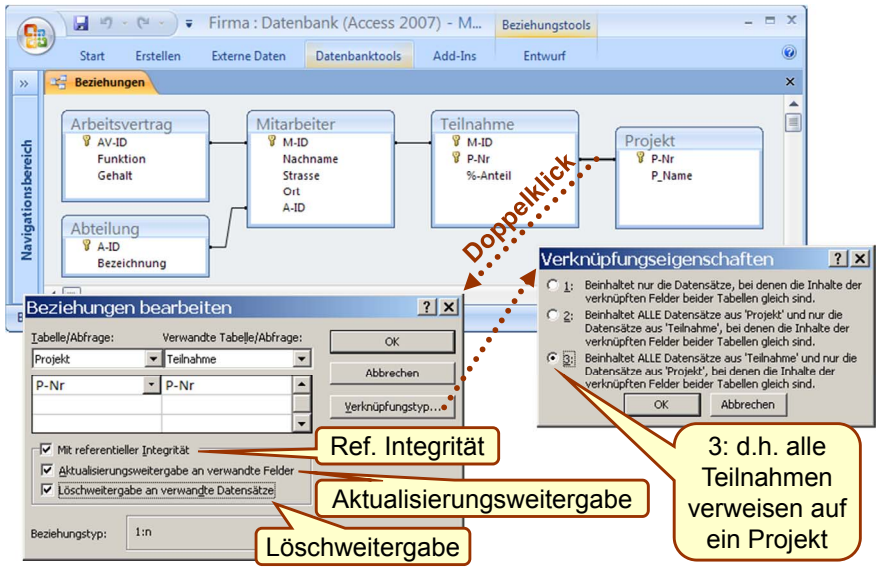


ANSICHT BEZIEHUNGEN: ANZEIGE VON VERKNÜPFUNGEN IN ACCESS

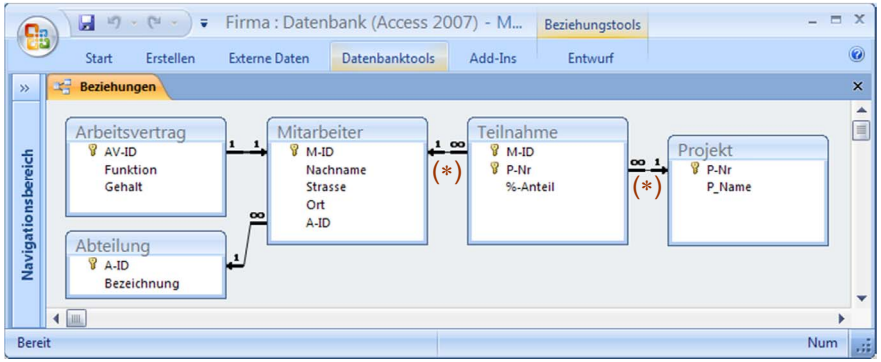


Ansicht Beziehungen, erreichbar über Menü „Datenbanktools – Beziehungen“. Die Grafik zeigt das **Datenmodell** mit der **Verknüpfung** von Fremd- und Primärschlüsseln.
Achtung: Dies ist **kein** Entitäten-Beziehungsmodell! Microsoft verwendet den Begriff Beziehungen in einem eigenen Sinn!

REFERENZIELLE INTEGRITÄTSBEDINGUNGEN



DEFINITION VON BEZIEHUNGSEIGENSCHAFTEN (1)



Alle Verknüpfungen wurden mit referenzieller Integrität und mit Aktualisierungsweitergabe definiert.

Die mit der Tabelle Teilnahme verbundenen Verknüpfungen (*) zusätzlich mit Löschweitergabe.

DEFINITION VON BEZIEHUNGSEIGENSCHAFTEN (2)

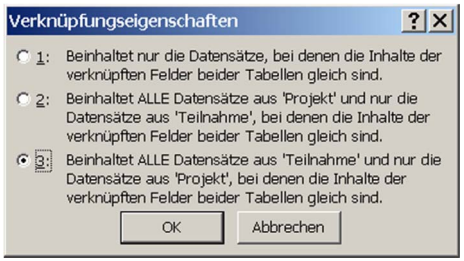
- **Achtung:** Access versteht unter „Beziehung“ jede Verknüpfung zwischen Tabellen über Schlüssel. Der Begriff wird also ein klein wenig anders verwendet als in einem Entitäten-Beziehungsmodell. Deshalb sollte man Beziehungen im Sinne von Access besser als **„Verknüpfungen“** bezeichnen.
- Praktisch immer müssen sog. 1:n-Verknüpfungen (in Grafik: 1-∞) verwendet werden. (Seltene Ausnahme: 1:1-Verknüpfung zwischen Primärschlüsseln)
- In Access gibt es keine „n:m-Verknüpfungen“. Da komplex-komplexe Beziehungsmengen durch eine Zwischentabelle dargestellt werden, treten diese in Access als zwei Verknüpfungen vom Typ 1-∞ in Erscheinungen.

DEFINITION VON BEZIEHUNGSEIGENSCHAFTEN (3)

- Es ist empfehlenswert, **referenzielle Integrität** zu verlangen.
  - ⇨ Tipp: Standardwerte für Fremdschlüssel vertragen sich meist nicht mit referenzieller Integrität, sie sollten im Tabellenentwurf gelöscht werden.
  - ⇨ Nullwerte (nicht besetzte Werte) in Fremdschlüsseln werden auch bei ref. Integrität von Access toleriert
- **Aktualisierungsweitergabe** hat Vorteile, falls sich je Primärschlüsselwerte ändern sollten. Diese Änderung wird dann in den entsprechenden Fremdschlüsseln nachgeführt. Allerdings ist es nicht empfehlenswert, Primärschlüssel zu ändern; daher ist diese Einstellung nicht so relevant.
- **Löschweitergabe** ist nützlich bei Beziehungsmengen-tabellen wie *Teilnahme*. Falls eine Entität gelöscht wird, wird auch der zugehörige Beziehungsdatensatz gelöscht.

VERKNÜPFUNGSEIGENSCHAFTEN

Bei den so genannten Verknüpfungseigenschaften gibt es die Wahl zwischen drei Möglichkeiten:



1 wird gewählt, wenn die Verknüpfung nur fallweise besteht.

2 bzw. 3, wenn allen Datensätzen einer Tabelle jeweils ein Datensatz der anderen Tabelle zugeordnet ist. Im Beziehungenfenster wird die Zuordnung dann als Pfeil vom Fremdschlüssel zum Primärschlüssel angezeigt, wobei die Fälle 2 und 3 für unterschiedliche Pfeilrichtungen stehen.

Das alles ist hier noch nicht so sehr wichtig und hat lediglich Bedeutung als Voreinstellung für die Definition von Abfragen (siehe Teil 3).

NORMALFORMEN (NORMAL FORMS)

HOCHSCHULE DER MEDIEN

- Eine Tabelle enthält **Redundanzen**, wenn einzelne (**redundante**) Datenwerte innerhalb der Tabelle ohne Informationsverlust weggelassen werden können.
- Redundanzen in Tabellen bewirken **Anomalien**, die die korrekte Repräsentation von Informationen in Form von Daten erschweren.
- Abhilfe: Einhaltung sogenannter **Normalformen** (engl.: *normal forms*)
  - ⇨ Untersuchung von Abhängigkeiten zwischen Feldern einer Tabelle
  - ⇨ Einhaltung bestimmter Regeln bei der Definition von Tabellen, ggf. Modifikation der Tabellenstruktur

Hinweis: Dieser Teil des Skripts bis Teil 3 kann übersprungen und zu einem späteren Zeitpunkt behandelt werden

NORMALFORMEN: ÜBERSICHT

HOCHSCHULE DER MEDIEN

1. Normalform (1NF): „Wertebereiche der Merkmale sind atomar“ (d.h. keine zusammengesetzten Werte)

2. Normalform (2NF): 1NF und „Nichtschlüsselmerkmale sind von **allen** Schlüsselmerkmalen voll funktional abhängig“

3. Normalform (3NF): 2NF und „kein Nichtschlüsselmerkmal ist von irgendeinem Schlüssel transitiv abhängig“

Aufzählung der weiteren Normalformen (zur Vollständigkeit):

Boyce-Codd-Normalform (BCNF): Weitere Verschärfung der 3. Normalform

4. Normalform (4NF): Berücksichtigung sog. „mehrwertiger Abhängigkeiten“

5. Normalform (5NF): Äußerste Möglichkeit der Aufgliederung in einzelne Tabellen ohne Informationsverlust

EINHALTUNG DER NORMALFORMEN

HOCHSCHULE DER MEDIEN

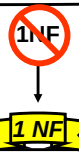
- Praktische Bedeutung besitzen vor allem die ersten drei Normalformen.
- Behandlung der ersten drei Normalformen in dieser Vorlesung in Form von Negativbeispielen und Anleitung zur Mängelbeseitigung. (Die genaue Definition der ersten drei Normalformen ist kein „Pflichtstoff“ dieser Vorlesung, wohl aber das Erkennen der Negativbeispiele und die Fähigkeit, deren Mängel zu beheben.)
- Die weiteren Normalformen werden in dieser Vorlesung nicht behandelt.
- Ein sauberer Entwurf eines Entitäten-Beziehungsmodells und die Beachtung der Abbildungsregeln bewirkt praktisch automatisch die Einhaltung der Normalformen.

1. NORMALFORM: NUR ATOMARE MERKMALSWERTE SIND ERLAUBT

HOCHSCHULE DER MEDIEN

Mitarbeiterkonto

Bankverbindung			M-Id	Mitarbeitername
123 456 234 567	Sparkasse Ulm Postbank Stgt.	(630 500 00) (600 100 70)	101	Abele
987 654 876 543	Sparda Stgt.	(600 908 00)	102	Bauer



1NF = darstellbar in Form einer Tabelle + Wertebereiche der Merkmale sind **atomar** (d.h. keine zusammengesetzten Werte)

Mitarbeiterkonto

Kontonr	BLZ	Geldinstitut	M-Id	Mitarbeitername
123 456	630 500 00	Sparkasse Ulm	101	Abele
234 567	600 100 70	Postbank Stgt.	101	Abele
987 654	600 908 00	Sparda Stgt.	102	Bauer
876 543	600 908 00	Sparda Stgt.	102	Bauer

2. NORMALFORM: ABHÄNGIG VON ALLEN SCHLÜSSELMERKMALEN

HOCHSCHULE DER MEDIEN

Mitarbeiterkonto

Kontonr	BLZ	Geldinstitut	M-Id	Mitarbeitername
123 456	630 500 00	Sparkasse Ulm	101	Abele
234 567	600 100 70	Postbank Stgt.	101	Abele
987 654	600 908 00	Sparda Stgt.	102	Bauer
876 543	600 908 00	Sparda Stgt.	102	Bauer

2NF

2NF = 1NF + Nichtschlüsselmerkmale sind von **allen** Schlüsselmerkmalen voll funktional abhängig. (Im Negativbeispiel war Geldinstitut nur von BLZ und nicht von Kontonr abhängig.)

2 NF

Mitarbeiterkonto

Kontonr	BLZ	M-Id	Mitarbeitername
123 456	630 500 00	101	Abele
234 567	600 100 70	101	Abele
987 654	600 908 00	102	Bauer
876 543	600 908 00	102	Bauer

2 NF

Geldinstitut

BLZ	Name
630 500 00	Sparkasse Ulm
600 100 70	Postbank Stgt.
600 908 00	Sparda Stgt.

3. NORMALFORM: NICHT ABHÄNGIG ÜBER UMWEGE

HOCHSCHULE DER MEDIEN

Mitarbeiterkonto

Kontonr	BLZ	M-Id	Mitarbeitername
123 456	630 500 00	101	Abele
234 567	600 100 70	101	Abele
987 654	600 908 00	102	Bauer
876 543	600 908 00	102	Bauer

Geldinstitut

BLZ	Geldinstitut
630 500 00	Sparkasse Ulm
600 100 70	Postbank Stgt.
600 908 00	Sparda Stgt.

3NF

3NF = 2NF + **Kein** Nichtschlüsselmerkmal ist von irgendeinem Schlüssel transitiv (d.h. **über Umwege**) **abhängig**. (Im Negativbeispiel ist dies über den Umweg M-Id der Fall)

3 NF

Mitarbeiterkonto

Kontonr	BLZ	M-Id
123 456	630 500 00	101
234 567	600 100 70	101
987 654	600 908 00	102
876 543	600 908 00	102

3 NF

Geldinstitut

BLZ	Name
630 500 00	Sparkasse Ulm
600 100 70	Postbank Stgt.
600 908 00	Sparda Stgt.

3 NF

Mitarbeiter

M-Id	Name
101	Abele
102	Bauer

NORMALFORMEN: KONSEQUENZEN (1)

HOCHSCHULE DER MEDIEN

- Die Einhaltung von Normalformen führt zu einer Verteilung der Sachverhalte auf mehrere Tabellen.
- Nichteinhaltung der Normalformen ist in der Regel dadurch verursacht, dass kein korrektes Entitäten-Beziehungsmodell gebildet wurde oder dieses nicht regelgemäß in Tabellen umgesetzt wurde.
  Oft wurde eine eigenständige Entitätsmenge übersehen.
- Manchmal gibt es Grenzfälle, in denen unklar ist, ob eine bestimmte Information durch ein einfaches Merkmal dargestellt werden kann oder eine eigene Entitätsmenge vorgesehen werden muss. Oft möchte man eine weitere Entitätsmenge und damit eine weitere Tabelle vermeiden.
  - ⇨ Unter Umständen ist das möglich (siehe hierzu folgende Folie).

NORMALFORMEN: KONSEQUENZEN (2)

HOCHSCHULE DER MEDIEN

Beispiel: In einer Personendatenbank soll den Personen eine Nationalität („deutsch“, „englisch“ usw.) zugeordnet werden.

- Soll die Nationalität als einfaches Merkmal vom Typ „Text“ oder als eine eigene Entitätsmenge dargestellt werden?

Die Antwort fällt differenziert aus:

- Wenn über die Nationalität nicht mehr gespeichert wird als ihr Name, genügt grundsätzlich die Darstellung als Merkmal.
- Soll aber über die Nationalität mehr gespeichert werden (z.B. Name=„deutsch“, Kürzel=„DE“), dann benötigt man eine eigene Entitätsmenge für die Nationalitäten. Dies hat auch Vorteile für die Einrichtung einer Nachschlagefunktion.
- Wenn eine Person mehrere Nationalitäten haben kann, benötigt man in jedem Fall eine eigene Entitätsmenge für Nationalitäten sowie eine komplex-komplexe Beziehungsmenge zwischen Personen und Nationalitäten.

TEIL 3: NUTZUNG VON DATEN-BANKEN - INHALTSÜBERSICHT

HOCHSCHULE DER MEDIEN

NUTZUNG VON DATENBANKEN: DATENMANIPULATION

HOCHSCHULE DER MEDIEN

- Die Nutzung von Datenbanken beruht auf dem Vorgang der **Datenmanipulation**. Dabei wird der **Inhalt** einer Datenbank **angezeigt** oder **verändert**.
- Datenmanipulation steht im Gegensatz zur Datendefinition, die Gegenstand des vorigen Vorlesungsabschnitts war. Zur Erinnerung: Unter dem Vorgang der **Datendefinition** verstehen wir die Festlegung des sogenannten **Datenbankschemas**, das die **Struktur** einer Datenbank beschreibt.
- Außer für die Nutzung von Datenbanken wird Datenmanipulation auch für den **Aufbau von Datenbanken** benötigt, und zwar für den 4. Schritt, das erstmalige **Füllen einer Datenbank mit Inhalten**.

