

Datenmanagement & -analyse

Datenmodellierung – Einführung und ER-Modellierung

Prof. Dr. Christoph M. Flath

Lehrstuhl für WI & BA

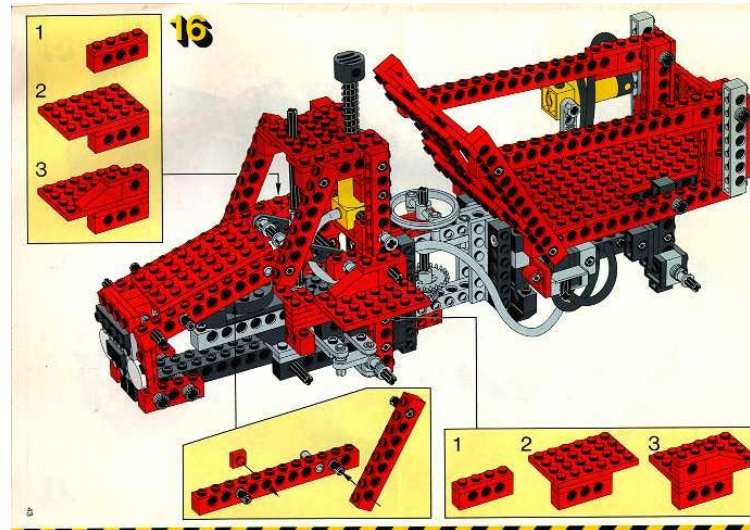
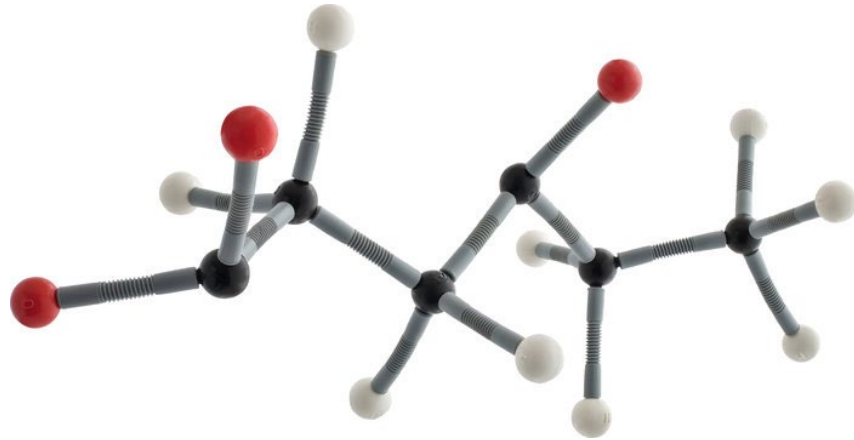
Julius-Maximilians-Universität Würzburg

Sommersemester 2021



- 1** Konzeptionelle Modellierung
- 2** Einführung Datenmodellierung
- 3** Grundlagen ER Modellierung

Modelle



▪ Abbildungsmerkmal

- Modelle sind Abbildungen, Repräsentationen natürlicher oder künstlicher Originale, die selbst wieder Modelle sein können

▪ Verkürzungsmerkmal

- Modelle erfassen nicht alle Attribute des Originals, sondern nur solche, die relevant scheinen

▪ Pragmatisches Merkmal

- Modelle sind ihren Originalen nicht eindeutig zugeordnet. Sie erfüllen ihre Ersetzungsfunktion
 - a) für bestimmte Subjekte
 - b) innerhalb bestimmter Zeitintervalle
 - c) unter Einschränkung auf bestimmte Operationen

Stachowiak (1973)