DATENMANIPULATION

HOCHSCHULE DER MEDIEN

Unter Datenmanipulation verstehen wir:

- Das **Abrufen** (retrieval) von Daten aus der Datenbank
- Das **Ändern** (modification) von Daten in der Datenbank
- Das **Eintragen** (insertion) von Daten in die Datenbank
- Das **Löschen** (deletion) von Daten in der Datenbank

Moderne relationale Datenbanksysteme bieten zwei Möglichkeiten der Datenmanipulation an:

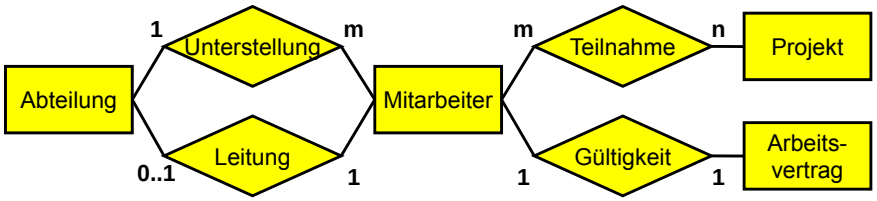
- **Interaktive Benutzungsoberflächen**: Datenblattansichten von Tabellen, Formulare, Berichte, Abfragen
- **Datenbankabfragesprache**: SQL (Structured Query Language)

BENUTZUNGSOBERFLÄCHEN ZUR DATENMANIPULATION

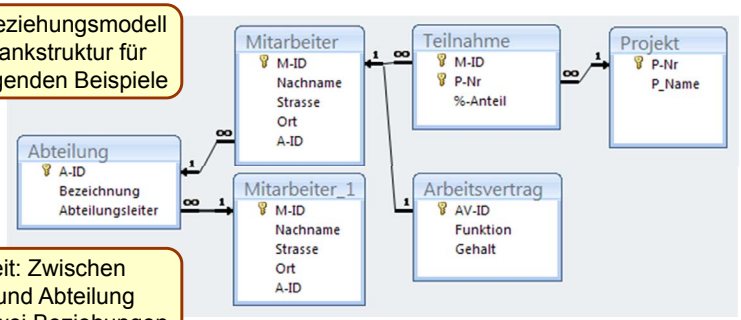
HOCHSCHULE DER MEDIEN

- **Datenblattansichten von Tabellen**: Editierbare Tabellen zum Anzeigen und Ändern von Tabelleninhalten
- **Formulare**: Anzeigen und Ändern von Tabelleninhalten
  - ⇨ Graphisch ansprechende Formulardarstellung
  - ⇨ Die Inhalte zusammengehöriger Tabellen können in einem Formular mit Unterformular dargestellt werden
- **Berichte**: Ähnlich wie Formulare, aber nur zum Anzeigen
- **Abfragen**:
  - ⇨ **Auswahlabfragen** zur (vorübergehenden) Erzeugung neuer Tabellen aus den vorhandenen durch Auswahl und Kombination nach bestimmten Kriterien
  - ⇨ **Aktionsabfragen** zum Ändern, Eintragen und Löschen von Datensätzen nach bestimmten Kriterien

DATENBANKSTRUKTUR FÜR DIE FOLGENDEN BEISPIELE

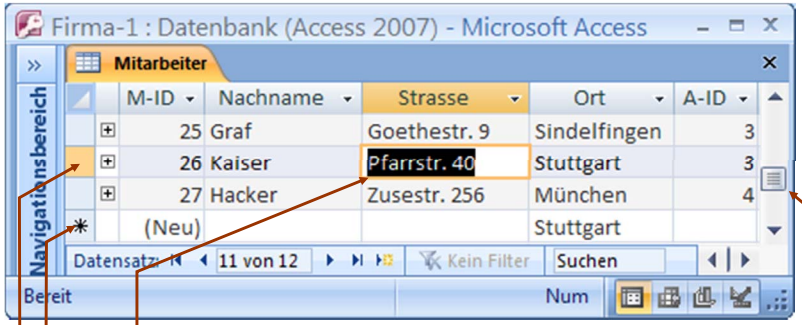


Entitäten-Beziehungsmodell und Datenbankstruktur für die nachfolgenden Beispiele



Besonderheit: Zwischen Mitarbeiter und Abteilung bestehen zwei Beziehungen

DATENBLATTANSICHT EINER TABELLE



Daten **abrufen** über ein rollbares Fenster

Daten **ändern** durch Überschreiben

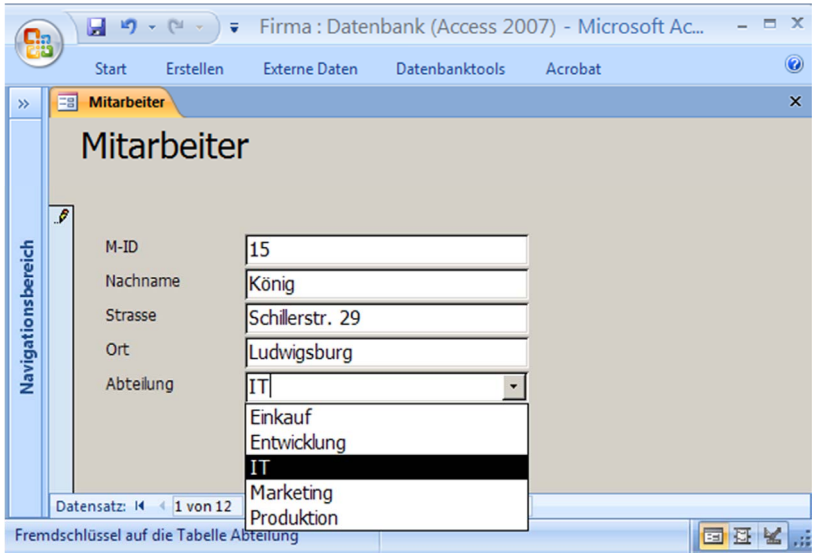
Daten **einfügen** durch direkten Eintrag

Daten **löschen** durch Selektieren und Betätigen der Löschtaste bzw. Menüauswahl Löschen

EIGENSCHAFTEN DER DATENBLATTANSICHT:

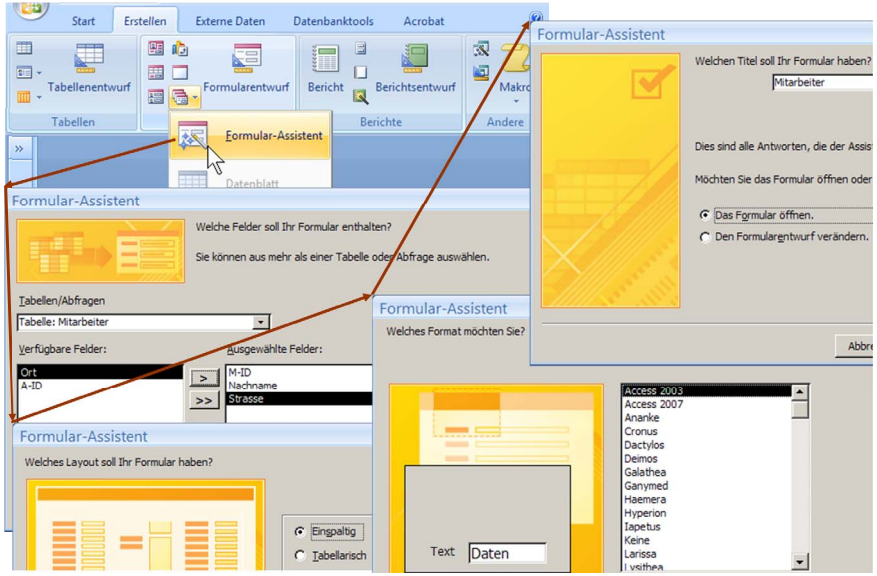
- Die Datenblattansicht eignet sich gut für Felder, die Texte und Zahlen als Datenwerte haben:
  - ⇨ Die Anzeige ist in der Regel intuitiv verständlich.
  - ⇨ Ändern ist einfach durch Überschreiben möglich.
- Probleme kann die Darstellung der Fremdschlüssel bereiten, vor allem, wenn diese keine sprechenden Schlüsselwerte (z.B. Abkürzungen) enthalten sondern künstliche Schlüsselwerte (z.B. automatisch generierte Nummern)
  - ⇨ Es ist schwer nachzuvollziehen, welcher Datensatz gemeint ist
  - ⇨ Hier ist Abhilfe möglich durch die Nachschlagefunktion (siehe vorangehendes Kapitel – referenzielle Integrität)

EIN EINFACHES FORMULAR ZUR DATENERFASSUNG



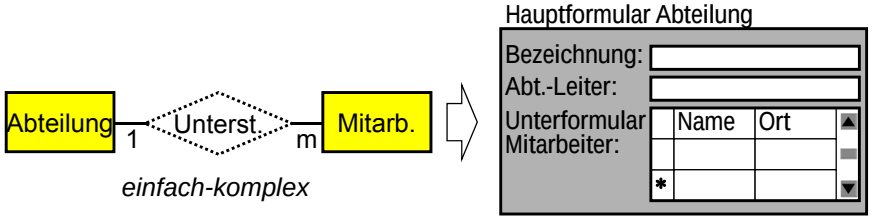
ENTWURF EINES EINFACHEN FORMULARS

HOCHSCHULE DER MEDIEN



EINFACH-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (1)

HOCHSCHULE DER MEDIEN

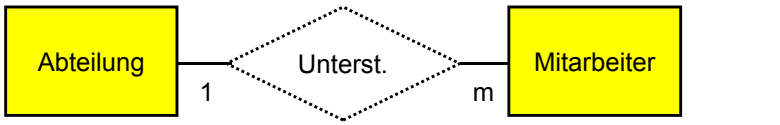


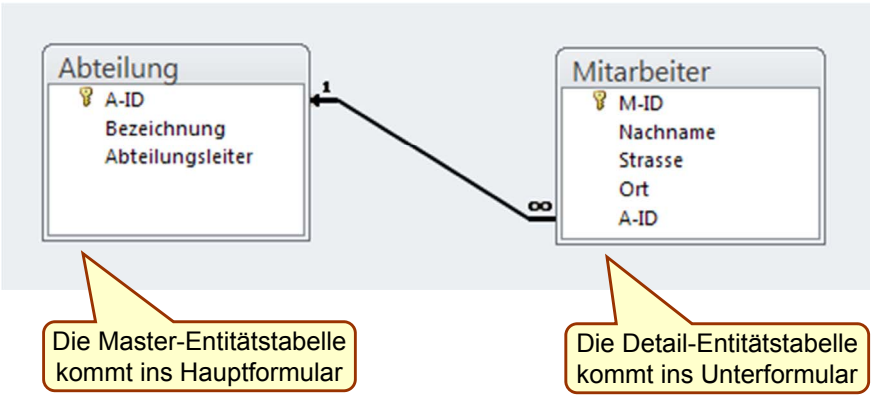
Für die Bearbeitung einfach-komplexer Beziehungen benötigt man komplexe Formulare (Formulare mit Unterformularen):

- sogenannte Master-Detailbeziehung,
  Beispiel: Abteilung - Mitarbeiter
- Die „Master"-Entitätstabelle (Abteilung) ins Hauptformular
- Die „Detail"-Entitätstabelle (Mitarbeiter) ins Unterformular
- Komplexe Formulare können mit „Formularassistenten" (Register „Erstellen" – „Weitere Formulare") definiert werden.

EINFACH-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (2)

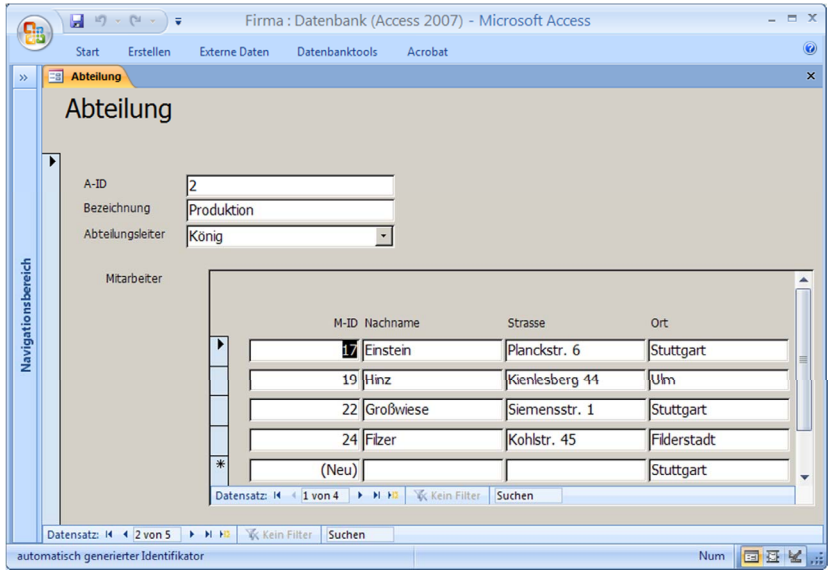
HOCHSCHULE DER MEDIEN





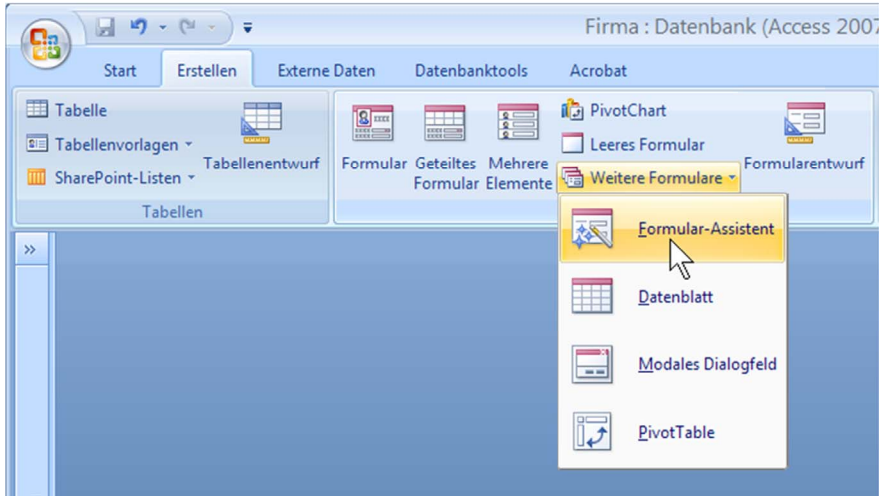
MASTER-DETAIL-BEZIEHUNG ALS FORMULAR MIT UNTERFORMULAR

HOCHSCHULE DER MEDIEN



FORMULAR MIT UNTERFORMULAR: FORMULAR-ASSISTENT STARTEN

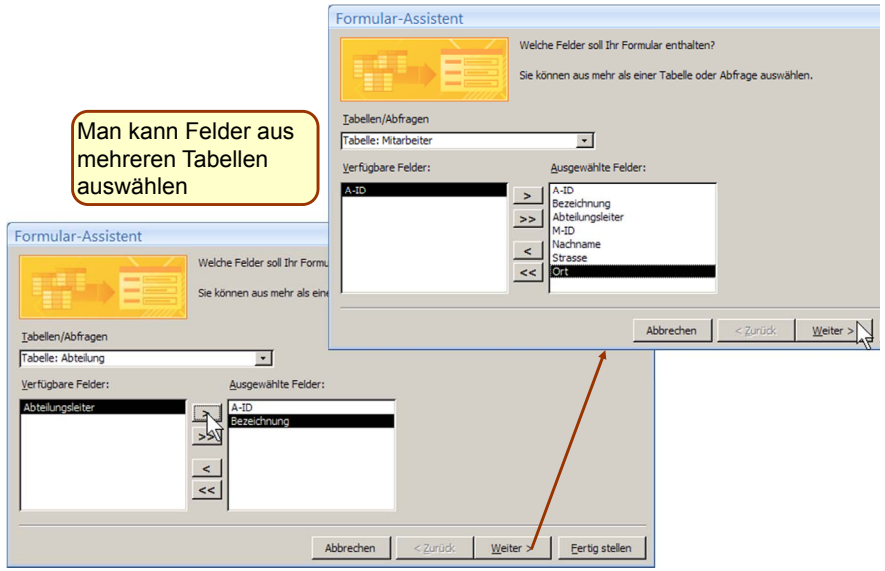
HOCHSCHULE DER MEDIEN



FORMULAR-ASSISTENT: FELDER AUSWÄHLEN

HOCHSCHULE DER MEDIEN

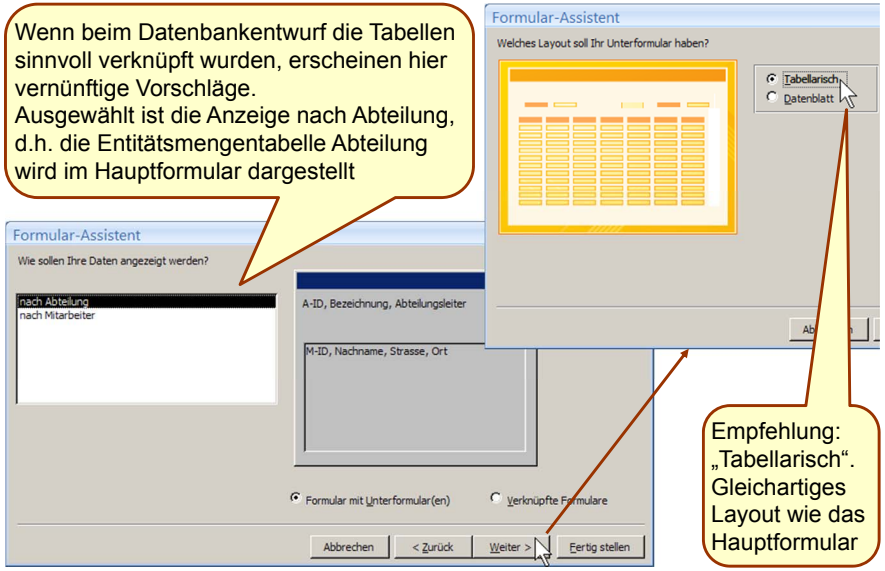
Man kann Felder aus mehreren Tabellen auswählen