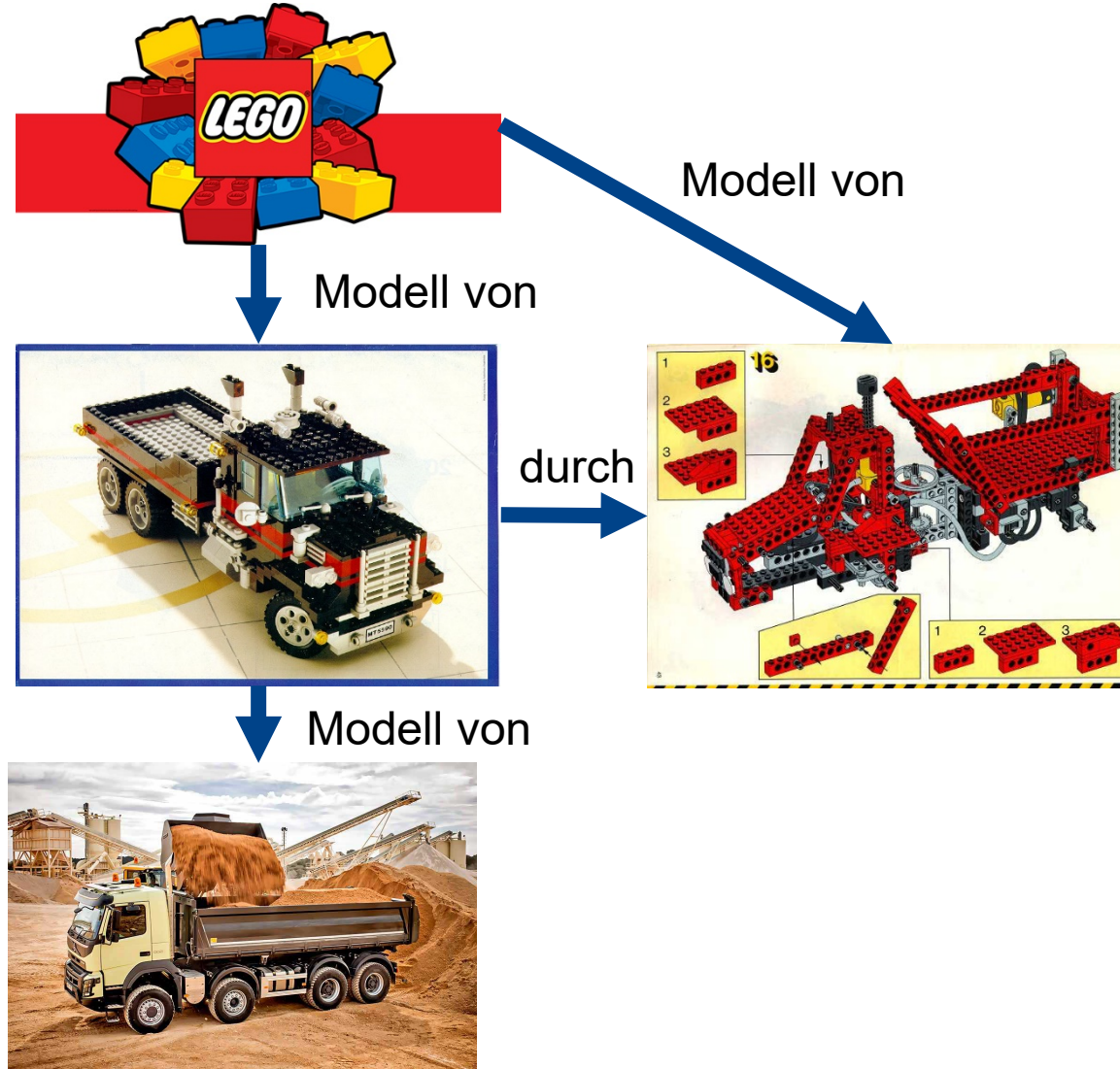
Ebene 1
Metamodell

Ebene 0
Modell

Miniwelt

Modelle

Modellierungs- prozesse

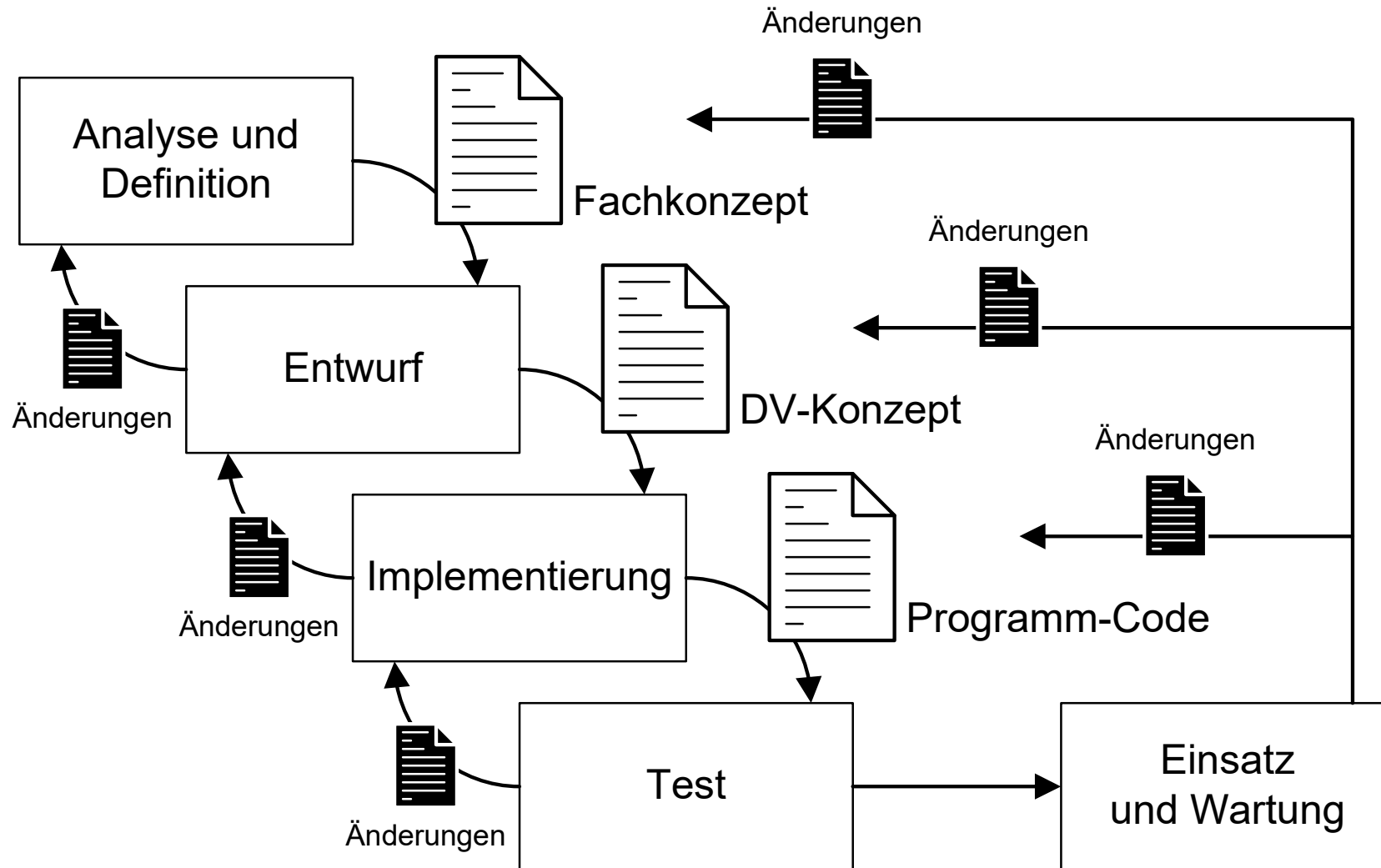


Aufgaben der konzeptionellen Modellierung

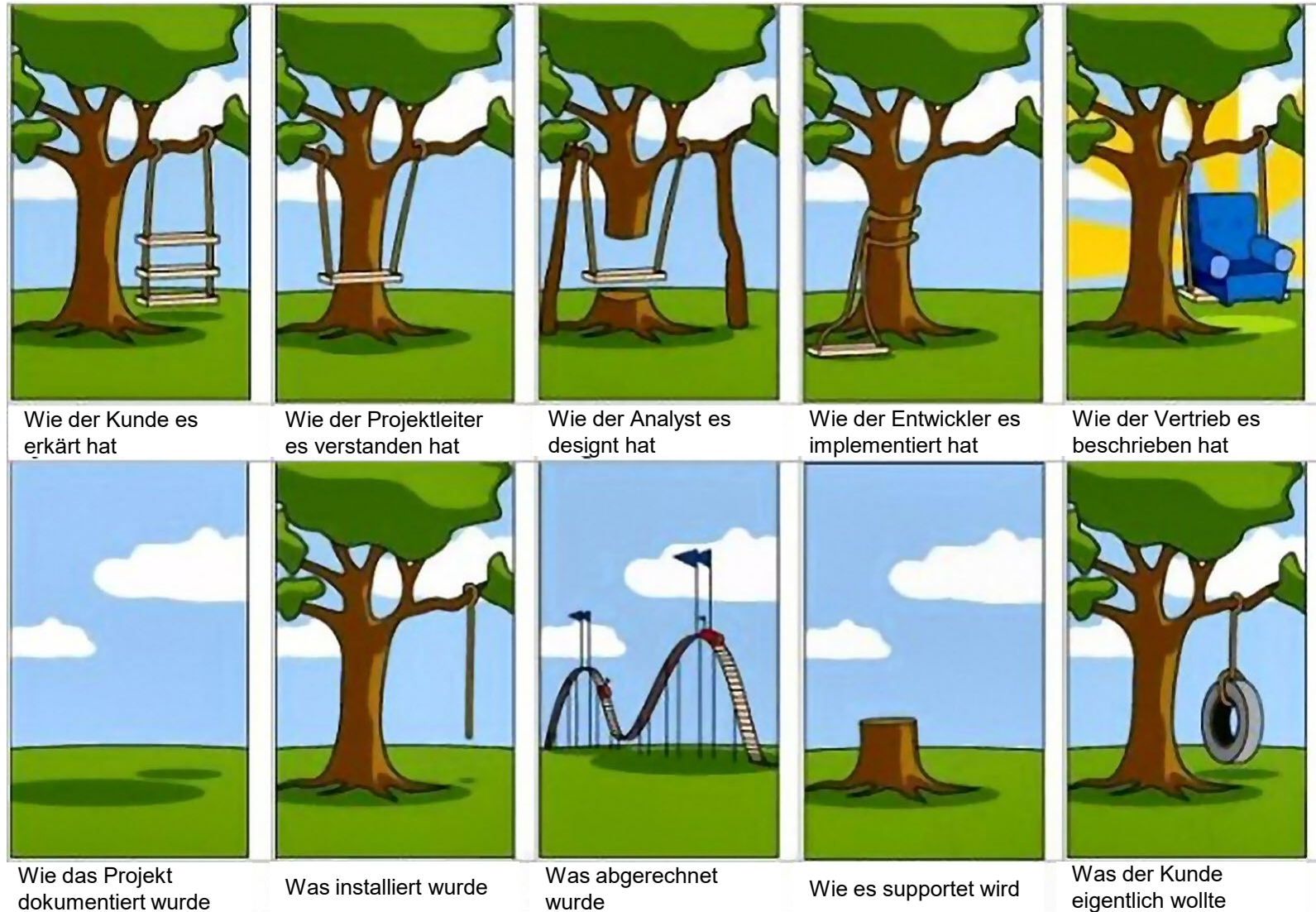
- Im Rahmen der Informationssystemgestaltung besteht ein **Komplexitätsproblem** durch
 - **Größe** des Realitätsausschnitts (Unternehmen)
 - **Diversität** der zu berücksichtigenden Elemente (Daten, Funktionen, Ressourcen, ...)
 - **Vielzahl** der Beschreibungsformen
- Architektur und Methode zur Abstraktion notwendig
 - als **Orientierungs- und Strukturierungsrahmen**
 - als **Vorgehensmodell**
 - für **vereinheitlichten** Methodeneinsatz



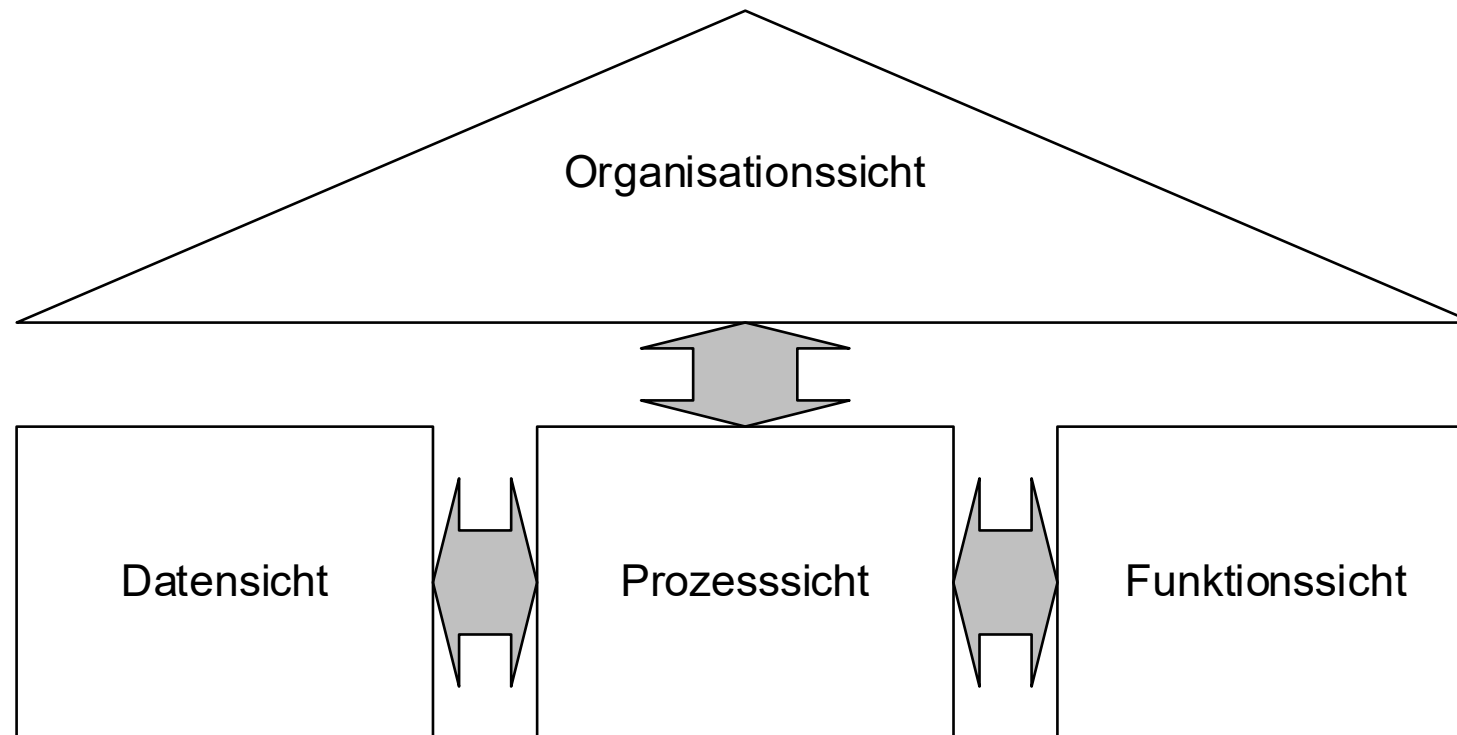
- **Management von Komplexität**
 - Reduktion der Anzahl von Elementen und Beziehungen
 - Konsolidierung/ Aggregation von Elementen und Beziehungen
 - Fokussierung / Vereinfachung
- **Strukturierung der Gegenwart (as-is) und Zukunft (to-be)**
 - Identifikation von bislang unerkannten, unbekanntem oder fehlenden Beziehungen, Ähnlichkeiten oder Unterschieden
 - Fehlererkennung
- **Dokumentation**
 - Gemeinsames Verständnis / Diskussionsunterstützung
 - Entscheidungen (Compliance)



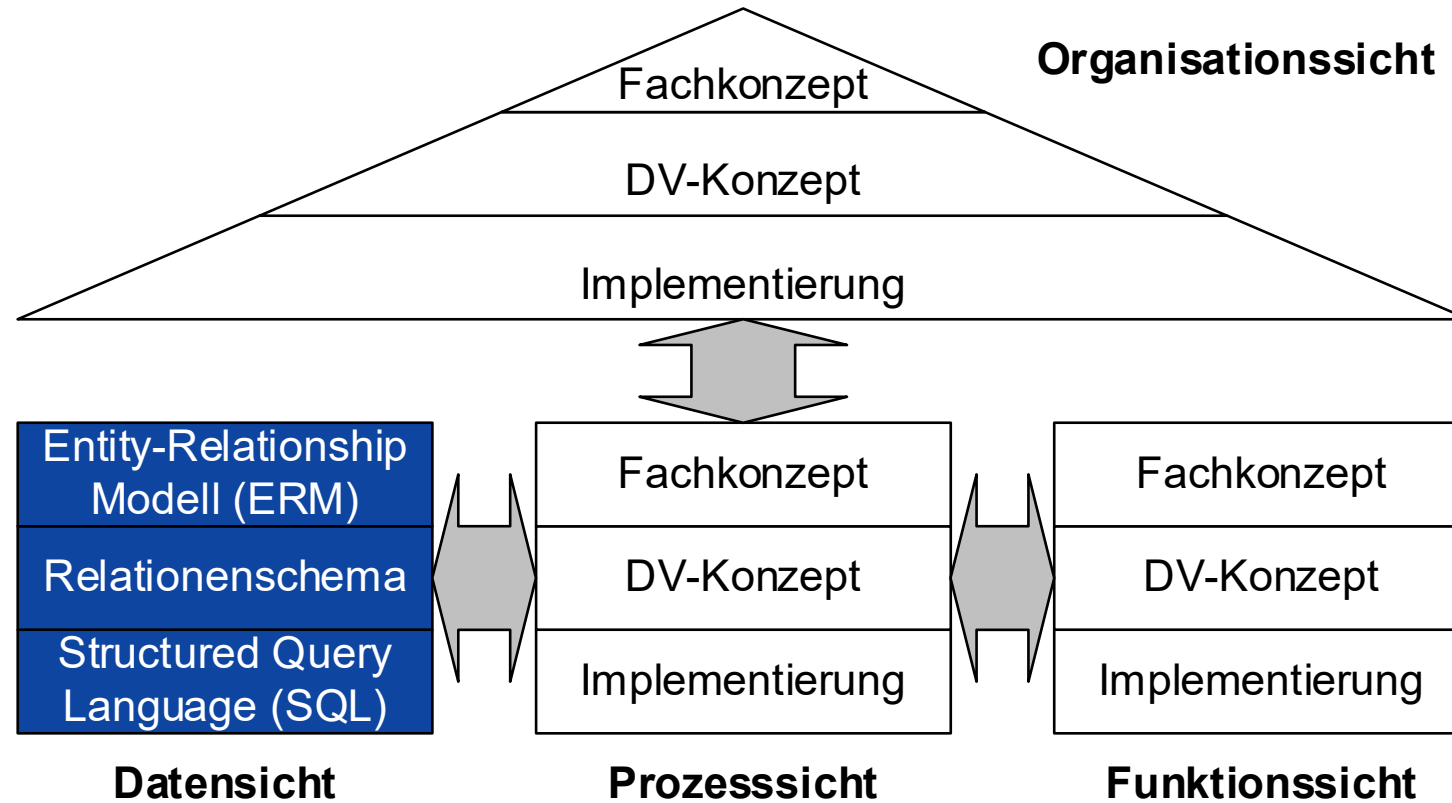
Das wäre mit Datenmodellen nicht passiert!