FORMULAR-ASSISTENT: ANZEIGEART UND LAYOUT

HOCHSCHULE DER MEDIEN

Wenn beim Datenbankentwurf die Tabellen sinnvoll verknüpft wurden, erscheinen hier vernünftige Vorschläge.
Ausgewählt ist die Anzeige nach Abteilung, d.h. die Entitätsmengentabelle Abteilung wird im Hauptformular dargestellt

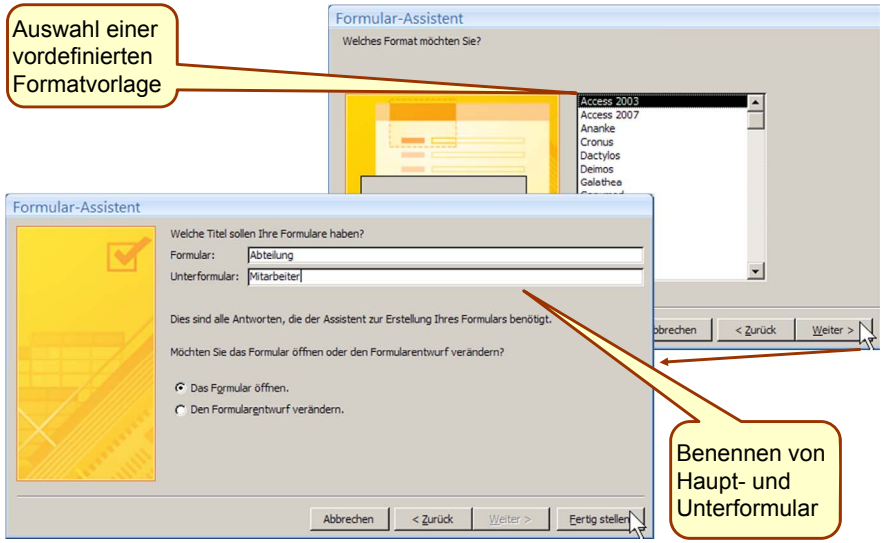


Empfehlung: „Tabellarisch". Gleichartiges Layout wie das Hauptformular

FORMULAR-ASSISTENT: FORMATVORLAGE UND NAMEN

HOCHSCHULE DER MEDIEN

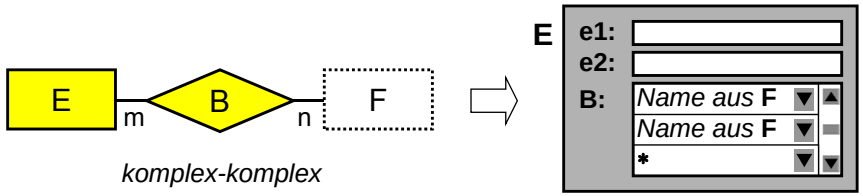
Auswahl einer vordefinierten Formatvorlage



Benennen von Haupt- und Unterformular

KOMPLEX-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (1)

HOCHSCHULE DER MEDIEN

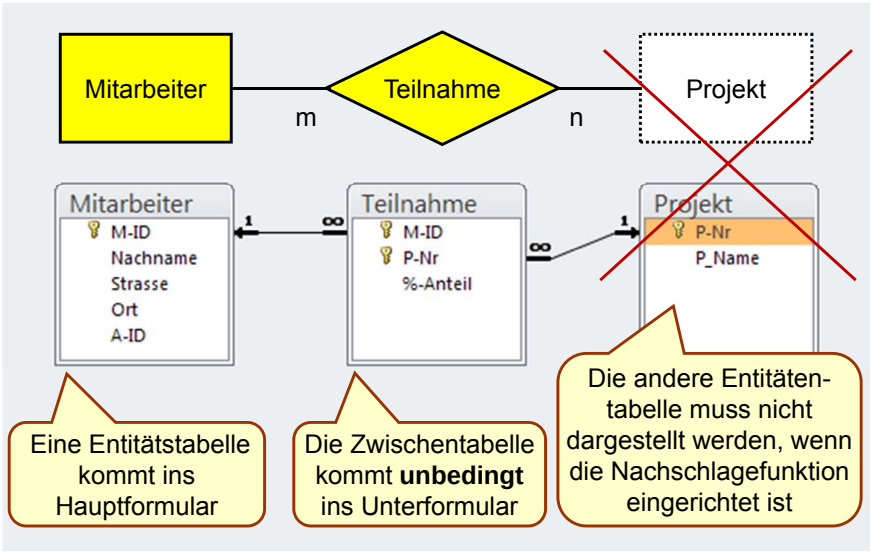


komplex-komplex

- Eine **Entitätstabelle (E)** wird im Hauptformular dargestellt.
- Wichtig: Die **Beziehungstabelle (B) muss unbedingt** im Unterformular dargestellt werden. Die andere Entitätstabelle (F) braucht **nicht** ins Formular aufgenommen zu werden.
- Wichtig ist, dass die **Nachschlagefunktion** in der Beziehungstabelle (B) eingerichtet ist. In Nachschlagefeldern können Beziehungen zur anderen Entitätsmenge (F) bearbeitet werden. Diese Beziehungen erscheinen in Form von Namen der Entitäten o.ä.

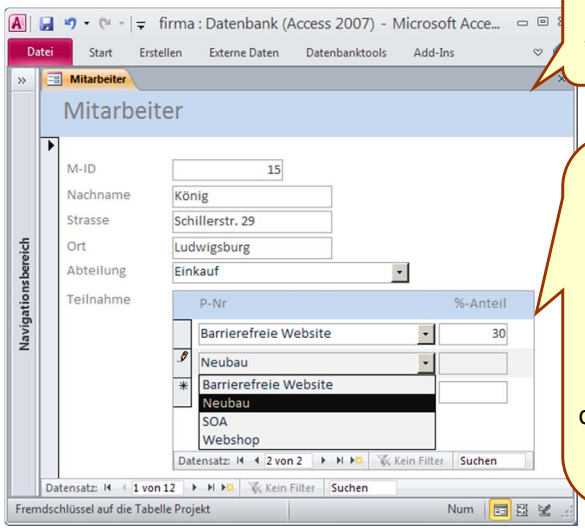
KOMPLEX-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (1)

HOCHSCHULE DER MEDIEN



KOMPLEX-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (2)

HOCHSCHULE DER MEDIEN

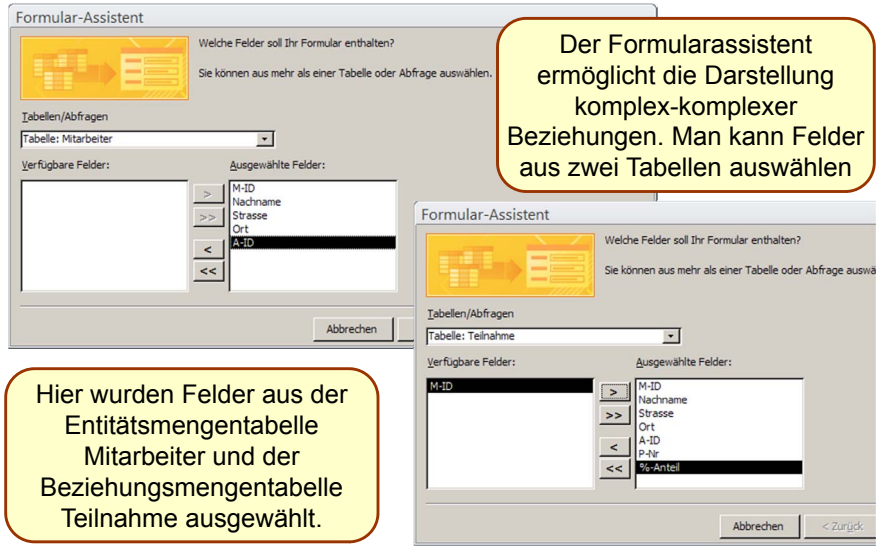


Eine Entitätstabelle, hier „Mitarbeiter“, kommt ins Hauptformular

Ins Unterformular kommt die komplex-komplexe Beziehungs-menge „Teilnahme“. Die in Beziehung stehende Entitäts-menge „Projekt“ wird nicht dargestellt. Durch die Nachschlagefunktion sind jedoch die Namen der Projekte sichtbar.

KOMPLEX-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (3)

HOCHSCHULE DER MEDIEN

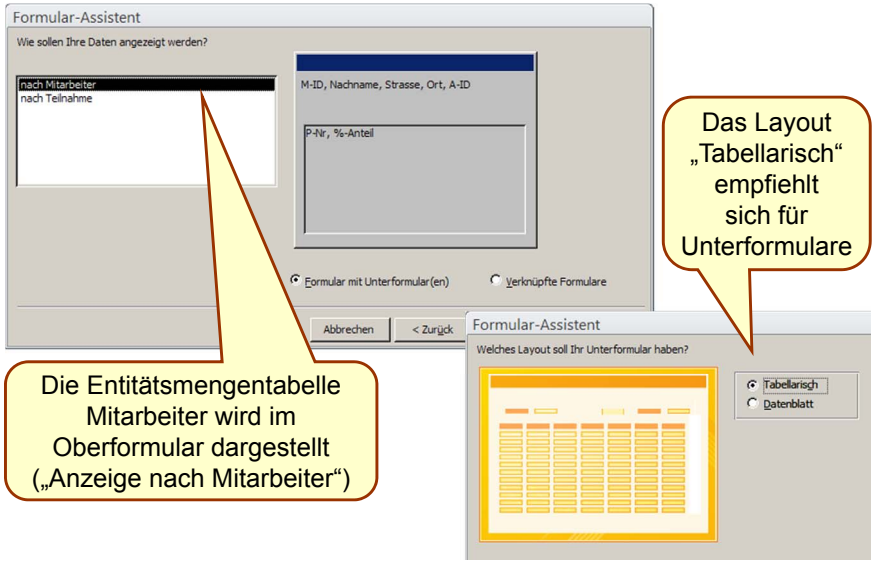


Der Formularassistent ermöglicht die Darstellung komplex-komplexer Beziehungen. Man kann Felder aus zwei Tabellen auswählen

Hier wurden Felder aus der Entitätsmengentabelle Mitarbeiter und der Beziehungsmengentabelle Teilnahme ausgewählt.

KOMPLEX-KOMPLEXE BEZIEHUNGEN IN FORMULAREN (4)

HOCHSCHULE DER MEDIEN



EMPFEHLUNGEN FÜR DAS ARBEITEN MIT FORMULAREN (1)

HOCHSCHULE DER MEDIEN

- Mit den Formularen **erst beginnen, wenn alle Nachschlagefelder definiert sind** (die Nachschlagefunktion wird dann auch in die Formulare übernommen.
- Mit Unterformularen können zwei Tabellen gleichzeitig dargestellt werden, die über Fremdschlüssel miteinander verknüpft sind:
  - ⇨ Einfach-komplexe Beziehungen: Die eine **Entitätsmenge**, der **„Master“** kommt ins Hauptformular, die andere **Entitätsmenge**, das **„Detail“** kommt ins Unterformular
  - ⇨ Komplex-komplexe Beziehungen: Eine **Entitätsmenge** kommt ins Hauptformular, die **Beziehungsmenge** kommt ins Unterformular (durch **Nachschlagefunktion** werden die Namen aus der anderen Entitätsmenge sichtbar), die andere Entitätsmenge selbst wird nicht angezeigt!

EMPFEHLUNGEN FÜR DAS ARBEITEN MIT FORMULAREN (2)

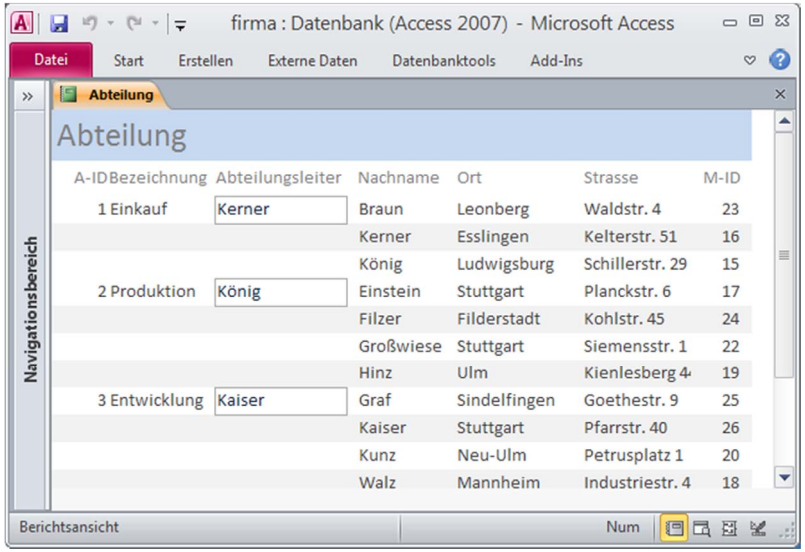
HOCHSCHULE DER MEDIEN

Detailverbesserungen in der Entwurfsansicht

- Eingabefelder: Ausrichten, Größe anpassen
- Schlüsselfelder, deren Werte sich automatisch ergeben (Primärschlüssel, Fremdschlüssel in Unterformularen) sollen nicht bearbeitbar sein oder ganz versteckt werden
  - ⇨ Eigenschaft „Aktiviert: Nein“ wählen
  - ⇨ Oder die Felder ganz aus dem Entwurf löschen
- In Formularen, die nur zur Ansicht dienen, ebenfalls die Eingabe sperren
- Eigenschaften des gesamten (Unter-)Formulars: Quadratisches Symbol links oben doppelklicken.
  - ⇨ Unterformulare als Endlosformular (nicht als Datenblatt)
  - ⇨ Navigationsschaltflächen evtl. löschen.