Architektur Integrierter Informationssysteme



Architektur Integrierter Informationssysteme



- 1 **Konzeptionelle Modellierung**
- 2 **Einführung Datenmodellierung**
- 3 **Grundlagen ER Modellierung**

- **Verfahren zur formalen Abbildung der relevanten Objekte eines Ausschnitts eines Interessenbereichs anhand bestimmter Methoden**
- zur Unterstützung
 - des Softwareentwicklungsprozesses,
 - der Anforderungsanalyse,
 - der Dokumentation von Datenstrukturen und
 - der Konzeption von Datenbankstrukturen.



- Ein Modell der **zu beschreibenden und verarbeitenden Daten** eines **Anwendungsbereichs** (...) und ihrer **Beziehungen** zueinander für einen bestimmten **Anwendungszweck**.
- Man unterscheidet entsprechend:
 - **Konzeptionelles** Datenmodell
 - Implementierungsunabhängiges Modell
 - **Logisches** Daten(bank)modell
 - Abbildung des konzeptionellen Datenbankschemas auf die Regeln des zu verwendenden Datenbankmanagementsystems
 - **Physisches** Datenbankschema
 - Anreicherung erforderliche Angaben für den technischen Betrieb

Entity-Relationship
Modell (ERM)

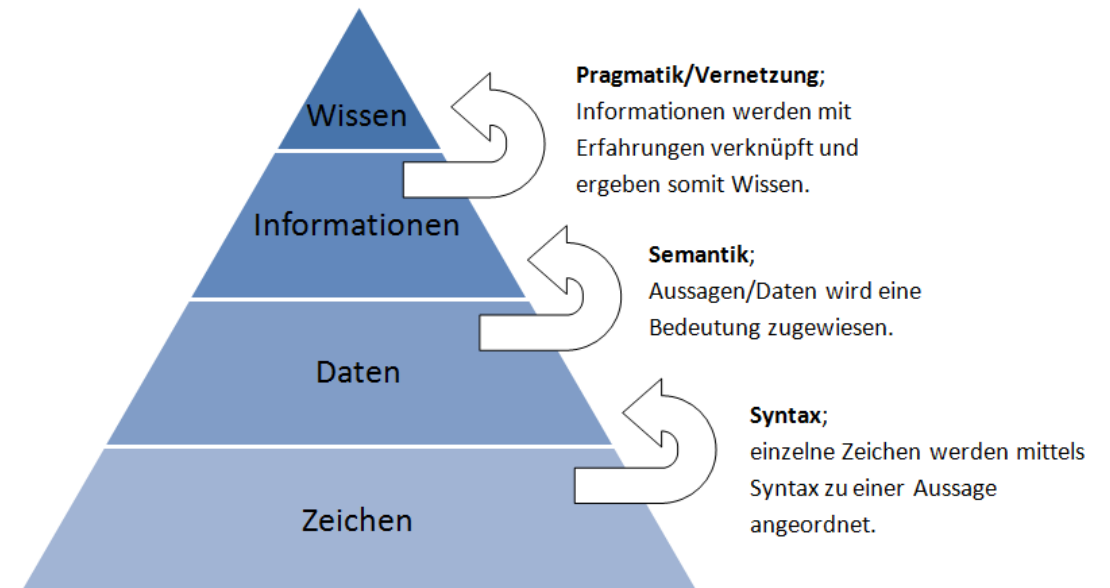
Relationenschema

Structured Query
Language (SQL)

Datensicht

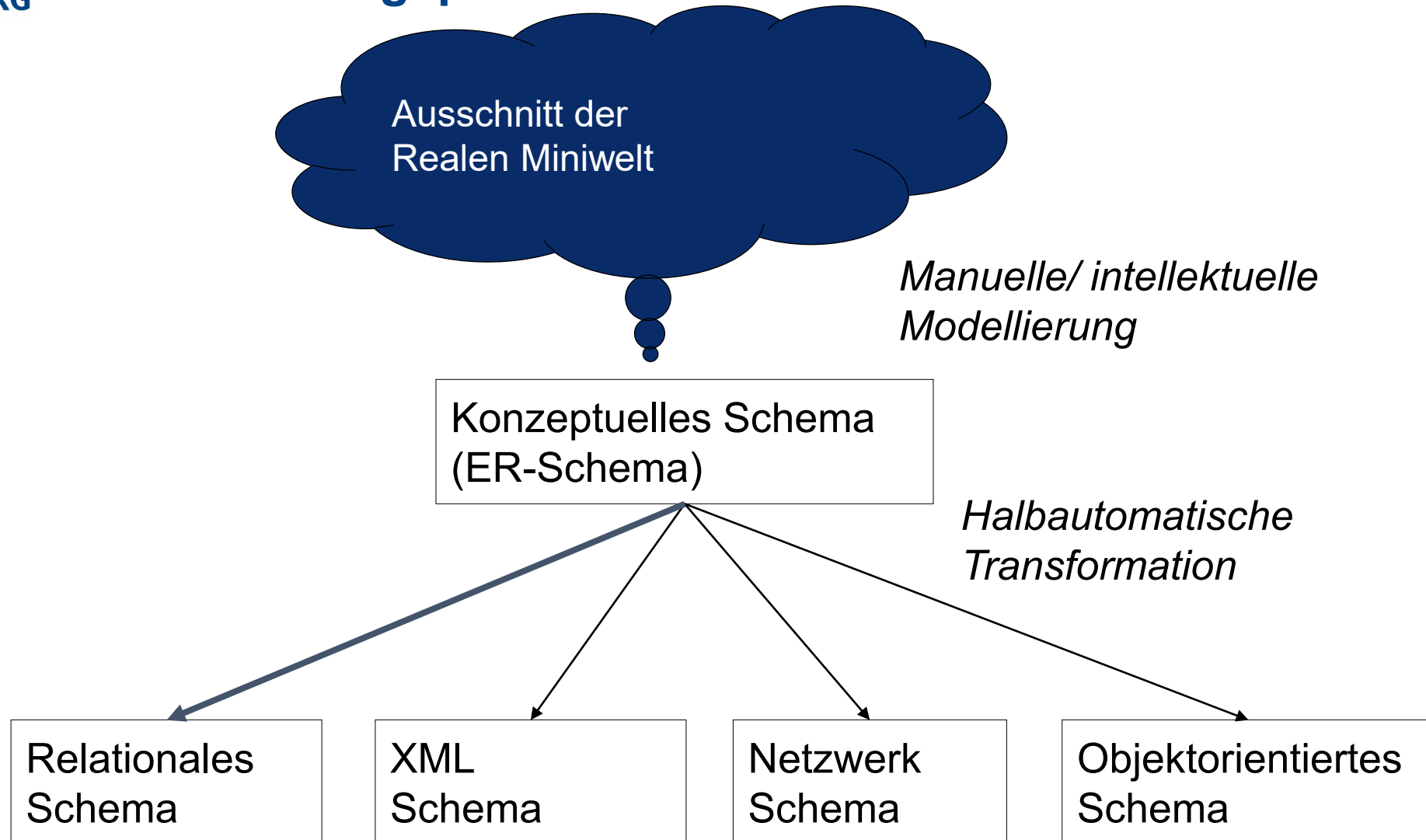
- Hierarchisches Modell
 - bspw. IBM IMS
- Netzwerkmodell
 - bspw. Siemens UDS
- **Semantisches Modell**
 - bspw. Entity-Relationship Model
- **Relationales Modell**
 - bspw. Access, DB2, MySQL, SAP MaxDB, Oracle DB, ...
- Objektorientiertes und objektrelationales Modell
 - bspw. PostgreSQL
- Semi-strukturierte Modell, nicht-relationales Modell
 - bspw. XML-Datenbanken, Amazon Dynamo DB

- Semantische Datenmodelle erlauben die Erstellung von Datenmodellen mit denen etwas mehr von der **Semantik** des Weltausschnitts erfasst werden
- Hierfür gibt es **keine Datenbanksysteme**
- Die semantische Modellierung ist grundsätzlich (datenbank-) **systemunabhängig**
- ER-Modelle können in (alle) anderen Datenbankmodelle **überführt** werden



(c) 2012 Raffael Herrmann; www.derwirtschaftsinformatiker.de

Modellierungsprozess



Kemper, Eickler (2013)

- Peter Pin-Shan Chen:**
 The Entity-Relationship Model:
 Toward a Unified View of Data.
 ACM Transactions on Database
 Systems 1 (1) 1976, S. 9-36.

The Entity-Relationship Model—Toward a Unified View of Data

PETER PIN-SHAN CHEN
Massachusetts Institute of Technology

A data model, called the entity-relationship model, is proposed. This model incorporates some of the important semantic information about the real world. A special diagrammatic technique is introduced as a tool for database design. An example of database design and description using the model and the diagrammatic technique is given. Some implications for data integrity, information retrieval, and data manipulation are discussed.

The entity-relationship model can be used as a basis for unification of different views of data: the network model, the relational model, and the entity set model. Semantic ambiguities in these models are analyzed. Possible ways to derive their views of data from the entity-relationship model are presented.

Key Words and Phrases: database design, logical view of data, semantics of data, entity-relationship model, relational model, Data Base Task Group, network model, data definition and manipulation, data integrity and consistency
CR Categories: 3.50, 3.70, 4.33, 4.34

1. INTRODUCTION

The logical view of data has been an important issue in recent years. Three data models have been proposed: the network model [2, 3, 7], the relational model [8], and the entity set model [25]. These models have their own strengths and weaknesses. The network model provides a more natural view of data entities and relationships (to a certain extent), but its capability to achieve a high degree of data independence has been challenged [8]. The relational model is based on set theory and can achieve a high degree of data independence, but it does not provide an important semantic information about the real world [12, 15, 23]. The entity set model, which is based on set theory, also achieves a high degree of data independence, but its viewing of values such as "3" or "red" may not be natural to some people [25].

This paper presents the entity-relationship model, which has more advantages of the above three models. The entity-relationship model provides a more natural view that the real world consists of entities and relationships.

Copyright © 1976, Association for Computing Machinery, Inc. General permission is granted by the Association to reproduce copies of this paper for personal or internal reference use only on a non-profit basis, but not for profit, all or part of this material is granted provided that ACM's copyright notice is given and that reference is made to the publication, to its date of issue, and to the reprinting privileges were granted by permission of the Association for Computing Machinery. A version of this paper was presented at the International Conference on Very Large Databases, Framingham, Mass., Sept. 22-24, 1975.
Author's address: Center for Information System Research, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA 02139.

ACM Transactions on Database Systems, Vol. 1, No. 1, March 1976

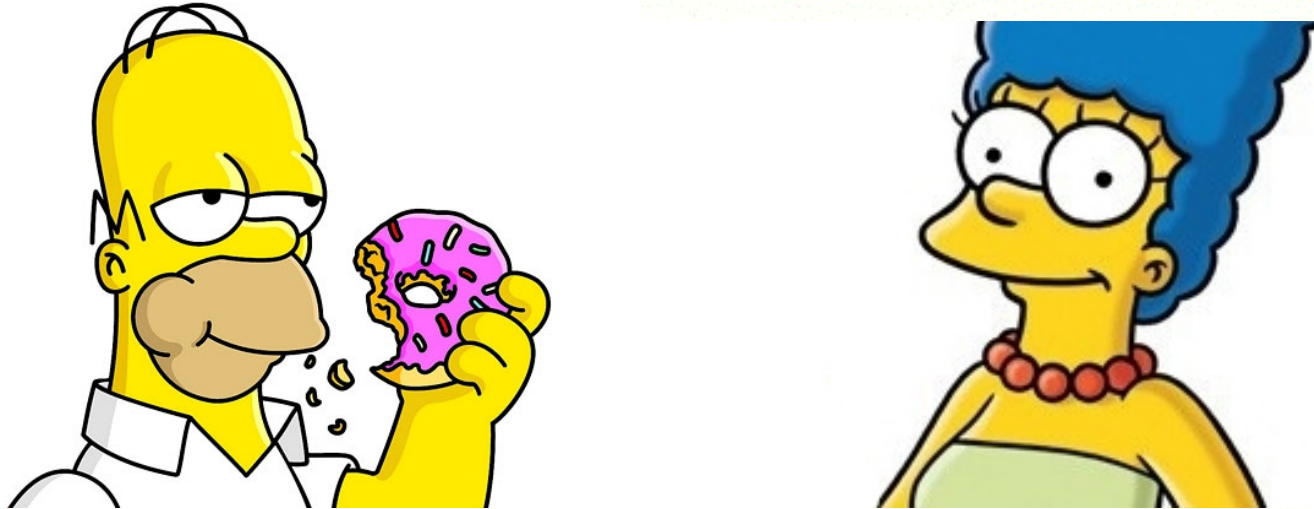


Chen (1976)

- 1** Konzeptionelle Modellierung
- 2** Einführung Datenmodellierung
- 3** Grundlagen ER Modellierung
 - 3.1** Attribute und Schlüssel
 - 3.2** Beziehungen
 - 3.3** Kardinalitäten
 - 3.4** Zeit in Datenmodellen
 - 3.5** Rekursive Beziehungen
 - 3.6** Mehrwertige Beziehungen

Entitäten und Beziehungen

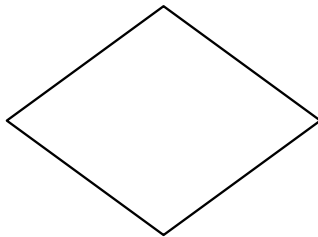
ENTITY AND RELATIONSHIP



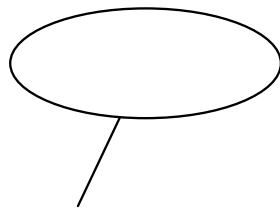
Entitäten, Beziehungen, Attribute



- **Entitätstyp (entity type)**
 - Gesamtheit aller gleichartigen Objekte
 - bspw. Personen, Vorlesungen



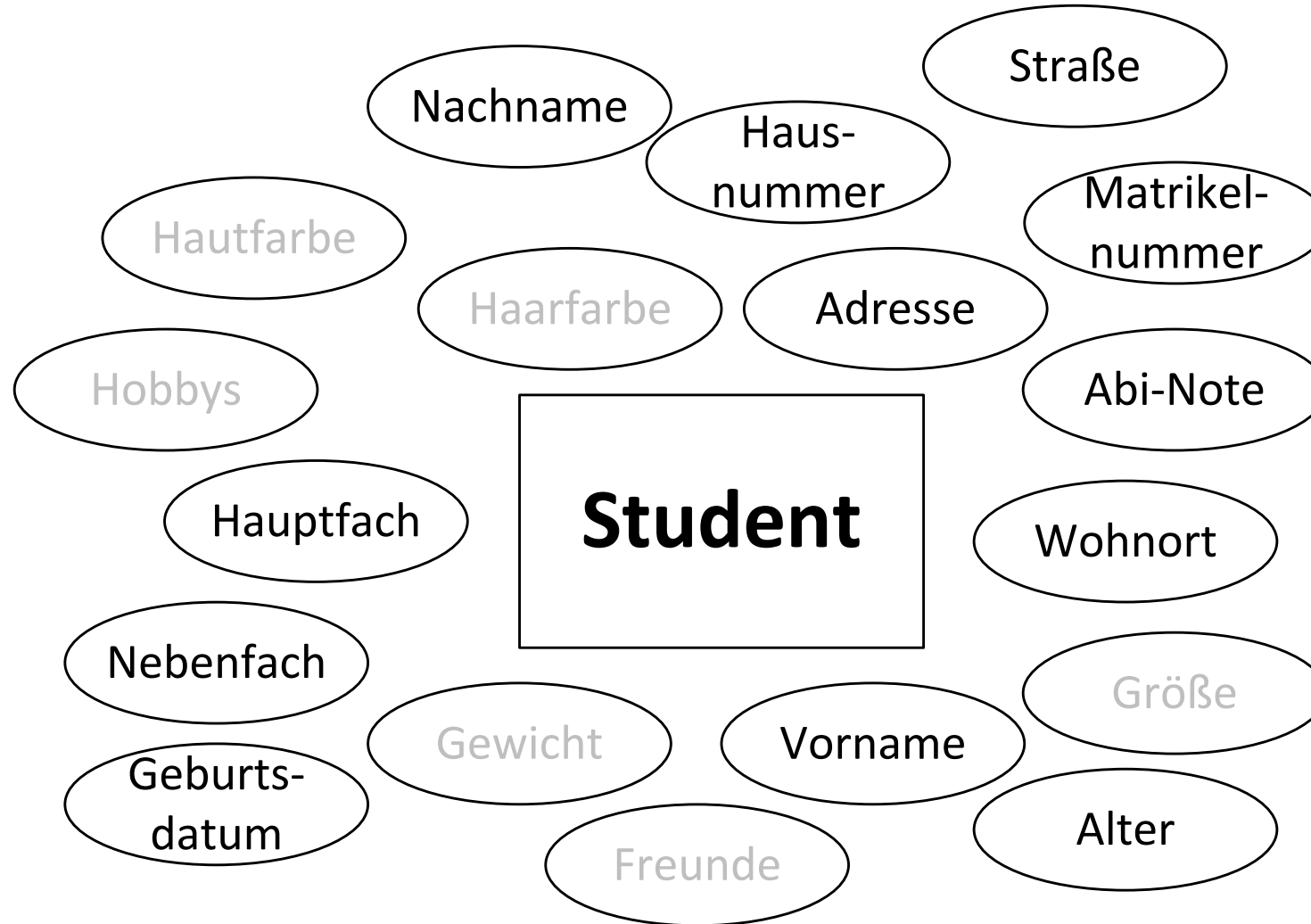
- **Beziehungstyp (relationship type)**
 - Gesamtheit aller Beziehungen von Objekten
 - bspw. “hört” oder “liest”



- **Attribut (attribute)**
 - Eigenschaft von einem Entitätstyp oder einem Beziehungstyp
 - bspw. Name, Datum, ID

Modellierung der Entität Student (1)

Modellierung der Entität Student (2)



- 1** Konzeptionelle Modellierung
- 2** Einführung Datenmodellierung
- 3** Grundlagen ER Modellierung
 - 3.1** Attribute und Schlüssel
 - 3.2** Beziehungen
 - 3.3** Kardinalitäten
 - 3.4** Zeit in Datenmodellen
 - 3.5** Rekursive Beziehungen
 - 3.6** Mehrwertige Beziehungen

Schlüsselkandidaten und Schlüsselattribut

- **Schlüsselkandidaten** sind diejenigen Attribute, die einen Entitätstyp eindeutig identifizieren **können**
- Ein **Schlüsselattribut** ist der Schlüsselkandidat, der zur **eindeutigen Identifizierung** verwendet wird
- Das Schlüsselattribut wird immer **unterstrichen** dargestellt

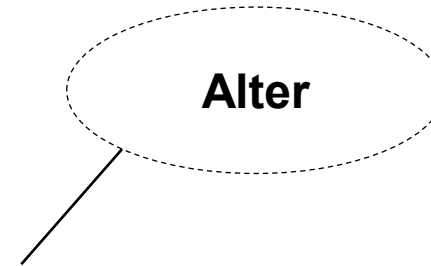


Mehrwertige Attribute

- Mehrwertige Attribute können **mehrere gleichartige** Informationen enthalten
- Mehrwertige Attribute haben eine **doppelte** Umrandung

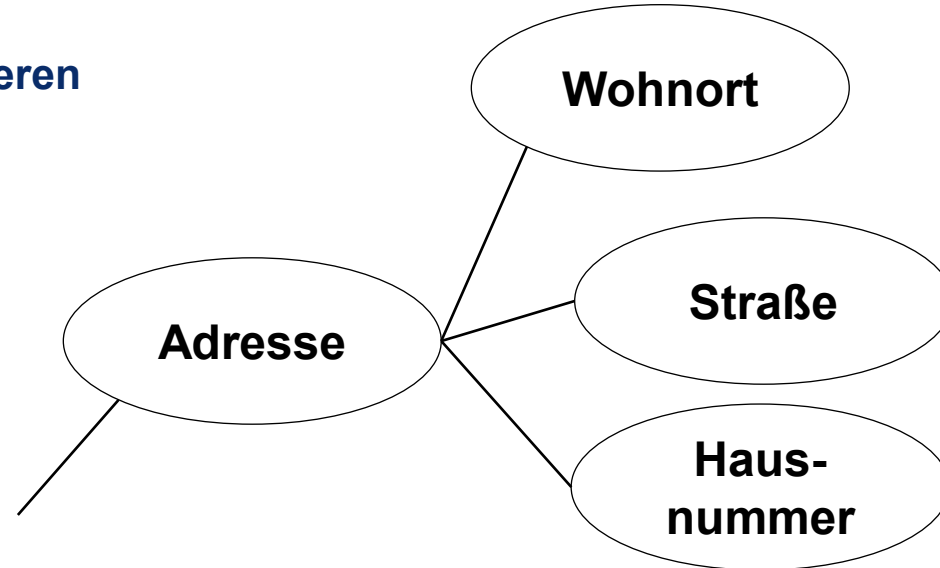


- Ein abgeleitetes Attribut **kann** aus vorhandenen Attributen **erzeugt werden**
- Ein abgeleitetes Attribut wird mit einer **gepunkteten** Umrandung dargestellt