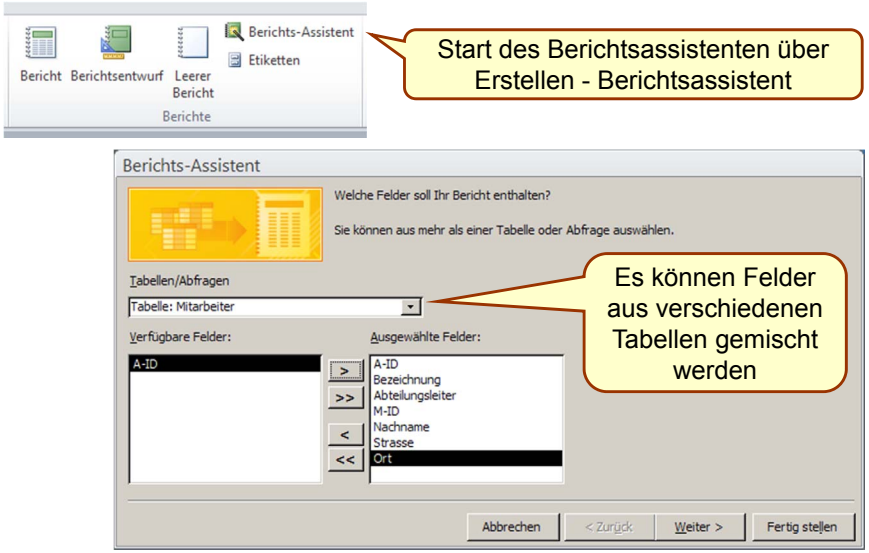
EIN BERICHT

HOCHSCHULE DER MEDIEN



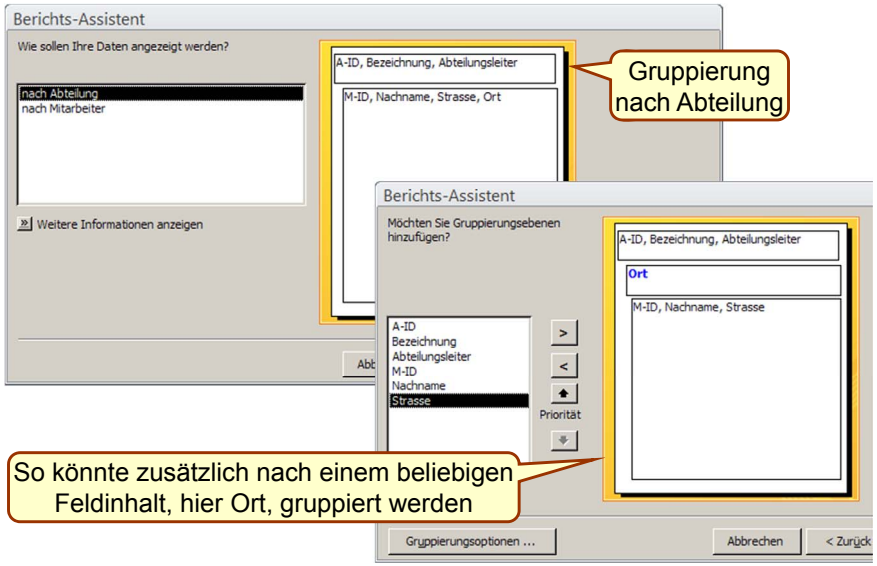
BERICHTE: ENTWURF MIT BERICHTSASSISTENT

HOCHSCHULE DER MEDIEN



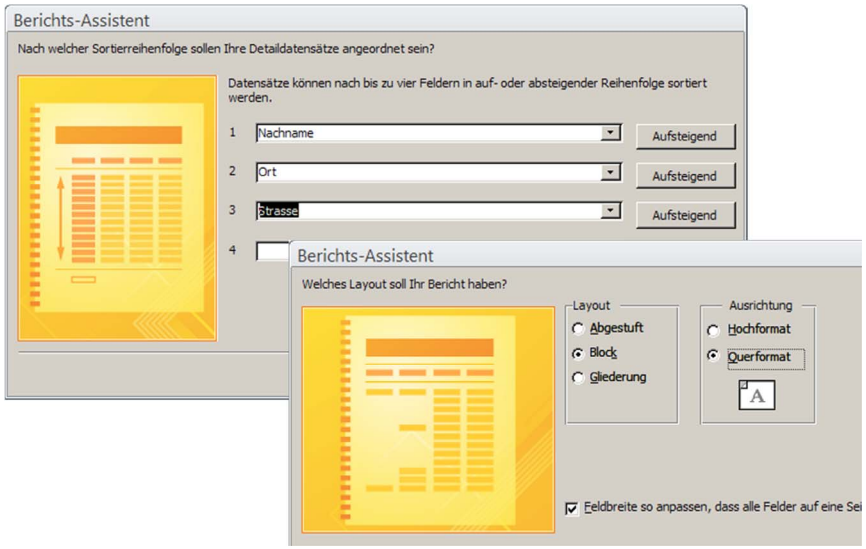
BERICHTE: GRUPPIERUNG VON BERICHTSELEMENTEN

HOCHSCHULE DER MEDIEN



BERICHTE: FESTLEGUNG VON SORTIERUNG UND LAYOUT

HOCHSCHULE DER MEDIEN



ARTEN VON ABFRAGEN AM BEISPIEL ACCESS

HOCHSCHULE DER MEDIEN

Es lassen sich folgende Arten von Abfragen unterscheiden:

- **Auswahlabfragen:**
  - ⇨ Daten aus ausgewählten Feldern einer oder mehrerer Tabellen werden nach bestimmten Kriterien selektiert.
  - ⇨ Das Ergebnis ist wieder eine Tabelle, die aber nur vorübergehende Existenz besitzt.
- **Aktionsabfragen:**
  Datenbankinhalte werden nach bestimmten Kriterien verändert. Wir unterscheiden:
  - ⇨ **Aktualisierungsabfragen:** Ändern von Datenwerten vorhandener Datensätze
  - ⇨ **Löschabfragen:** Löschen von Datensätzen
  - ⇨ **Anfügeabfragen:** Anfügen von neuen Datensätzen an vorhandene Tabellen
  - ⇨ **Tabellenerstellungsabfragen:** Wie Auswahlabfragen, aber Ergebnis wird als dauerhafte Tabelle gespeichert

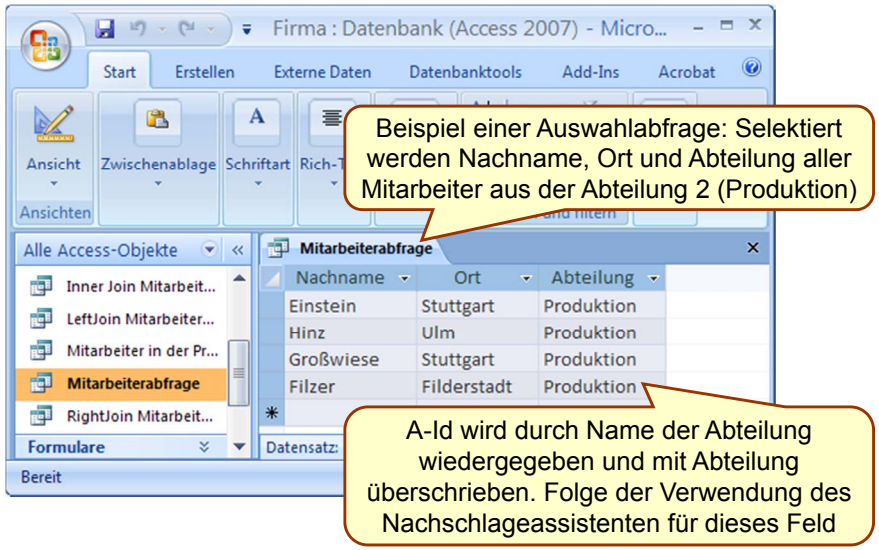
DEFINITION VON ABFRAGEN MIT „QUERY BY EXAMPLE“ (QBE)

HOCHSCHULE DER MEDIEN

- Abfragen werden heute oft ohne eine Abfragesprache wie SQL (Structured Query Language) definiert.
- Dies geschieht mit einer Technik, die den Namen „Query by Example“ trägt (QBE, vorgestellt von M.M. Zloof im Jahr 1977). Die Abfrage wird definiert, indem repräsentative Beispiele für die Lösung gegeben werden.
- Es können Auswahl- und Aktionsabfragen mit dieser Technik definiert werden.
- Auch Access unterstützt QBE.
- Bei der Verwendung von QBE wird intern eine SQL-Abfrage erzeugt, die man sich auch anschauen und die man sogar verändern kann.

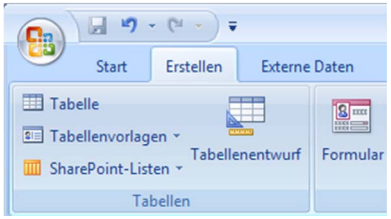
EINE AUSWAHLABFRAGE

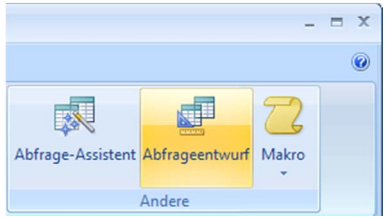
HOCHSCHULE DER MEDIEN



ABFRAGEN ERSTELLEN IN MICROSOFT ACCESS

HOCHSCHULE DER MEDIEN



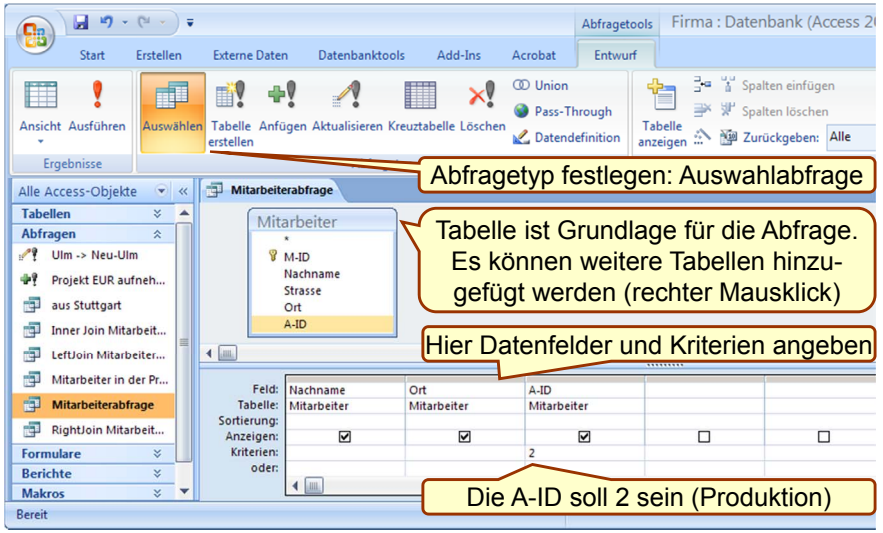


Abfragen in Access werden folgendermaßen erstellt:

- Erstellen – Abfrageentwurf
- Tabellen auswählen und hinzufügen
- Es können auch mehrere Tabellen hinzugefügt werden

ENTWURFSANSICHT EINER ABFRAGE

HOCHSCHULE DER MEDIEN

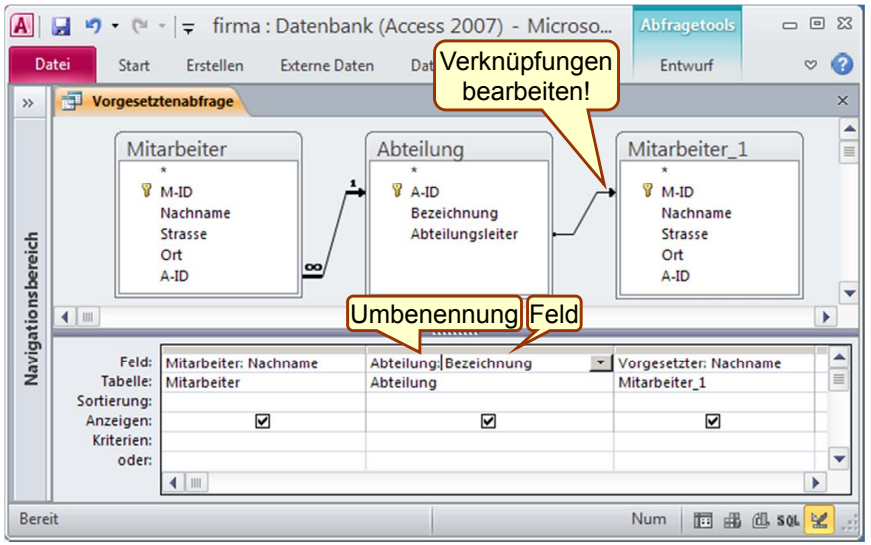


VORGESETZTER: EINE ABFRAGE ÜBER MEHRERE TABELLEN.

HOCHSCHULE DER MEDIEN

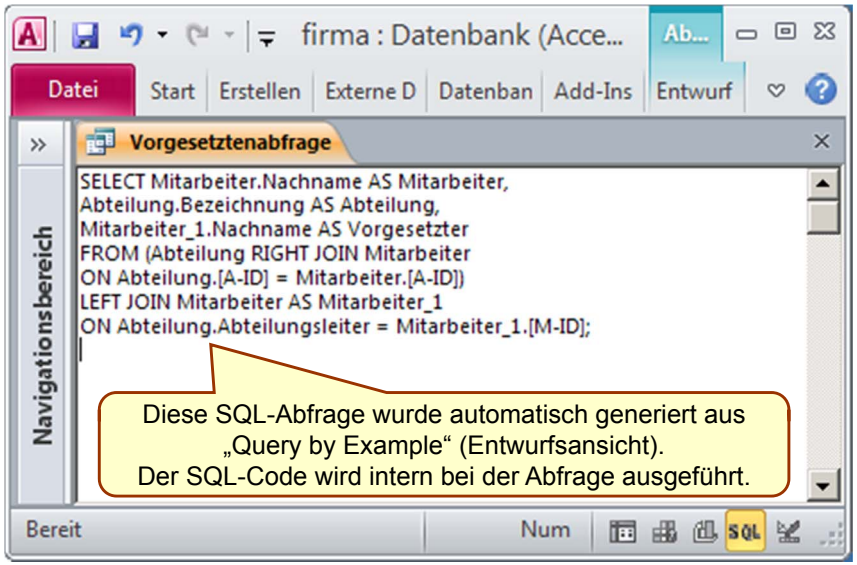
VORGESETZTER: ENTWURFSANSICHT DER ABFRAGE

HOCHSCHULE DER MEDIEN



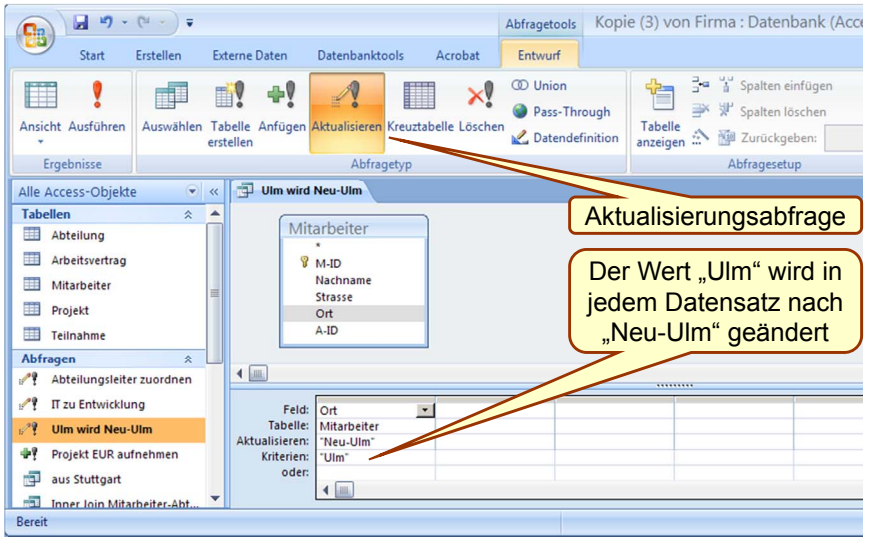
VORGESETZTER: SQL-ANSICHT DER ABFRAGE

HOCHSCHULE DER MEDIEN

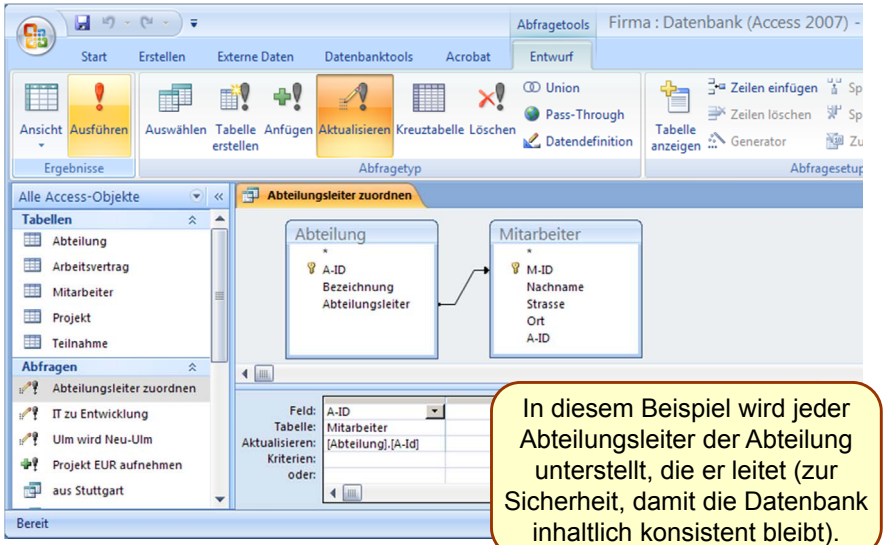


EINE AKTUALISIERUNGSABFRAGE

HOCHSCHULE DER MEDIEN

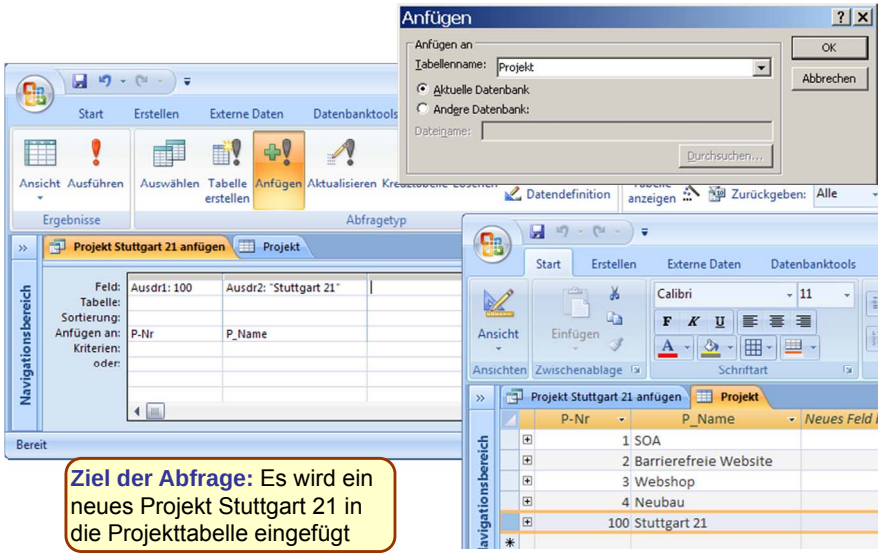


EINE KOMPLEXERE AKTUALISIERUNGSABFRAGE



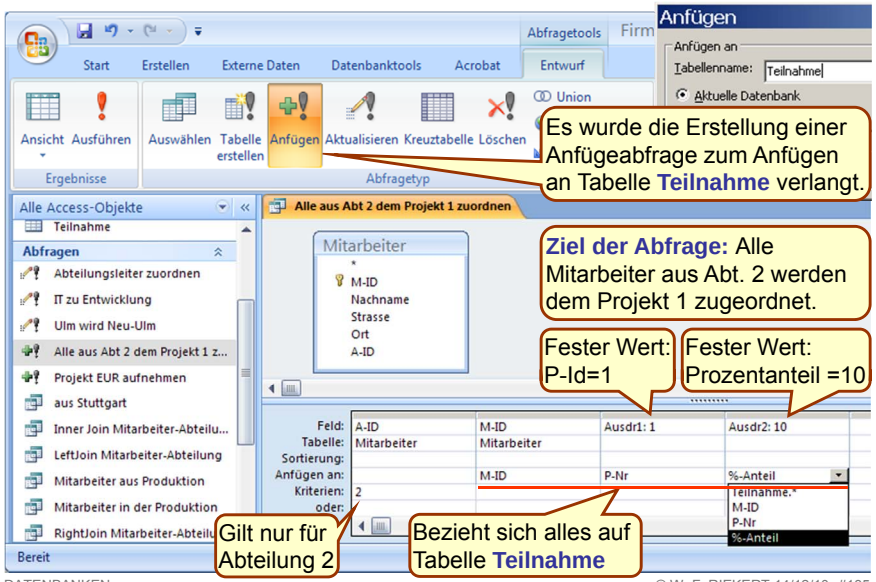
In diesem Beispiel wird jeder Abteilungsleiter der Abteilung unterstellt, die er leitet (zur Sicherheit, damit die Datenbank inhaltlich konsistent bleibt).

EINE EINFACHE ANFÜGEABFRAGE



Ziel der Abfrage: Es wird ein neues Projekt Stuttgart 21 in die Projekttabelle eingefügt

EINE ANFÜGEABFRAGE: BEISPIEL PROJEKTZUORDUNG



Es wurde die Erstellung einer Anfügeabfrage zum Anfügen an Tabelle **Teilnahme** verlangt.

Ziel der Abfrage: Alle Mitarbeiter aus Abt. 2 werden dem Projekt 1 zugeordnet.

Fester Wert: P-Id=1

Fester Wert: Prozentanteil =10

Gilt nur für Abteilung 2

Bezieht sich alles auf Tabelle **Teilnahme**

STRUCTURED QUERY LANGUAGE (SQL)

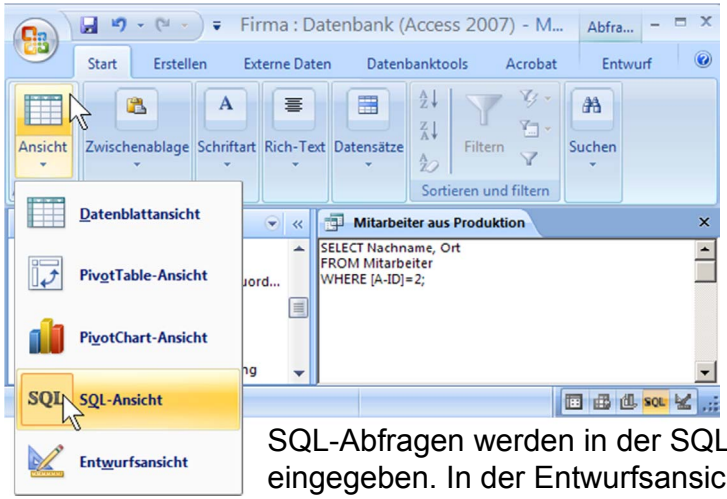
- Structured Query Language (SQL): in relationalen Datenbanksystemen verwendete formale Sprache
  - ⇨ zur Datendefinition (Definition von Tabellen)
  - ⇨ und zur Datenmanipulation (Definition von Abfragen)
- Hier wird nur auf die Datenmanipulation mit SQL (SQL als Abfragesprache) eingegangen, da man für kompliziertere Abfragen das Verständnis von SQL braucht.
- Für die Datendefinition stehen meist komfortablere Tools als die Verwendung der Sprache SQL zur Verfügung.
- Es gibt verschiedene Standards:
  - ⇨ SQL86, SQL89, SQL92, SQL3
  - ⇨ und innerhalb der Standards verschiedene „Level“
- Die nachfolgenden Beispiele sind konform mit SQL92

DATENMANIPULATION MIT SQL

Unterschiedliche SQL-Anweisungen ermöglichen verschiedene Arten von Abfragen zur Datenmanipulation:

SQL-Anweisung	Zweck
SELECT	Auswahl- oder Selektionsabfragen Abrufen von Daten
UPDATE	Aktualisierungsabfragen Ändern von Daten
INSERT	Anfügeabfragen Eintragen von neuen Daten
DELETE	Löschabfragen Löschen von Daten
SELECT ... INTO	Tabellenerstellungsabfragen

SQL-ABFRAGEN IN ACCESS



SQL-Abfragen werden in der SQL-Ansicht eingegeben. In der Entwurfsansicht sieht man die QBE- Darstellung der Abfrage, in der Datenblattansicht das Abfrageergebnis.

AUSWAHLABFRAGEN IN SQL

- Auswahlabfragen in SQL werden mit SELECT-Anweisungen festgelegt.
- Es gibt verschiedene Arten von Auswahlabfragen, z.B.:
  - ⇨ Projektion: Auswahl von Tabellenspalten
  - ⇨ Selektion: Auswahl von Tabellenzeilen
  - ⇨ Joins (Verbünde): Vereinigung der Spalten zweier Tabellen unter bestimmten Bedingungen
  - ⇨ Union (Vereinigung): Vereinigung der Zeilen zweier Tabellen
  - ⇨ Diese Arten können auch kombiniert werden.
- Es gibt verschiedene Abfrageoptionen:
  - ⇨ Sortieren, Gruppieren von Datensätzen
  - ⇨ Umbenennen von Feldern

PROJEKTION

Feld-auswahl

SELECT Name, Ort FROM Mitarbeiter;

ABFRAGEOPTIONEN BEI DER PROJEKTION

HOCHSCHULE DER MEDIEN

- Alle Spalten auswählen:
  SELECT * FROM Mitarbeiter;
- Nur unterschiedliche Datensätze:
  SELECT DISTINCT Ort, Strasse FROM Mitarbeiter;
- Spalten umbenennen:
  SELECT Name AS Nachname, Ort AS Heimatort
  FROM Mitarbeiter;
- Konstante Spalten einfügen:
  SELECT Name, Ort, "USA" AS Land FROM Mitarbeiter;
- Rechnen mit Spalten (arithmetische Operatoren: **+ - * /**):
  SELECT anzahl * einzelpreis AS preis FROM Bestellung;
- Textoperationen (Verknüpfungsoperator **&**):
  SELECT name & ", " & vorname AS anzeigename FROM Mitarbeiter;

AGGREGATFUNKTIONEN

HOCHSCHULE DER MEDIEN

Aggregatfunktionen fassen die Werte in einer Spalte zu einem kombinierten („aggregierten“) Wert zusammen. Beispiel:

SELECT COUNT(*), SUM(preis), MAX(einzelpreis)
FROM Bestellung;

Ergebnis:
eine Zeile, bestehend aus Anzahl der Datensätze, Summe aller Preise und Maximum aller Einzelpreise.

MIN(feldname), MAX(feldname)	Minimum. Maximum
SUM(feldname), AVG(feldname)	Summe, Mittelwert
COUNT(*)	Anzahl der Datensätze

SORTIER- UND GRUPPIERFUNKTIONEN

HOCHSCHULE DER MEDIEN

Datensätze lassen sich mit der Option ORDER BY numerisch und alphabetisch sortieren. Absteigend sortiert wird durch den Zusatz DESC.

SELECT * FROM Mitarbeiter ORDER BY gehalt DESC, name

Die Mitarbeiter werden nach Gehalt absteigend sortiert, innerhalb einer Gehaltsgruppe alphabetisch nach dem Namen.

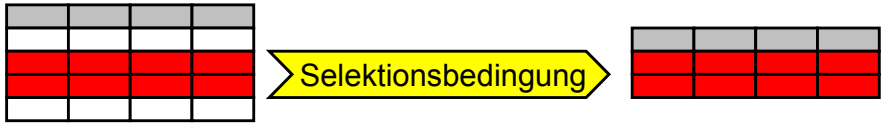
Datensätze lassen sich nach gleichen Feldwerten zu Gruppen zusammenfassen. Auf Felder, die innerhalb der Gruppe variieren, können Aggregatfunktionen angewandt werden

SELECT ort, AVG(gehalt) FROM Mitarbeiter GROUP BY ort

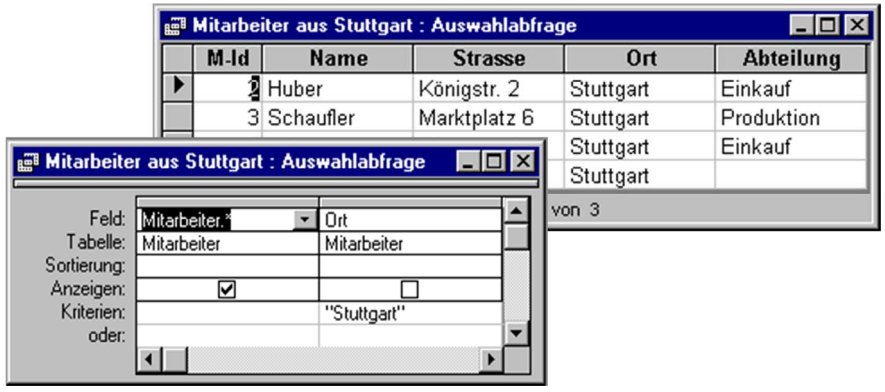
Für jeden Wohnort wird das durchschnittliche Gehalt der dort wohnenden Mitarbeiter bestimmt.

SELEKTION

HOCHSCHULE DER MEDIEN



SELECT * FROM Mitarbeiter WHERE Ort="Stuttgart";



SELEKTIONSBEDINGUNGEN IN DER „WHERE“-KLAUSEL

HOCHSCHULE DER MEDIEN

- Alle Mitarbeiter aus Stuttgart
  SELECT * FROM Mitarbeiter WHERE Ort="Stuttgart";
- Alle Mitarbeiter, die mehr als 5000 verdienen
  SELECT * FROM Mitarbeiter WHERE Gehalt > 5000;
- Alle, deren Name alphabetisch sortiert nach „Maier“ kommt:
  SELECT * FROM Mitarbeiter WHERE Name > "Maier";
- Bedingungen lassen sich auch kombinieren:
  ... WHERE Gehalt > 5000 AND NOT Ort="Stuttgart";

= gleich	**<** kleiner	**<=** kleiner oder gleich
<> ungleich	**>** größer	**>=** größer oder gleich
AND und	**OR** oder	**NOT** nicht

WEITERE ARTEN VON SELEKTIONSBEDINGUNGEN

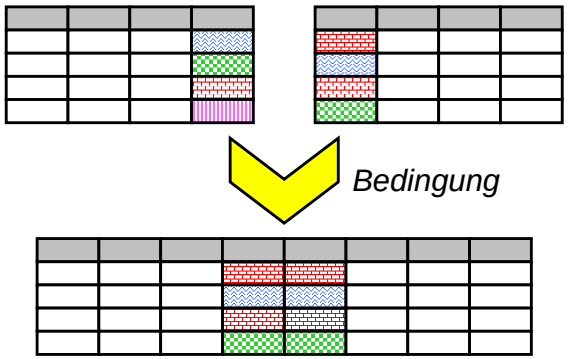
HOCHSCHULE DER MEDIEN

- Alle Mitarbeiter mit Gehalt zwischen 5000 und 10000
  SELECT * FROM Mitarbeiter
  WHERE Gehalt BETWEEN 5000 AND 10000;
- Alle Mitarbeiter aus verschiedenen Städten:
  SELECT * FROM Mitarbeiter
  WHERE Ort IN ("Stuttgart","Esslingen", "Ludwigsburg");
- Alle Mitarbeiter, für die keine Telefonnr. gespeichert ist:
  SELECT * FROM Mitarbeiter WHERE telnr IS NULL;
- Alle, deren Name mit „Ma“ beginnt:
  SELECT * FROM Mitarbeiter WHERE Name LIKE "Ma*";
  In manchen Datenbanksystemen auch ... LIKE "Ma%";

BETWEEN ... AND ... zwischen	**IN (...,...,...)** in
IS NULL „Nullwert“, nicht gesetzt	**LIKE** Textvergleich

JOIN

HOCHSCHULE DER MEDIEN



Der Join verbindet zwei einzelne Tabelle zu einer großen Tabelle unter Berücksichtigung bestimmter Bedingungen. Beim „Equivalent Join“ oder „Equi-Join“ werden Datensätze zusammengefügt, bei denen die Werte ausgewählter Felder miteinander übereinstimmen.