Zusammengesetzte Attribute

- Ein zusammengesetztes Attribut besteht aus **weiteren** Attributen

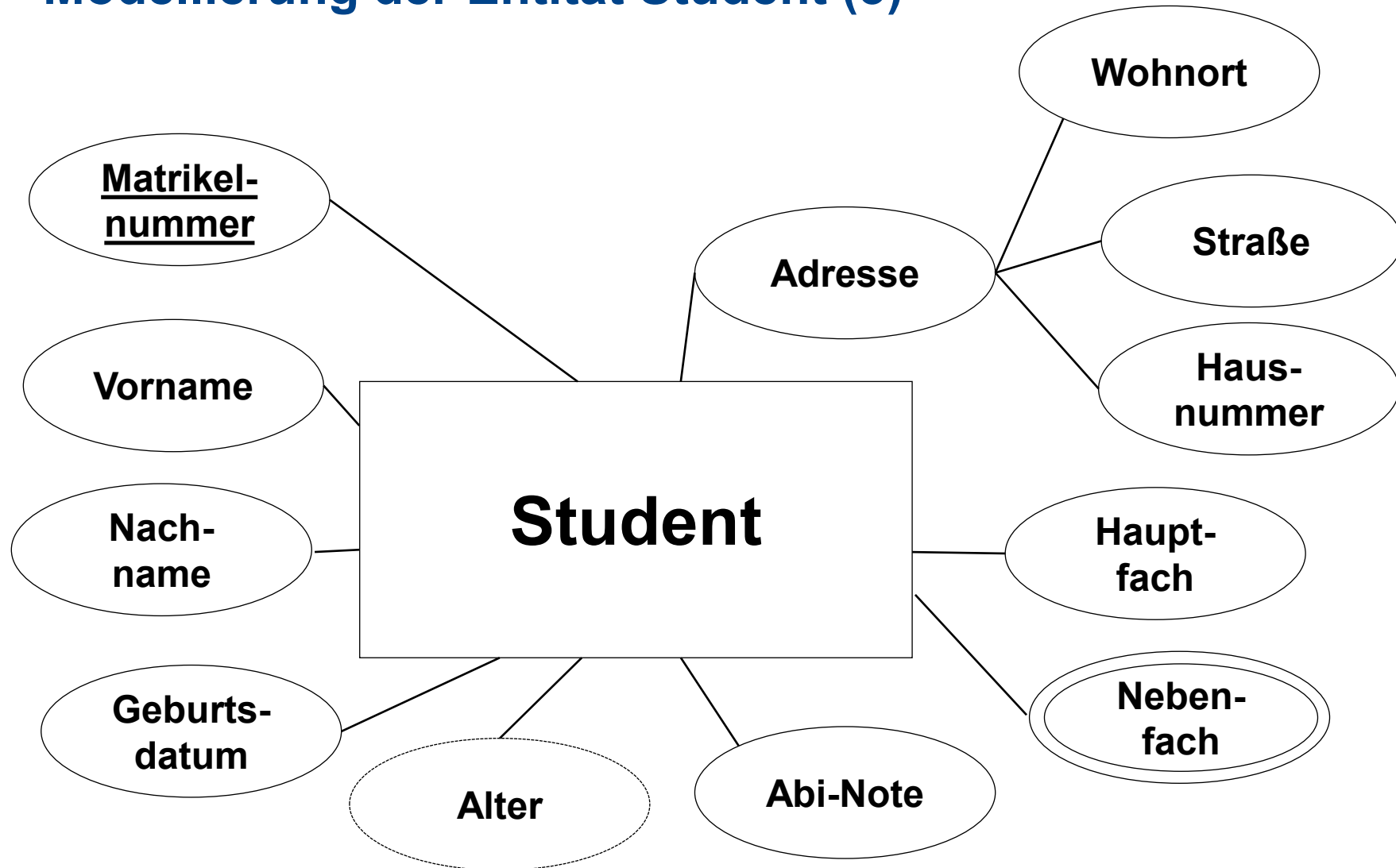


Anmerkung: Nullwerte

- Es gibt Fälle, in denen ein Attribut für eine Entität **keinen Wert** enthält
- Es gibt Fälle, in denen der **Wert** für ein Attribut für eine Entität **nicht bekannt** ist
- 0 ist kein Nullwert!
- Notwendigkeit von Nullwerten sollte immer kritisch hinterfragt werden

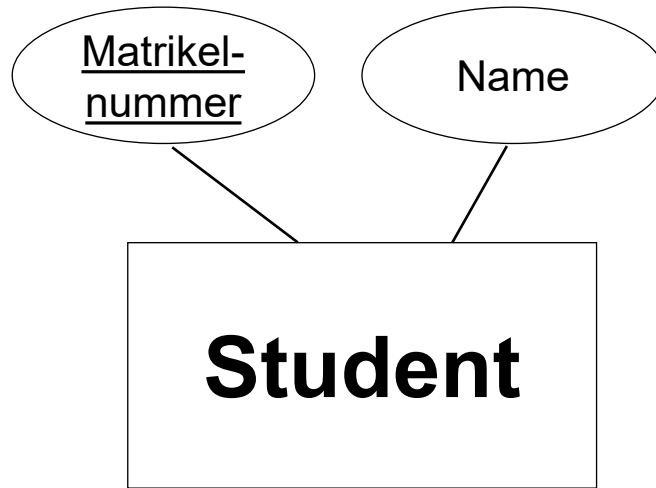


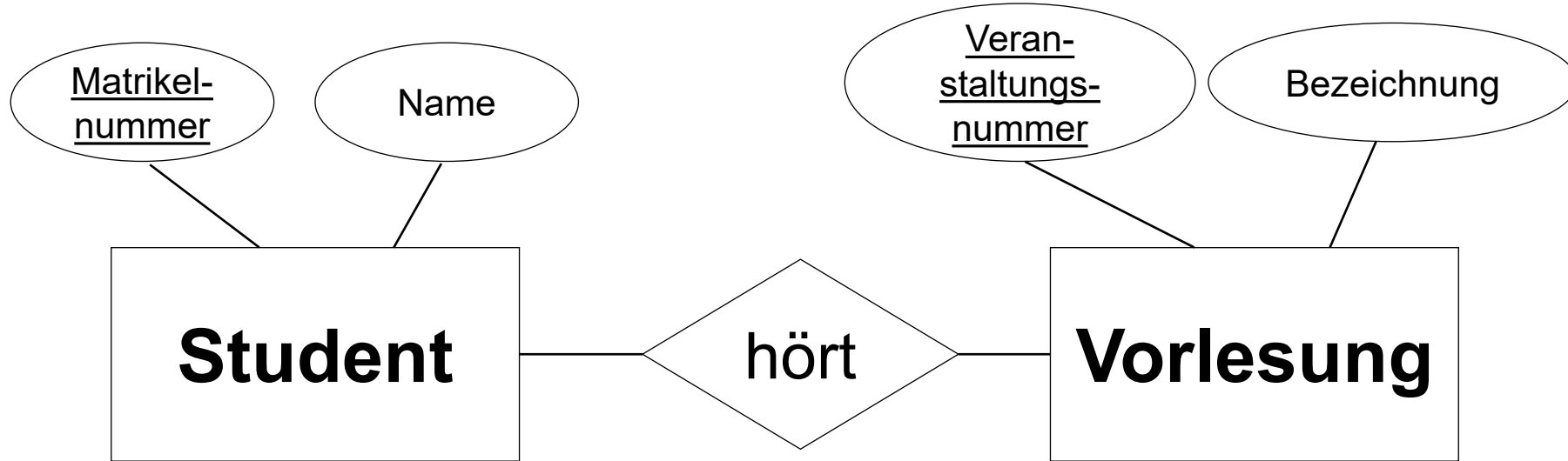
Modellierung der Entität Student (3)



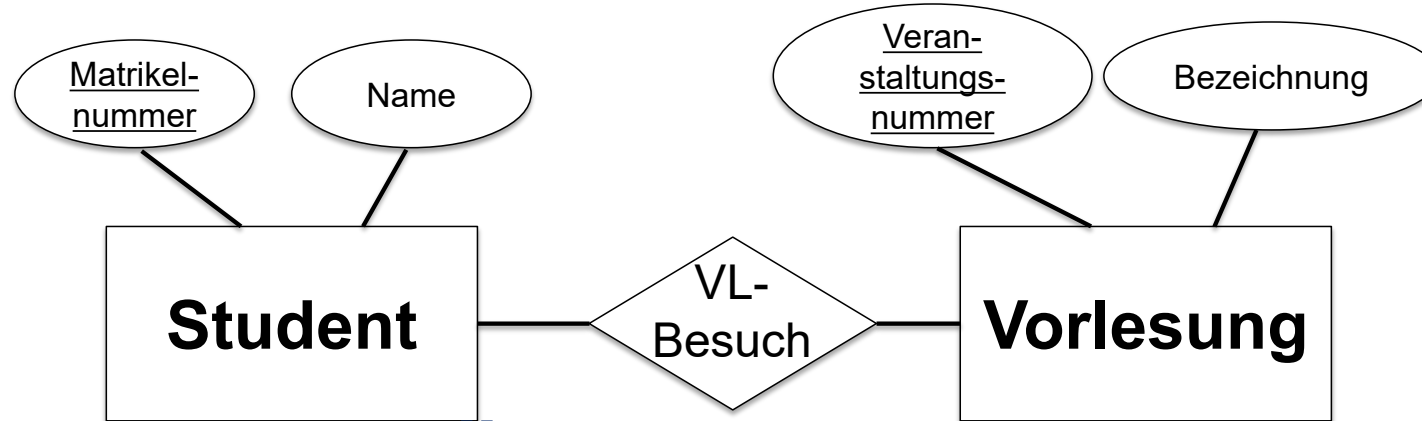
- 1** Konzeptionelle Modellierung
- 2** Einführung Datenmodellierung
- 3** Grundlagen ER Modellierung
 - 3.1** Attribute und Schlüssel
 - 3.2** Beziehungen
 - 3.3** Kardinalitäten
 - 3.4** Zeit in Datenmodellen
 - 3.5** Rekursive Beziehungen
 - 3.6** Mehrwertige Beziehungen

Beziehungen

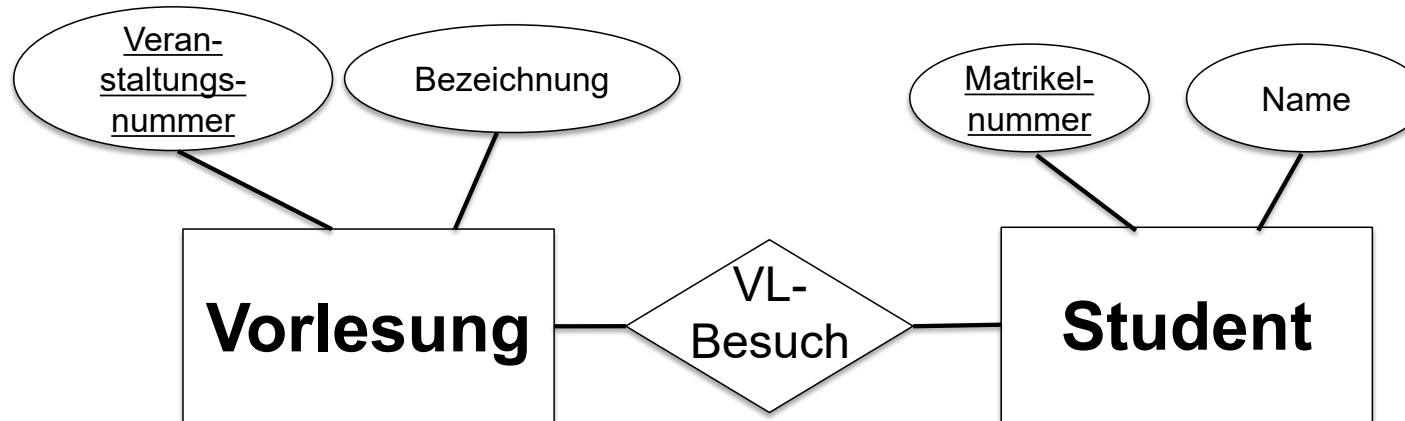




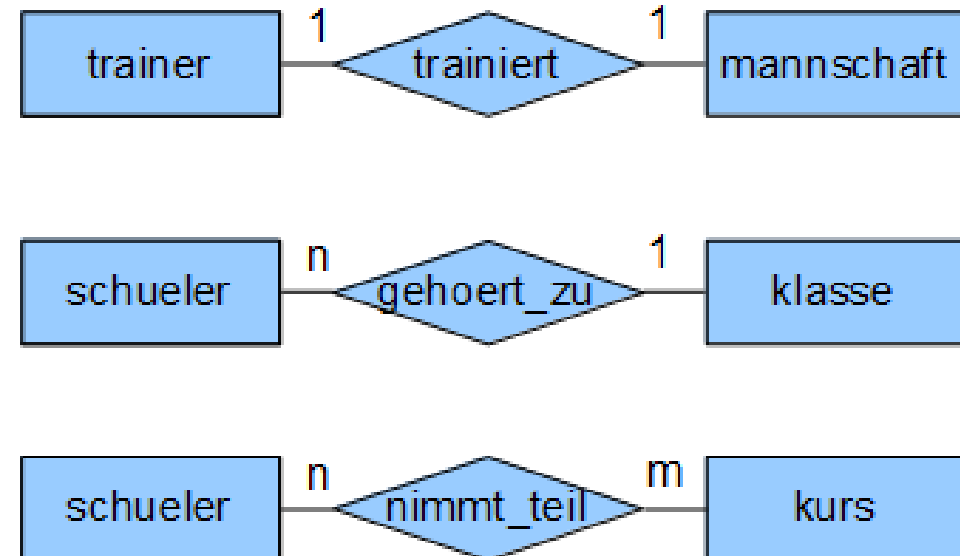
Leserichtung?