INNER UND OUTER JOINS AUSGANGSTABELLEN

HOCHSCHULE DER MEDIEN

Ausgangstabellen

Mitarbeiter

M-Id	Name	Strasse	Ort	Abteilung
1	Maier	Bahnhofstr. 3	Stuttgart	1
2	Huber	Königstr. 2	Stuttgart	1
3	Schaufler	Marktplatz 6	Stuttgart	1
4	Schreiber	Schloßstr. 20	Böblingen	3
5	König	Schillerstr. 29	Ludwigsburg	2
6	Kerner	Kelterstr. 51	Esslingen	3
7	Einstein	Planckstr. 6	Karlsruhe	2
8	Walz	Industriestr. 4	Mannheim	1
9	Braun	Waldstr. 4	Leonberg	2
10	Filzer	Kohlstr. 45	Filderstadt	3
11	Graf	Goethestr. 9	Sindelfingen	3
12	Kaiser	Pfarrstr. 40	Stuttgart	1
13	Hacker	Zusestr. 200	München	4

Abteilung

A-Id	Bezeichnung	Abteilungsleiter-ID
1	Einkauf	12
2	Produktion	5
3	Entwicklung	11
4	EDV	13
5	E-Commerce	

INNER UND OUTER JOINS

HOCHSCHULE DER MEDIEN

Ziel: Wir möchten die Tabellen Mitarbeiter und Abteilung durch einen Equivalent Join über die Abteilungsleiter-Id verknüpfen.

Probleme:

- Es gibt Mitarbeiter, die keine Abteilungen leiten. Sollen diese im Ergebnis mit aufgeführt werden?
- Es gibt Abteilungen, für die der Abteilungsleiter nicht festgelegt wurde. Gehören die auch zum Ergebnis?

Je nach der Antwort auf diese Frage muss man einen Inner oder einen Outer Join verwenden.

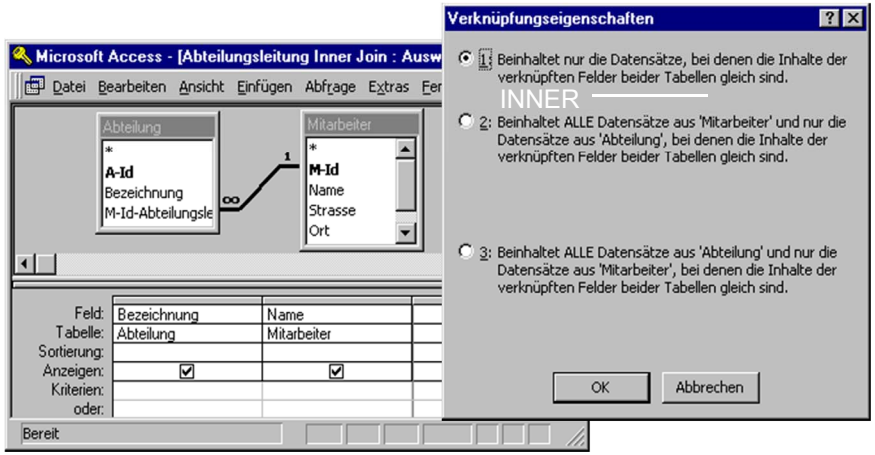
- Der „Left Outer Join“ (oder kurz: „Left Join“) schließt alle Datensätze der linken (ersten) Tabelle ein, der „Right Outer Join“ (oder kurz: „Right Join“) alle der rechten.
- Der „Inner Join“ enthält nur die Datensätze, bei denen in beiden Tabellen die betreffenden Felder existieren.

INNER UND OUTER JOINS ERGEBNISSE

HOCHSCHULE DER MEDIEN

INNER JOIN

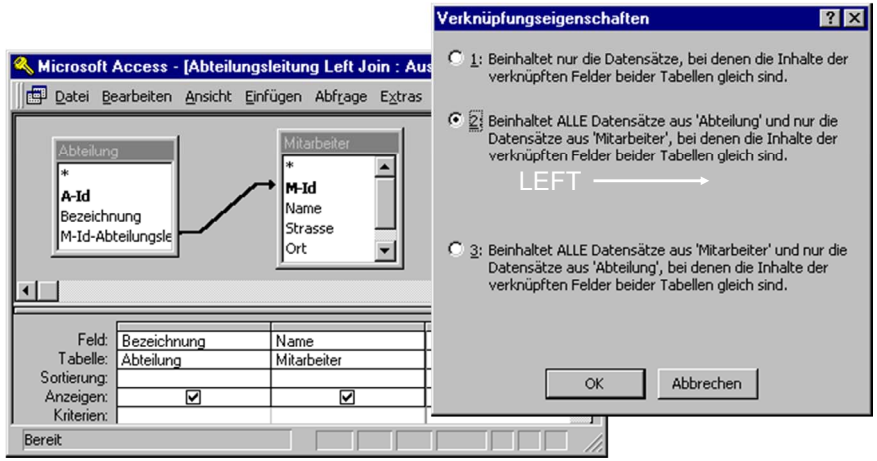
HOCHSCHULE DER MEDIEN




```
SELECT Abteilung.Bezeichnung, Mitarbeiter.Name
FROM Abteilung INNER JOIN Mitarbeiter
ON Abteilung.[M-Id-Abteilungsleiter] = Mitarbeiter.[M-Id];
```

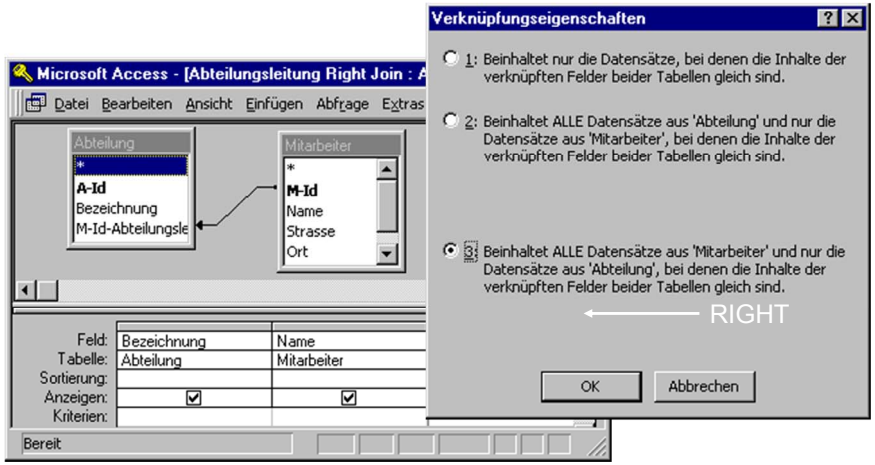

LEFT JOIN

HOCHSCHULE DER MEDIEN




```
SELECT Abteilung.Bezeichnung, Mitarbeiter.Name
FROM Abteilung LEFT JOIN Mitarbeiter
ON Abteilung.[M-Id-Abteilungsleiter] = Mitarbeiter.[M-Id];
```

RIGHT JOIN




```
SELECT Abteilung.Bezeichnung, Mitarbeiter.Name
FROM Abteilung RIGHT JOIN Mitarbeiter
ON Abteilung.[M-Id-Abteilungsleiter] = Mitarbeiter.[M-Id];
```


DEFINION VON ABFRAGEN UND FENSTER „BEZIEHUNGEN“

In der Entwurfsansicht von Abfragen werden die beteiligten Tabellen in einer Darstellung gezeigt, die dem Fenster „Beziehungen“ entspricht:

- Dabei werden bereits in den Fenstern „Beziehungen“ und „Verknüpfungseigenschaften“ eingerichtete Verknüpfungen als Voreinstellungen für Join-Abfragen verwendet.
- In der Entwurfsansicht für die jeweilige Join-Abfrage können diese Voreinstellungen überschrieben werden, falls diese nicht gewünscht sind (z.B. Übergang von einem Inner Join zu einem Left Join)
- Diese Änderungen wirken sich aber nur auf die jeweilige Join-Abfrage in der Entwurfsansicht aus und nicht auf das allgemeine Datenbankschema im Fenster „Beziehungen“

UNION

Das Ergebnis zweier SELECT-Abfragen kann vereinigt werden, wenn die Felddatentypen der jeweiligen Ergebnisspalten miteinander übereinstimmen:

SELECT … UNION SELECT … ;

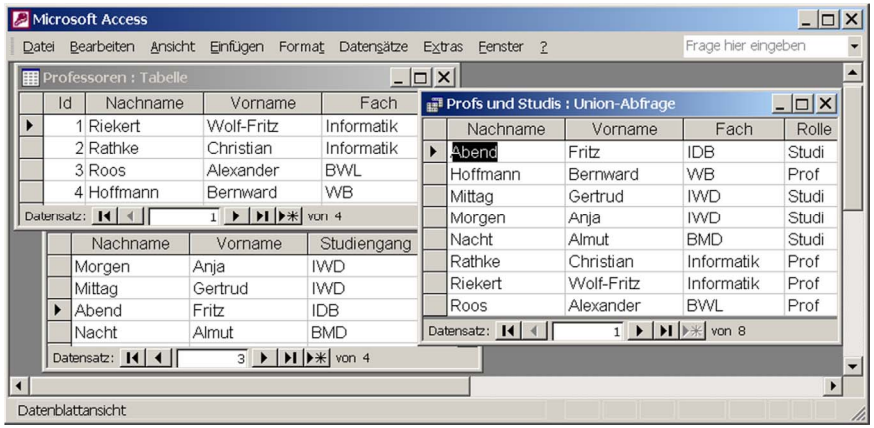
Aufgabe:

Die Tabelle Studierende besitzt die Felder Nachname, Vorname und Studiengang.

Die Tabelle Professoren besitzt die Felder Nachname, Vorname und Fach.

Gesucht wird eine Tabelle, die beide Personengruppen mit allen Angaben enthält. Es soll vermerkt werden, ob die betreffende Person Professor(in) oder Studierende(r) ist.

UNION: LÖSUNG DER AUFGABE




```
SELECT Nachname, Vorname, Fach, "Prof" AS Rolle FROM Professoren
UNION
SELECT Nachname, Vorname, Studiengang, "Studi" FROM Studierende;
```

UNTERABFRAGE MIT IN

Nach einer IN-Klausel kann eine ganze Abfrage (eine Unterabfrage) kommen, die eine einspaltige Tabelle ergibt.

```
SELECT  name FROM Mitarbeiter
        WHERE M-Id IN
                SELECT  M-Id-Abteilungsleiter
                        FROM Abteilung
```

Fragen:

1. Was bedeutet diese Abfrage?
2. Mit einem Join erzielt man dasselbe Ergebnis. Mit welchem?
3. Löse auf ähnliche Weise:
   Welche Mitarbeiter arbeiten in einem Projekt mit?

AUSWAHLABFRAGEN – BEISPIELE: AUSGANGSTABELLE

Microsoft Access - [Gebrauchtwagen : Tabelle]

Datei Bearbeiten Ansicht Einfügen Format Datensätze Extras Fenster ?

Nr	Marke	Typ	Baujahr	km	Haendler	Preis
1	Volkswagen	Passat	2000	40000	Hahn	14.000,00 €
2	Opel	Astra	1999	70000	Schott	10.000,00 €
3	Mercedes	C190	2001	20000	Mercedes Stuttgar	30.000,00 €
4	Volkswagen	Sharan	1998	80000	Schott	15.000,00 €
5	Opel	Corsa	1997	80000	Schott	8.000,00 €
6	Ford	Explorer	2002	5000	Schwabengarage	40.000,00 €
7	Mercedes	A170	2000	15000	Mercedes Stuttgar	16.000,00 €
8	Opel	Sintra	2000	30000	Hahn	19.000,00 €
9	Mercedes	SLK	1999	100000	Schott	45.000,00 €
10	Volkswagen	Golf	2001	25000	Hahn	25.000,00 €
11	Volkswagen	Passat	2001	30000	Hahn	28.000,00 €
(Wert)			0	0		0,00 €

Datensatz: 12 von 12

Automarke, z.B. Volkswagen, Opel, Mercedes NF

AUSWAHLABFRAGEN – BEISPIELE: FRAGEN

- Die Namen aller Händler
  - ⇨ Ohne Doppelnennung von Namen
  - ⇨ Alphabetisch sortiert
- Wie viele Autos bieten die Händler jeweils an
- Um welche Gesamtsumme bieten die Händler ihre Autos an
- Wie viel kostet das teuerste Auto jedes Händlers
- Berechnen Sie die Bruttopreise der Autos (+ 16% MwSt)
- Welches ist das teuerste Auto überhaupt
- Welche Autos kosten weniger als der Durchschnitt der Autos

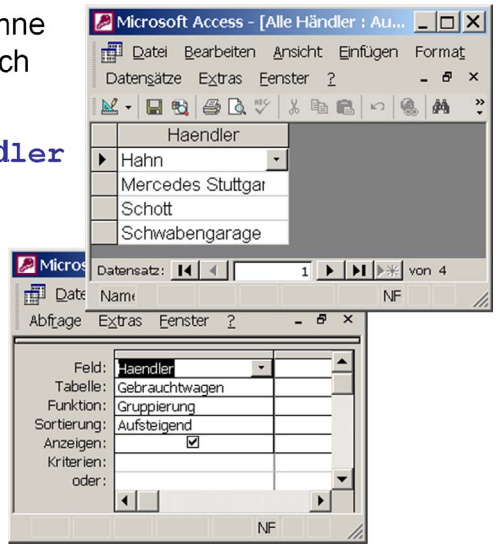
AUSWAHLABFRAGEN – BEISPIELE: LÖSUNGEN (1)

Die Namen aller Händler, ohne Doppelnennung, alphabetisch sortiert:

```
SELECT DISTINCT Haendler
FROM Gebrauchtwagen
ORDER BY Haendler;
```

oder

```
SELECT Haendler
FROM Gebrauchtwagen
GROUP BY Haendler
ORDER BY Haendler;
```



AUSWAHLABFRAGEN – BEISPIELE: LÖSUNGEN (2)

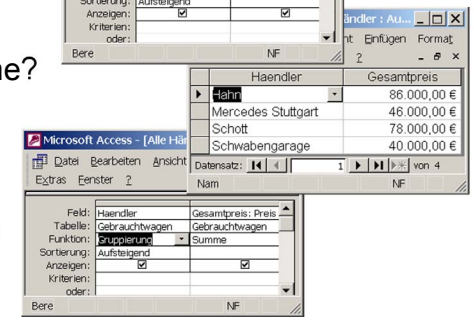
HOCHSCHULE DER MEDIEN

Wie viele Autos bieten die Händler jeweils an?

```
SELECT Haendler,
COUNT(*) AS Anzahl
FROM Gebrauchtwagen
GROUP BY Haendler
ORDER BY Haendler;
```

Um welche Gesamtsumme?

```
SELECT Haendler,
SUM(Preis)
  AS GESAMTPREIS
FROM Gebrauchtwagen
GROUP BY Haendler
ORDER BY Haendler;
```