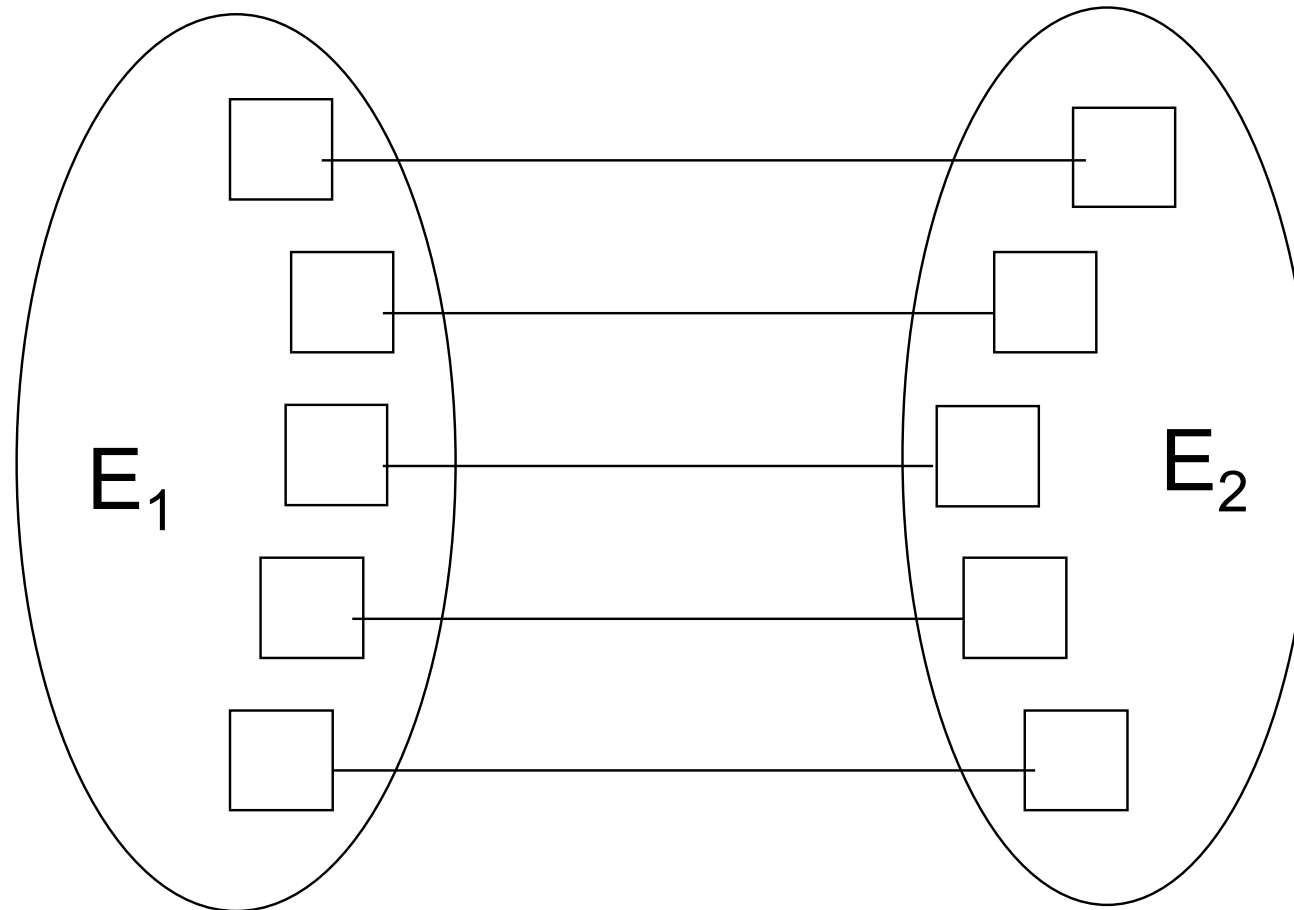


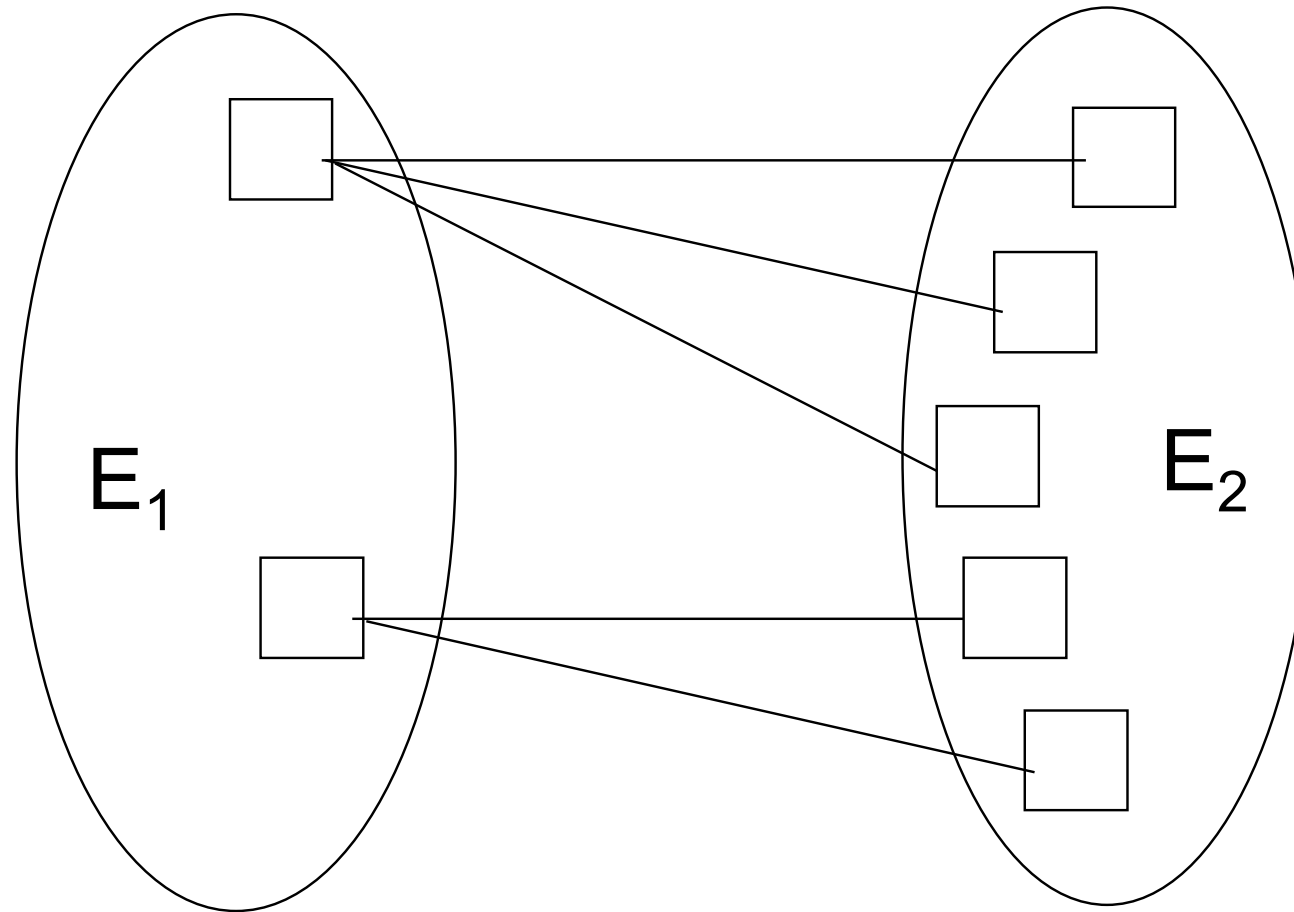
Äquivalent !

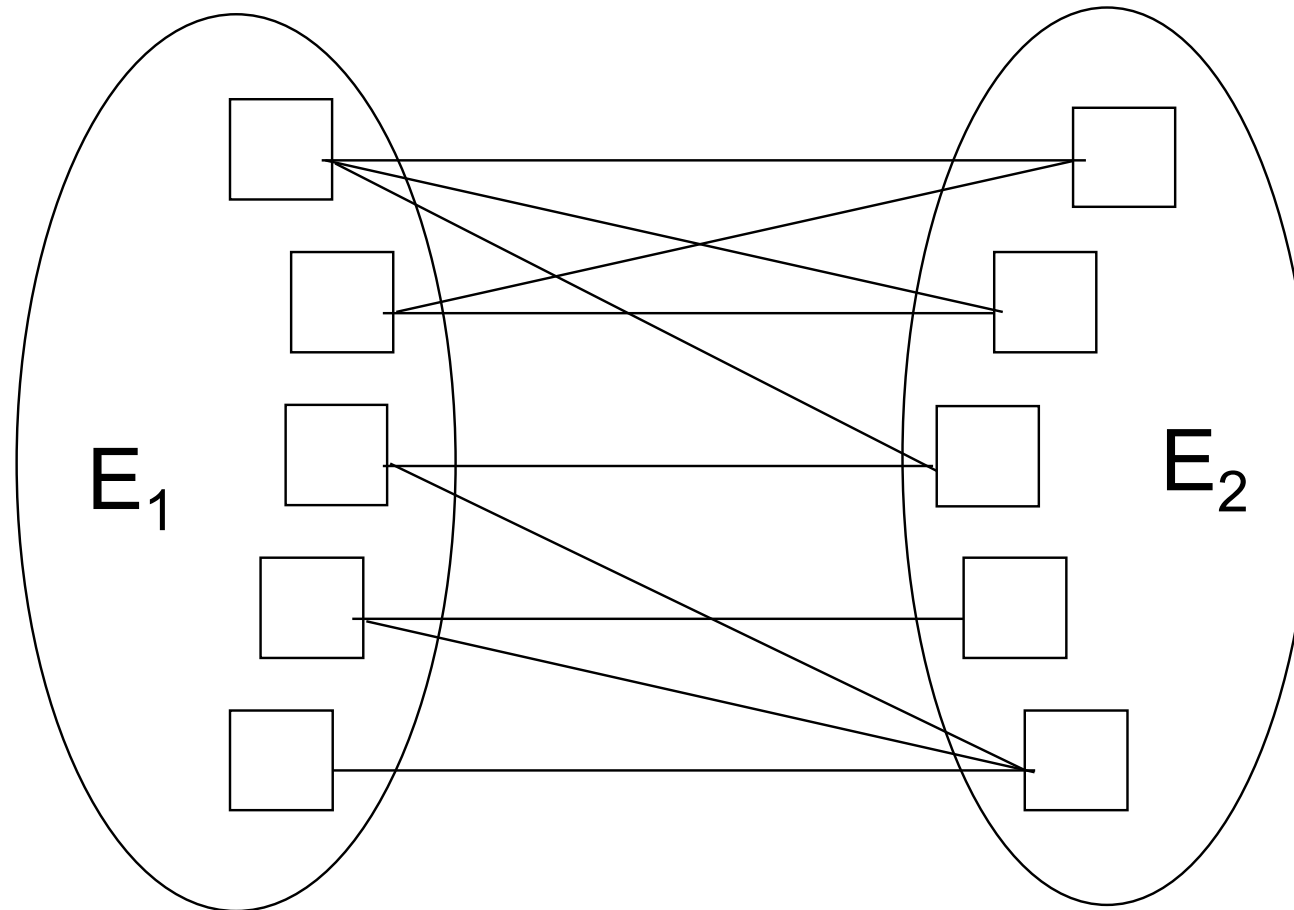


- Kardinalitäten legen fest, **wieviele Entitäten** eines Entitätstypen mit einer Entität des verbundenen Entitätstypen (und umgekehrt) **in Beziehung stehen können oder müssen**





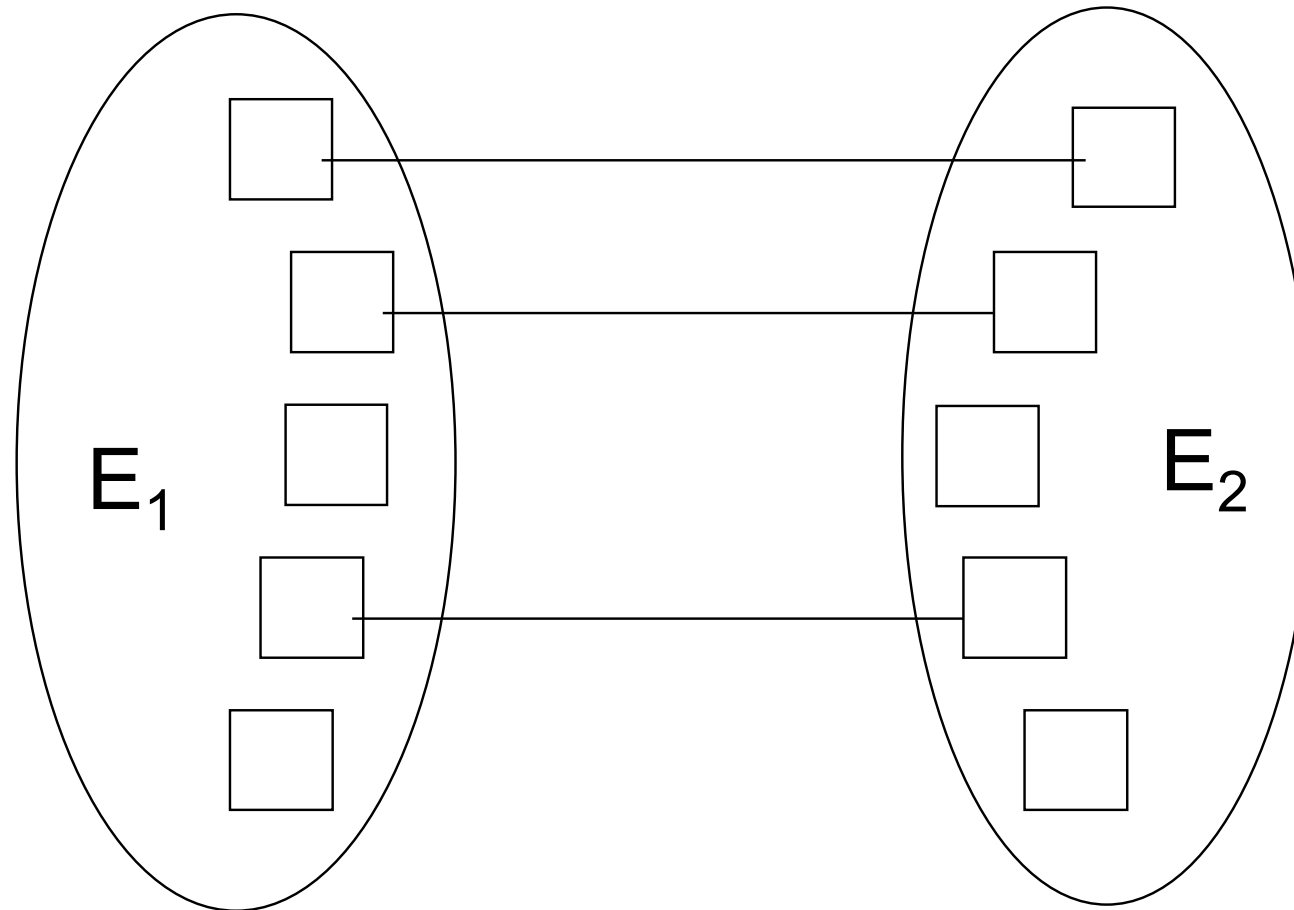




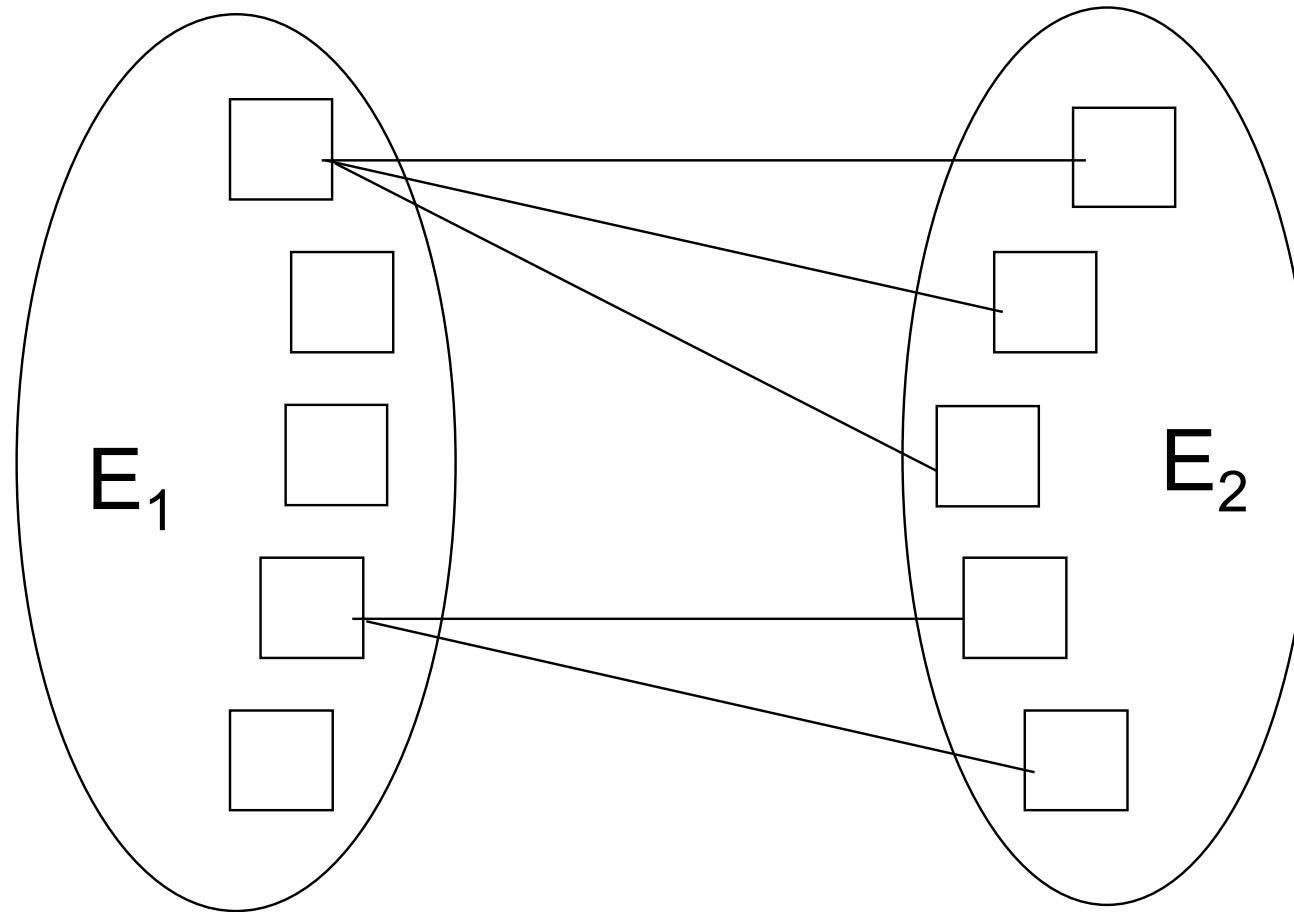
- 1** Konzeptionelle Modellierung
- 2** Einführung Datenmodellierung
- 3** Grundlagen ER Modellierung
 - 3.1** Attribute und Schlüssel
 - 3.2** Beziehungen
 - 3.3** Kardinalitäten
 - 3.4** Zeit in Datenmodellen
 - 3.5** Rekursive Beziehungen
 - 3.6** Mehrwertige Beziehungen

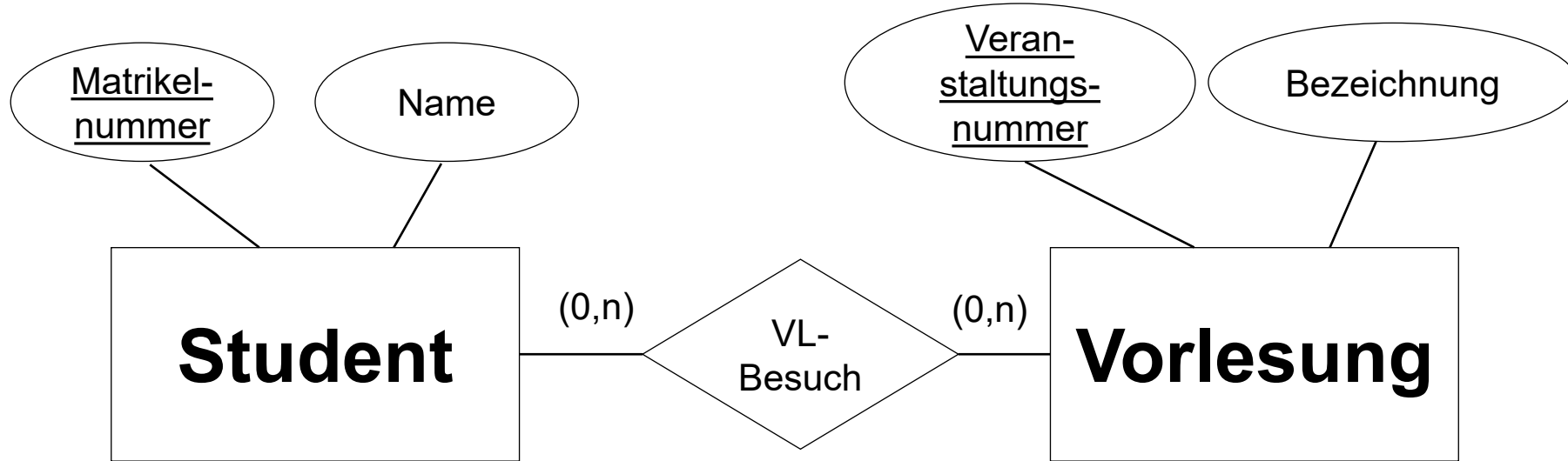
- Aber wie unterscheidet man **kann** von **muss**?
- **Min-Max-Notation**
- Wie oft eine muss Entität mindestens eine Beziehung eingehen und wie oft kann sie maximal eine Beziehung eingehen
- 2 (min, max)-Paare notwendig
- Bei einer Minimum-Kardinalität mit dem Wert 1, spricht man von einer **existenziellen Abhängigkeit**