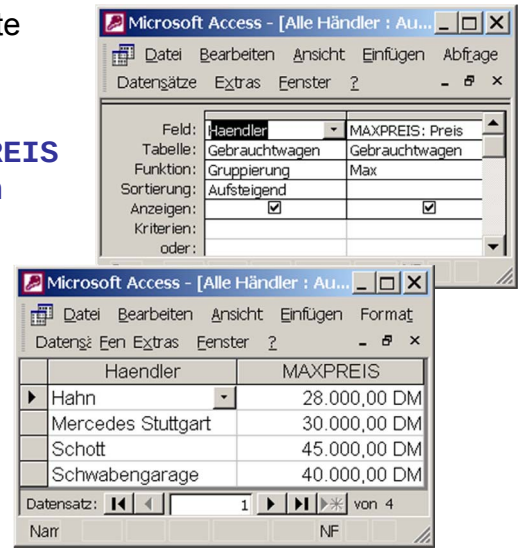


AUSWAHLABFRAGEN – BEISPIELE: LÖSUNGEN (3)

HOCHSCHULE DER MEDIEN

Wieviel kostet das teuerste Auto jedes Händlers?

```
SELECT Haendler,
MAX(Preis) AS MAXPREIS
FROM Gebrauchtwagen
GROUP BY Haendler
ORDER BY Haendler;
```



AUSWAHLABFRAGEN – BEISPIELE: LÖSUNGEN (4)

HOCHSCHULE DER MEDIEN

Berechnen Sie die Bruttopreise der Autos (+ 16% MwSt)

```
SELECT Marke, Typ,
Preis * 1.16 AS Brutto
FROM Gebrauchtwagen;
```

AUSWAHLABFRAGEN – BEISPIELE: LÖSUNGEN (5)

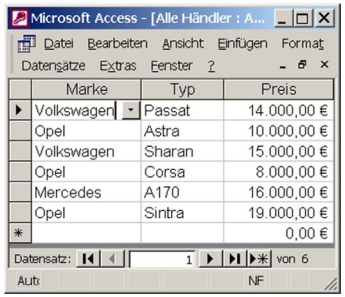
HOCHSCHULE DER MEDIEN

Was ist das teuerste Auto überhaupt?

```
SELECT Marke, Typ, Preis
FROM Gebrauchtwagen
WHERE Preis =
 (SELECT MAX(Preis)
  FROM Gebrauchtwagen);
```

Welche Autos kosten weniger als der Durchschnitt der Autos

```
SELECT Marke, Typ, Preis
FROM Gebrauchtwagen
WHERE Preis <
 (SELECT AVG(Preis)
  FROM Gebrauchtwagen);
```



AKTIONSABFRAGEN

HOCHSCHULE DER MEDIEN

Aktionsabfragen verändern Datenbankinhalte (im Gegensatz zu Auswahlabfragen)

Anfügeabfragen:

```
INSERT INTO Mitarbeiter (name, ort, strasse)
        VALUES ("Study", "Stuttgart", "Wolframstrasse");
```

Aktualisierungsabfragen:

```
UPDATE Studierende SET hochschule = "HdM"
        WHERE hochschule = "HBI";
```

Löschabfragen:

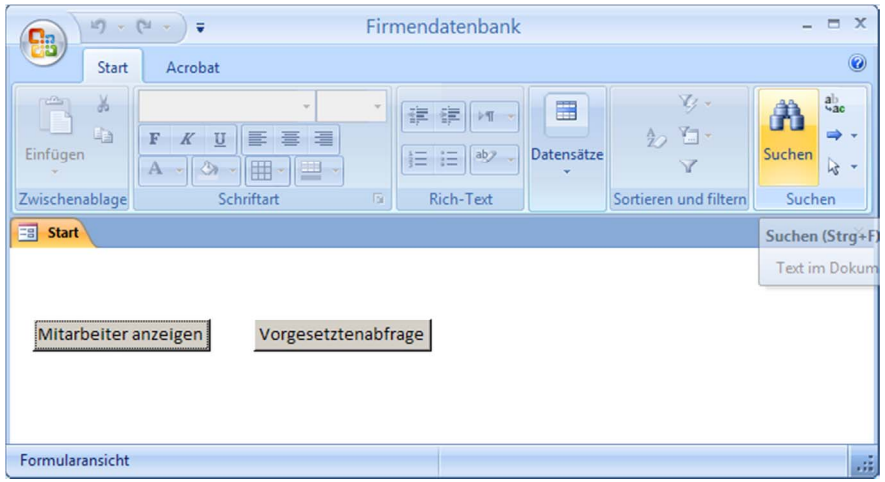
```
DELETE FROM Studierende WHERE Jahrgang = 1990;
```

Tabellenerstellungsabfrage:

```
SELECT ((irgendeine Abfrage)) INTO tabellenname;
```

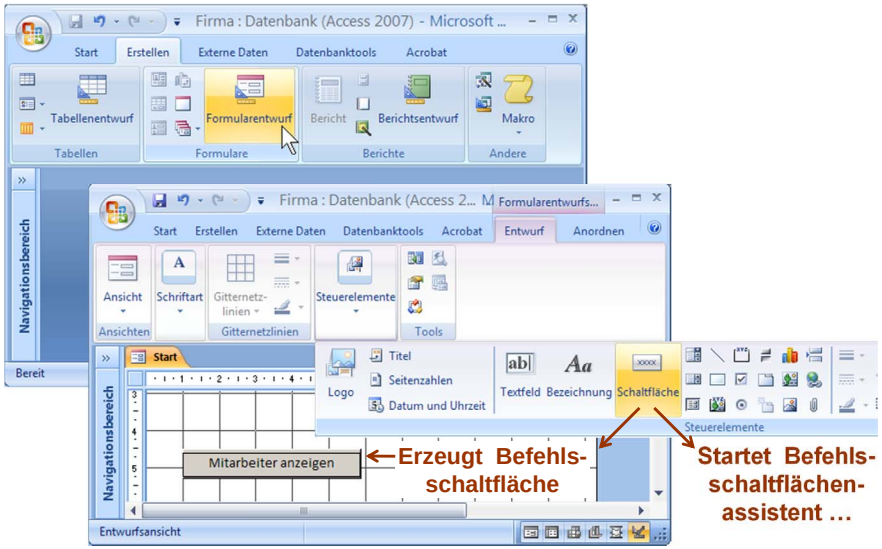

FORMULARE ZUR DIALOGSTEUERUNG

HOCHSCHULE DER MEDIEN



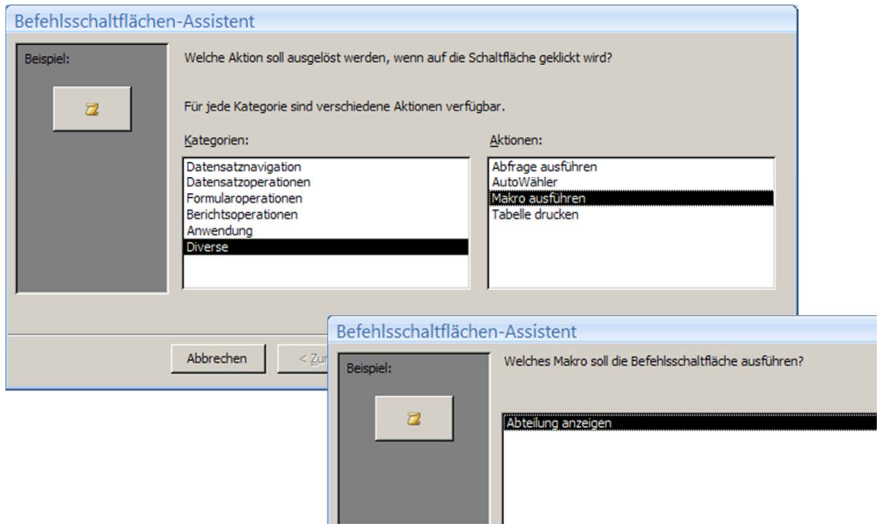
DEFINITION EINER BEFEHLSSCHALTFLÄCHE

HOCHSCHULE DER MEDIEN

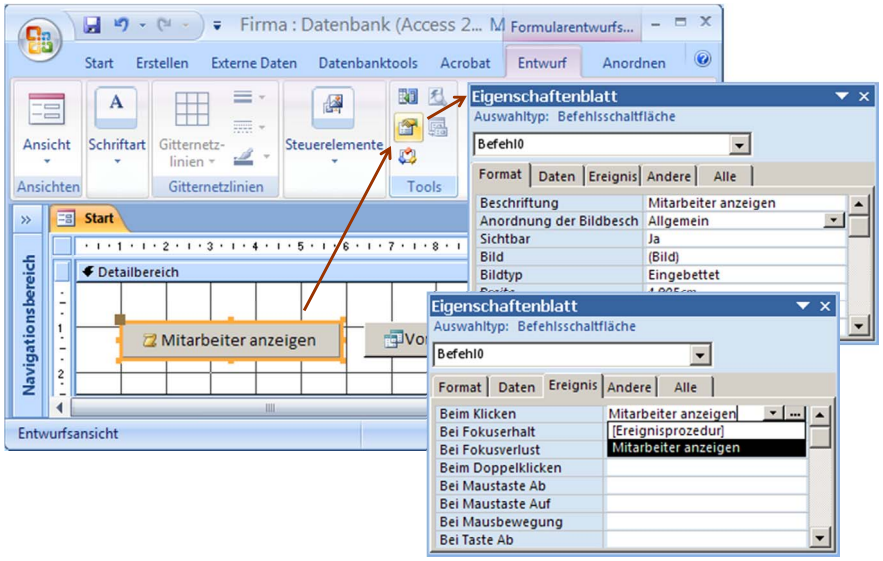


BEFEHLSSCHALTFLÄCHEN-ASSISTENT

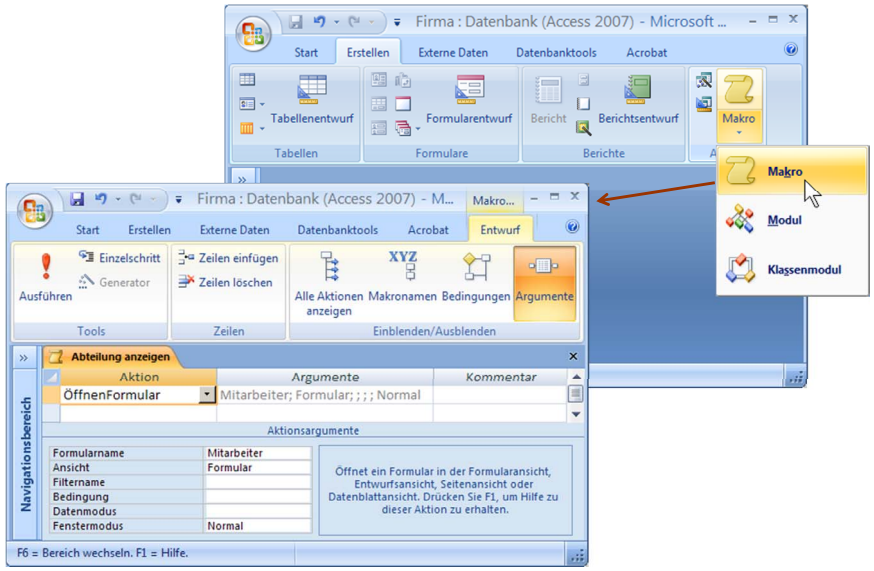
HOCHSCHULE DER MEDIEN



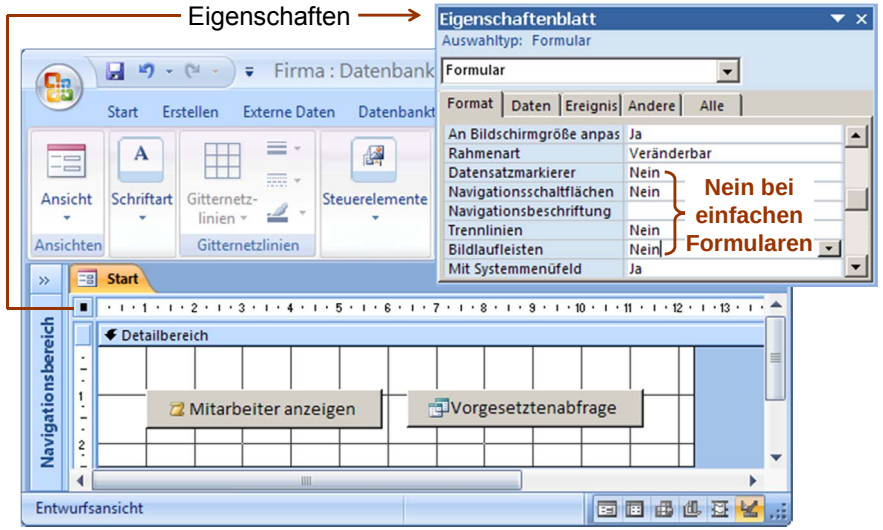
EIGENSCHAFTEN EINER BEFEHLSSCHALTFLÄCHE



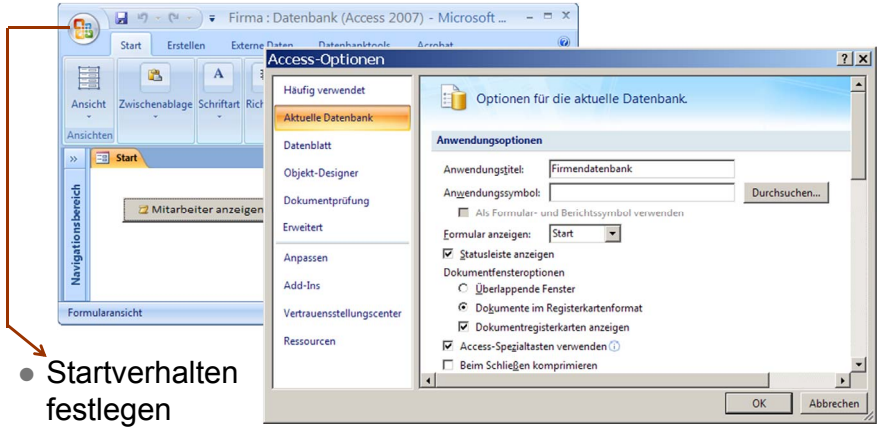
DEFINITION VON AKTIONEN MIT HILFE VON MAKROS



EIGENSCHAFTEN EINES EINFACHEN FORMULARS



VERHALTEN DER DATENBANK BEIM START



- Startverhalten festlegen
  mit Menü Office – Access-Optionen – Aktuelle Datenbank
- Besonderes Startverhalten deaktivieren durch Öffnen der Datenbank mit gedrückter Umschalt-Taste

ÜBUNGSAUFGABE: MULTIPLE-CHOICE-FRAGEN

HOCHSCHULE DER MEDIEN

a) Datenbankinhalte können geändert werden:

- durch Auswahlabfragen
- durch Aktionsabfragen

b) Was gilt für die Datenwerte in einer einzelnen Zeile einer Tabelle

- Die Datenwerte gehören alle zum selben Datensatz
- Die Datenwerte gehören alle zum selben Feld

c) Welche Vorgänge zählen zur Datenmanipulation:

- Definition eines Entitäten-Beziehungsmodells
- Definition von Tabellen
- Anzeige und Veränderung von Tabelleninhalten
- Ausführen einer Abfrage

d) Wozu dienen Normalformen

- Vermeidung von Redundanzen in Datenbanken
- Herstellung der referenziellen Integrität
- Vermeidung von Anomalien in Datenbanken

ÜBUNGSAUFGABE: AUFBAU EINER DATENBANK

HOCHSCHULE DER MEDIEN

Zeichnen Sie ein Entitäten-Beziehungsmodell, das nachfolgende Sachverhalte wiedergeben kann.