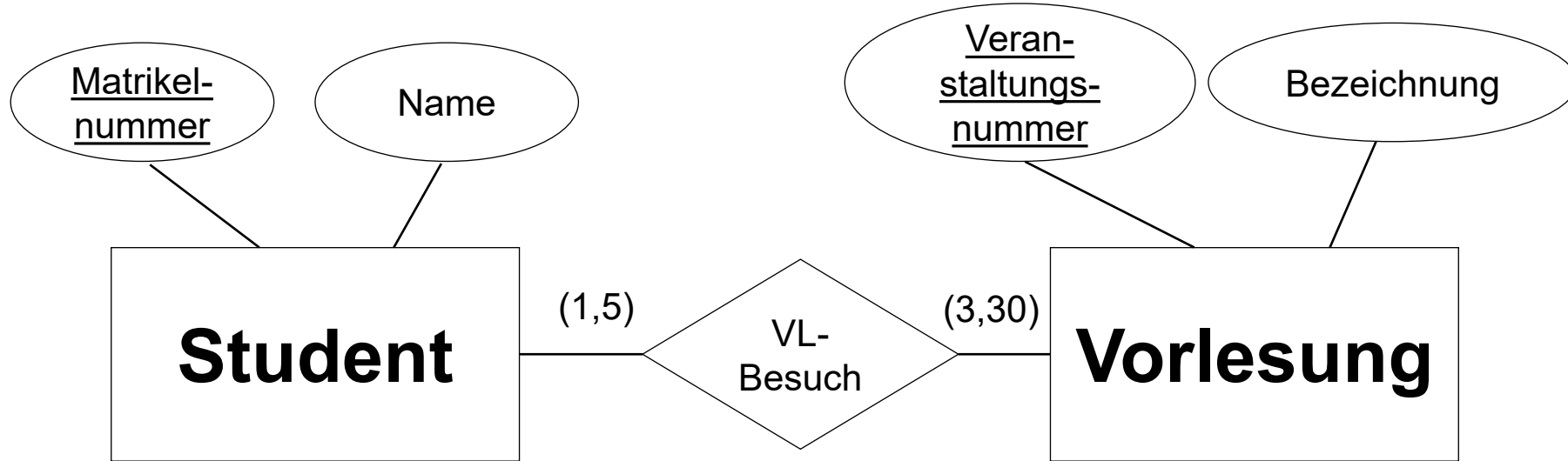
$(0,1) : (0,1)$



(0,n) : (1,1)

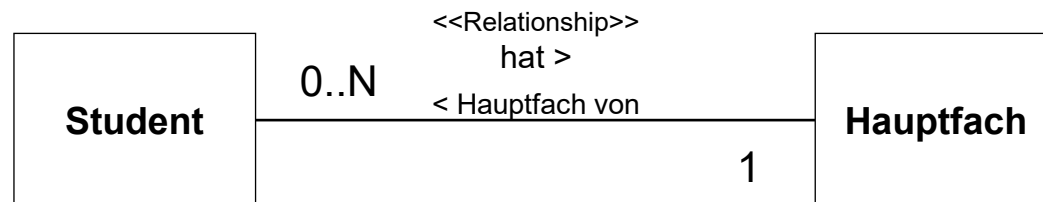
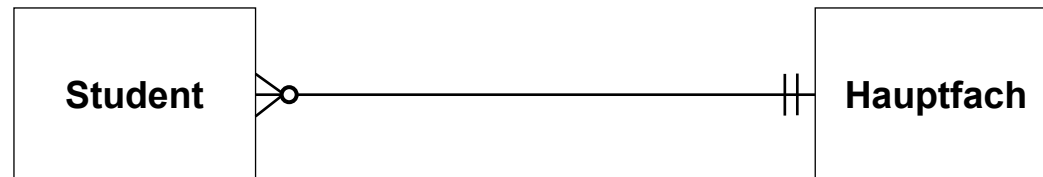
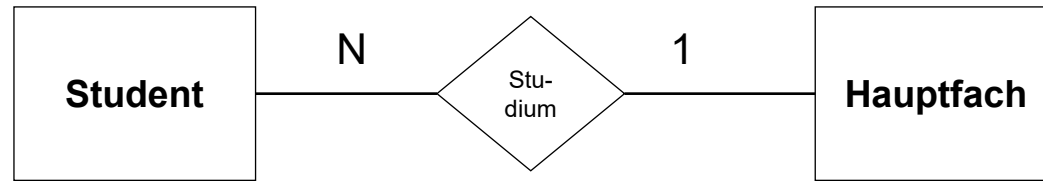
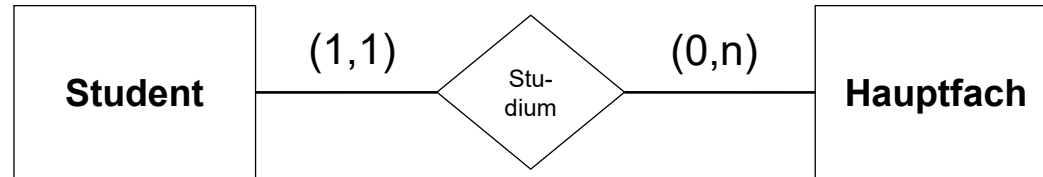






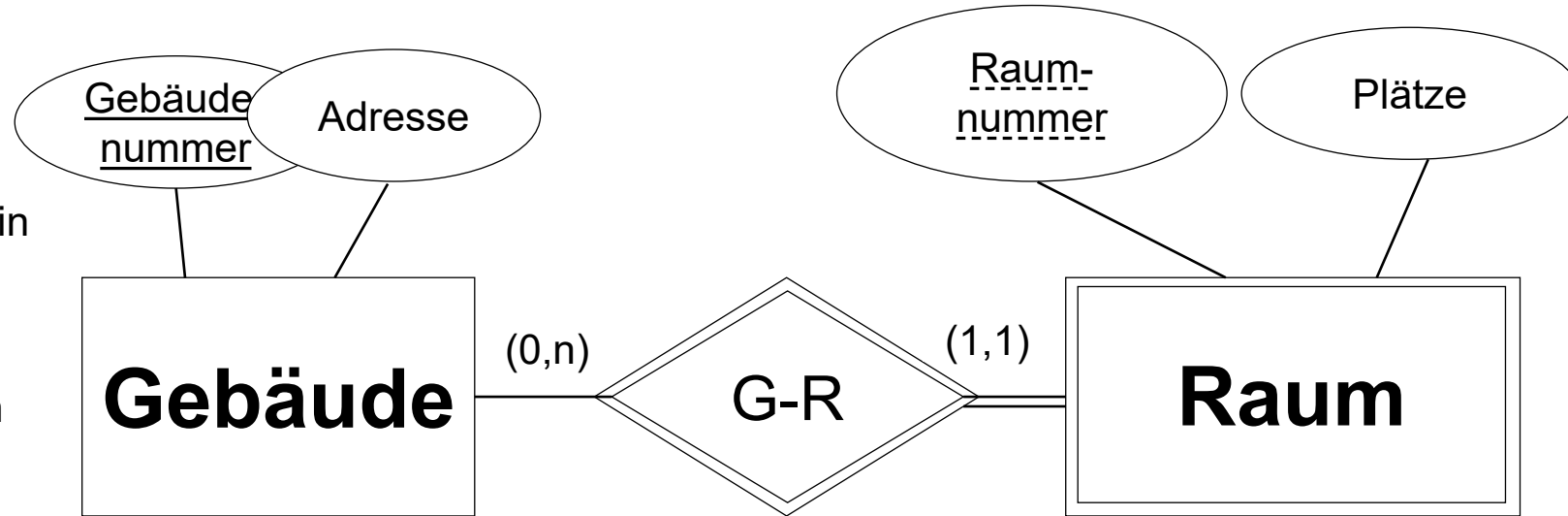
Notationsvergleich

- Min-Max / ISO
- Chen
- Krähenfuß
- UML

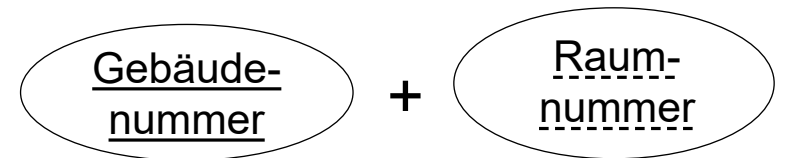


Schwache Entitäten

- Sind **abhängig** vom zugeordneten Entitätstyp
- Können selbst wiederum mit weiteren normalen oder schwachen Entitätstypen in Beziehung stehen
- Haben **keinen** eindeutigen Schlüssel
- Eindeutiger Schlüssel nur in Kombination mit zugeordnetem Entitätstyp
- Modellierung
 - Werden durch doppelte Umrandung dargestellt [in der Praxis oft nicht...]
 - Beziehungstyp zum abhängigen Entitätstyp ebenfalls mit doppelter Umrandung
 - Schlüssel (nicht eindeutig) wird durch gestrichelte Linie hervorgehoben



Schlüssel:



- 1** Konzeptionelle Modellierung
- 2** Einführung Datenmodellierung
- 3** Grundlagen ER Modellierung
 - 3.1** Attribute und Schlüssel
 - 3.2** Beziehungen
 - 3.3** Kardinalitäten
 - 3.4** Zeit in Datenmodellen
 - 3.5** Rekursive Beziehungen
 - 3.6** Mehrwertige Beziehungen

Die Beziehung zwischen Entitätstypen ist abhängig von der **Betrachtungsweise** und vom **gewählten Zeithorizont**

Beispiel Bibliothek

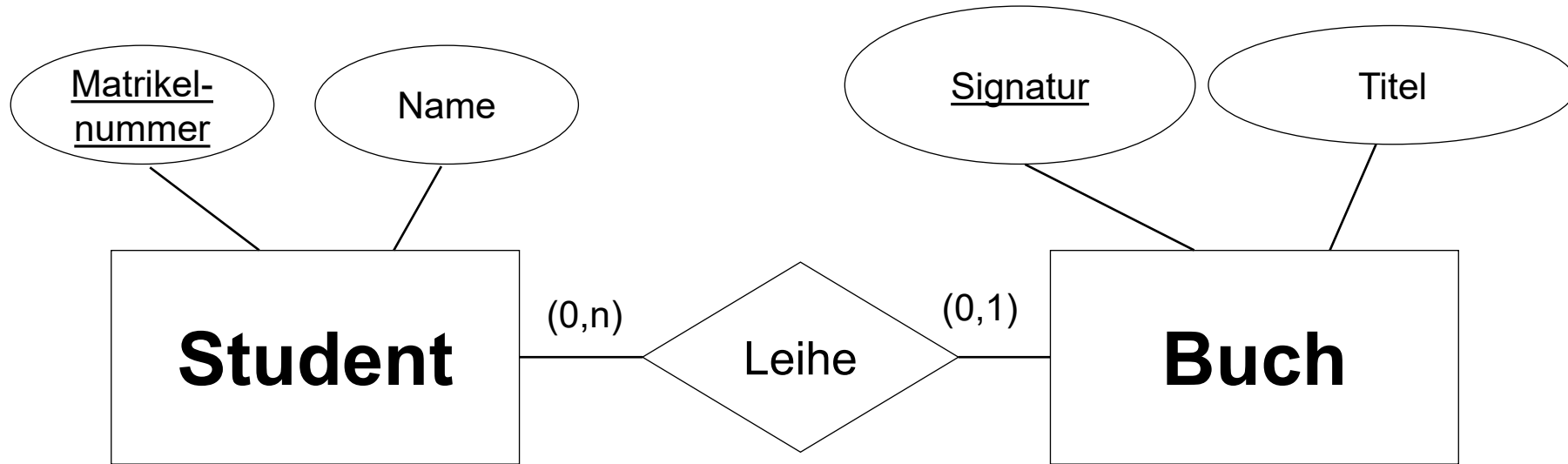
- Ein Student kann mehr als ein Buch ausleihen, aber **ein Buch kann zur gleichen Zeit immer nur von einem Studenten ausgeliehen werden**

Student : Buch = 1 : n

- Betrachtet man einen längeren Zeitraum, so **kann dasselbe Buch nacheinander von mehreren Studenten ausgeliehen werden**

Student : Buch = m : n

Beziehungen ohne Attribut



Modellierung der Zeit: Beziehungen mit Attribut