1. In einem Dokumentenverwaltungssystem gibt es Dokumente, die durch eine eindeutige Nummer und durch einen Titel gekennzeichnet sind.
2. Außerdem gibt es sogenannte Themen, die durch eine eindeutige Nummer und einen Namen gekennzeichnet sind.
3. Jedem Dokument können mehrere Themen zugeordnet werden und umgekehrt.
4. Themen stehen in einer einfachen Hierarchiebeziehung: Jedes Thema kann ein übergeordnetes Thema besitzen. Jedem Thema können mehrere Themen untergeordnet sein.

Vermerken Sie bei den Beziehungsmengen auch die jeweils zutreffenden Mächtigkeiten (**`1`**, **`m`**, **`n`**), am besten mit Hilfe von Intervalldarstellungen (z.B. **`0..1`** oder **`0..m`**), sofern sinnvoll.

Setzen Sie das Entitäten-Beziehungsmodell in Tabellen um.

TEIL 4: WEITERFÜHRENDE THEMEN - INHALTSÜBERSICHT

HOCHSCHULE DER MEDIEN

- **Mehrbenutzerbetrieb**
  - ⇨ Transaktionen
- **Objektorientierte Datenbanksysteme**
- **Erweiterungen des Entitäten-Beziehungsmodells**
  - ⇨ Generalisation
  - ⇨ Aggregation

MEHRBENUTZERBETRIEB AM BEISPIEL ÜBERWEISUNGEN

HOCHSCHULE DER MEDIEN

100 € von Konto Nr. 1 auf Nr. 999

Kontostand von Konto Nr. 1 lesen

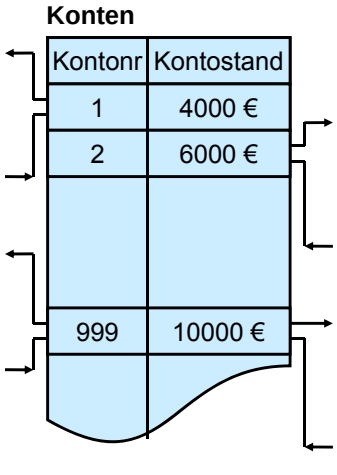
100 € subtrahieren

Konto Nr. 1 aktualisieren

Kontostand von Konto Nr. 999 lesen

100 € addieren

Konto Nr. 999 aktualisieren



200 € von Konto Nr. 2 auf Nr. 999

Kontostand von Konto Nr. 2 lesen

200 € subtrahieren

Konto Nr. 2 aktualisieren

Kontostand von Konto Nr. 999 lesen

200 € addieren

Konto Nr. 999 aktualisieren

LOST UPDATE PROBLEM

HOCHSCHULE DER MEDIEN

Beispiel: Zwei Überweisungen von den Konten Nr. 1 und 2 in Höhe von 100 € und 200 € auf das Spendenkonto Nr. 999. Dabei wird die Datenbank gleichzeitig von zwei Arbeitsplätzen (z.B. PCs der Kunden beim Homebanking) angesprochen.

Ohne besondere Vorkehrungen könnte folgendes passieren

- Bei der zweiten Überweisung wird festgestellt, dass Konto 999 den Stand 10000 € besitzt.
- Zwischenzeitlich wird das Konto Nr. 999 durch die erste Überweisung verändert und hat nun den Stand 10100 €.
- Daraufhin wird als Ergebnis der zweiten Überweisung das Konto 999 auf 10000 € + 200 € = 10200 € gesetzt.
- Das heißt, die erste Überweisung wirkt sich nicht auf Konto 999 aus ⇒ **Lost Update Problem**

KONSISTENZERHALTUNG DURCH TRANSAKTIONEN

HOCHSCHULE DER MEDIEN

- **Mehrbenutzerbetrieb** von Datenbanken bringt besondere **Konsistenzprobleme** mit sich:
  - ⇨ Der Vorgang einer Überweisung im vorigen Beispiel darf nicht durch eine andere Überweisung unterbrochen werden.
  - ⇨ Andernfalls wäre die Konsistenz der Datenbank gestört (Lost Update Problem).
- Abhilfe durch das Konzept der **Transaktion**:
  - ⇨ Die Überweisungsvorgänge sind Beispiele für sogenannte Transaktionen
  - ⇨ Transaktionen sollen ungestört von anderen Transaktionen ablaufen (wie in Einbenutzersystemen)
  - ⇨ Im Zweifelsfall müssen Transaktionen nacheinander ablaufen (Serialisierung)

EIGENSCHAFTEN VON TRANSAKTIONEN

HOCHSCHULE DER MEDIEN

- **Isolation** von Transaktionen: Gleichzeitig ablaufende Transaktionen müssen voneinander isoliert ablaufen, d.h. ihre Wirkung muss so sein, als ob sie nacheinander abliefen (Serialisierbarkeit)
- **Konsistenz** von Transaktionen: Während der Transaktion sind u.U. Konsistenzbedingungen verletzt. Am Ende jeder Transaktion ist aber die Konsistenz wieder hergestellt.
- **Atomarität** von Transaktionen: Transaktionen sollen entweder ganz oder garnicht ausgeführt werden.
- **Dauerhaftigkeit** von Transaktionen: Korrekt abgeschlossene Transaktionen haben dauerhafte Wirkung. Das muss auch gelten, wenn gleich nach Abschluß einer Transaktion ein Computerausfall stattfindet.

REALISIERUNG VON TRANSAKTIONEN

HOCHSCHULE DER MEDIEN

- In Computerprogrammen, die Datenbanken verändern, müssen Transaktionen gekennzeichnet werden.
- Dies geschieht durch zwei Operationen: BEGIN_OF_TRANSACTION und END_OF_TRANSACTION
- Grundidee (es gibt noch ausgeklügeltere Verfahren):
  - ⇨ Wenn eine Transaktion auf einen Datensatz zugreift, wird dieser für andere Transaktionen gesperrt.
  - ⇨ Wenn eine Transaktion auf einen gesperrten Datensatz zugreifen will, wird die Transaktion verzögert, bis der Datensatz wieder freigegeben wird.
  - ⇨ Am Ende einer Transaktionen werden alle Sperren aufgehoben.

OBJEKTORIENTIERTE DATENBANKSYSTEME (OODBMS)

OODBMS besitzen ein objektorientiertes Datenmodell (kein relationales). Das zugrundeliegende Konzept ist das **Objekt** (anstelle der *Tabelle* in relationalen DBMS):

- Objekte gehören einer **Klasse** an (z.B. Buch).
- Objekte haben eine eindeutige **Identität** (z.B. Inventar-Nr.).
- Objekte besitzen **Merkmale** und Merkmalswerte, diese können auch komplex sein (z.B. Menge anderer Objekte).
- **Aggregation**: Objekte können sich aus anderen Objekten zusammensetzen (z.B. ein Buch besteht aus Kapiteln).
- **Generalisierung**: Klassen können eine Hierarchie aus Ober- und Unterklassen bilden (z.B. Publikation ist eine Oberklasse von Buch und Zeitschrift).
- Objekte zeigen „Verhalten“, d.h. sie besitzen **Methoden** (z.B. „Entleihen“, „Zurückgeben“).

BEISPIEL FÜR EIN OBJEKT

Beispiel: das Buch „Object-oriented Analysis and Design“:

- **Klasse:** Buch
- **Identität:** eine Nummer
- **Merkmale** (die Merkmalswerte können Zahlen, Texte, Objekte oder komplexe Merkmalswerte, d.h. Mengen sein):
  - ⇨ Jahr = 1986 — Zahl
  - ⇨ Titel = „Object-oriented Analysis and Design“ — Text
  - ⇨ Verlag = Prentice-Hall — Objekt
  - ⇨ Autoren = (J.Martin, J.J.Odell) — Menge von Objekten
- **Methoden** (Programme, die auf Anforderung ausgeführt werden, und das Objekt zum Gegenstand haben):
  - ⇨ Entleihen
  - ⇨ Zurückgeben

VORTEILE OBJEKTORIENTIERTER DATENBANKSYSTEME

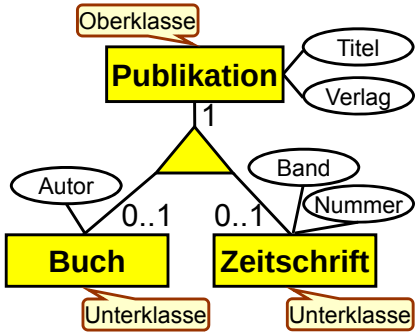
- Das Konzept eines Objekts ist anschaulicher als das einer Tabelle. Wir sind eher gewohnt, in Objekten zu denken. In einem Objekt bilden Zustand (Merkmale) und Verhalten (Methoden) eine Einheit. In relationalen Datenbanken müssen Objekte hingegen auf unterschiedlichen Datensätze verteilt werden.
- Objektmerkmale können auch komplex sein. Dies wird im Bibliotheksbereich oft gewünscht (z.B. Autorenliste).
- Objekte können auf der Benutzungsoberfläche als Bildschirmobjekte (Fenster, Symbole) dargestellt werden. Die Merkmale der Objekte lassen sich so visualisieren, die Methoden können durch Zeigeoperationen (Anklicken, Menüauswahl, Ziehen) mit der Maus ausgelöst werden.
- OODBMS lassen sich gut mit objektorientierten Programmiersprachen (z.B. Java, C++) koppeln.

OBJEKTOR. DATENBANKEN UND ENTITÄTEN-BEZIEHUNGSMODELLE

Datenbanken, die mit Hilfe eines Entitäten-Beziehungsmodells aufgebaut wurden, haben bereits viele Eigenschaften objektorientierter Datenbanken:

- Objekte entsprechen Entitäten.
- Objektklassen entsprechen Entitätsmengen.
- Identität eines Objekts entspricht Identifikationsschlüssel.
- Merkmale entsprechen den Merkmalen einer Entitätsmenge; komplexe Merkmalswerte können durch Beziehungsmengen dargestellt werden.
- Generalisierung und Aggregation können durch Erweiterung des Entitäten-Beziehungsmodells realisiert werden.
- Methoden sind jedoch im Entitäten-Beziehungsmodell nicht vorgesehen.

REGEL 6: ABBILDUNG VON KLASSENHIERARCHIEN



Publikation

Nr	Titel	Verlag

Buch

Nr	Autor

Zeitschrift

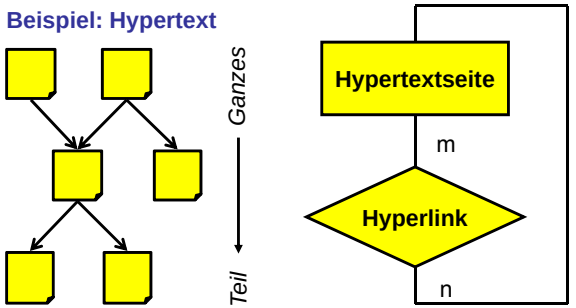
Nr	Band	Nummer

Jede Entitätsmenge einer Generalisierungshierarchie verlangt eine eigenständige Tabelle, wobei der Primärschlüssel der übergeordneten Tabelle auch Primärschlüssel der untergeordneten Tabellen wird.

In diesem Beispiel ist ein Buch durch je einen Datensatz mit demselben Primärschlüsselwert in den Tabellen Publikation und Buch repräsentiert. Eine Zeitschrift benötigt je einen Datensatz in Zeitschrift und Publikation.

REGEL 7: ABBILDUNG ZUSAMMENGESETZTER OBJEKTE

Beispiel: Hypertext



Hypertextseite

URL	Titel

Hyperlink

Von-URL	Nach-URL

Bei einer Aggregation (Bildung von zusammengesetzten Objekten) müssen sowohl die Entitätsmenge als auch die Beziehungsmenge je als eigene Tabelle definiert werden, falls der Beziehungstyp komplex-komplex ist. Die Tabelle der Beziehungsmenge enthält dann zweimal den Schlüssel aus der zugehörigen Entitätsmengentabelle mit entsprechenden Rollennamen.
Im Falle einer einfach-komplexen Beziehung (einfache Hierarchie) kann die Beziehungsmenge mit der Entitätsmenge in einer einzigen Tabelle kombiniert werden.

ÜBUNGSAUFGABE: THESAURUS

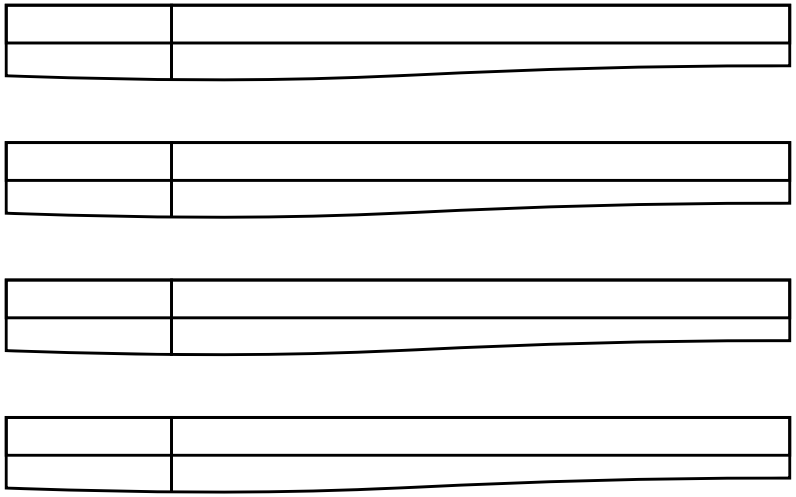
Ein Thesaurus ist eine Struktur zur Verwaltung von Deskriptoren, die zur Verschlagwortung von Dokumenten verwendet werden können, und von Non-Deskriptoren, die als Synonyme für die Deskriptoren dienen.

- Jeder Deskriptor besitzt eine eindeutige Nummer, einen Namen und eine Beschreibung.
- Deskriptoren bilden eine „Polyhierarchie“, d.h. jeder Deskriptor kann mehrere Deskriptoren als Oberbegriffe und mehrere Deskriptoren als Unterbegriffe haben.
- Non-Deskriptoren besitzen wie die Deskriptoren eine eindeutige Nummer, einen Namen und eine Beschreibung.
- Jeder Non-Deskriptor ist das Synonym für genau einen Deskriptor; Deskriptoren können mehrere Non-Deskriptoren als Synonyme besitzen.

LÖSUNG DER ÜBUNGSAUFGABE: ENTITÄTEN-BEZIEHUNGSMODELL

LÖSUNG DER ÜBUNGSAUFGABE: TABELLENSTRUKTUR

HOCHSCHULE DER MEDIEN



QUELLEN

HOCHSCHULE DER MEDIEN

Chen, P.: “The Entity-Relationship Model - Toward a Unified View of Data.” *ACM Transactions on Database Systems*, 1(1), March 1976, pp. 9-36. [Erste Publikation über die Entitäten-Beziehungsmodellierung]

Hansen, H.R. und Neumann, G.: *Wirtschaftsinformatik I*. 8. Auflage. UTB 802. Lucius & Lucius 2002. [Kapitel 11: Datenstrukturen und Datenspeicherung, gut geeignet als begleitende und weiterführende Lektüre]

Meier, A.: *Relationale Datenbanken: Leitfaden für die Praxis*. 4. überarbeitete und erweiterte Auflage. Springer, 2001. [Gut geeignet als begleitende und weiterführende Lektüre. Vorsicht: Die Assoziationstypen sind gegenüber diesem Skript und der Mehrzahl der Veröffentlichungen spiegelbildlich dargestellt!]

Neuhold, E.: *Informationssysteme I*. Vorlesungsskript. Universität Stuttgart, Institut für Informatik. 1977. [Nicht mehr erhältlich]

Stallings, W.: *Data and Computer Communication*, Macmillan, New York. 1994 (zitiert nach Müller, G.; Kohl, U. und Schoder, D.: *Unternehmenskommunikation: Telematiksysteme für vernetzte Unternehmen.* Addison-Wesley, Bonn [u.a.] 1997)

Zloof, M.M.: „Query by Example: A Data Base Language“. *IBM Systems Journal*. 16(4) pp. 324-343, 1977.