



Datenbanken und SQL

(Woche 1 - Tag 1)

Agenda

Grundbegriffe

- Daten
- Datenbank (Datenbanksystem)
- Entität und Entitätstyp
- Attribut
- Datensatz
- Relation

Entity-Relationship-Modell/Diagramm (erste Einführung)

- Definition
- Motivation (+ kurzer Exkurs zum Datenbankentwurfs-Prozess)
- Vorstellung der Symbole (Chen-Notation)
- Verwendung von DIA

ERM
=
ERD

Grundbegriffe

Daten (Einzahl: **Datum**)

- Unter einem „Datum“ verstehen wir **umgangssprachlich** eine „Termin-Angabe“.
- Fachsprachlich wird der Begriff hingegen weiter gefasst und meint: „**Information**“.
Damit erklärt sich dann aber auch die große Bedeutung von Daten für unser Fach, denn nicht umsonst sprechen wir bei unserer Disziplin von der ...
 - „**IT**“ (Informationstechnologie) *oder auch*
 - „**EDV**“ (Elektronische **Daten**verarbeitung)
- Da durch Daten sehr unterschiedliche Informationen angesprochen werden können, können diese auch von sehr unterschiedlichem **Typ** sein:
 - **Zahlen-Typ** (z.B. Seitenzahl, Streckenlänge, Gewicht ...)
 - **Text-Typ** (Wort/Name, Satz, Kapitel ...)
 - **Audiodatei-Typ** (Sprachnachricht, Musikstück ...)
 - **Grafik/Videodatei-Typ** (Zeichnung, Foto, Film...)

diese Typen werden von uns später tatsächlich verwendet

diese Typen benennen wir nur der Vollständigkeit halber

Datenbank / Datenbanksystem

- Unter einer „Datenbank“ versteht man (im Allgemeinen) eine „Sammlung von Daten“.
- Für **Datenbanken** (aus Sicht der EDV) werden hingegen 3 weitere Forderungen gestellt:
 - Zu einer Datenbank müssen Daten **hinzugefügt** werden können (*man muss weitere Daten „**einpflegen**“ können*)
 - In einer Datenbank müssen Daten **korrigiert** oder **gelöscht** werden können (*man muss die Datenbank „**pflegen**“ können*)
 - Aus einer Datenbank müssen Daten **abgefragt** werden können
- Um innerhalb der EDV deutlich zu machen, dass man unter einer Datenbank mehr als eine „reine Datensammlung“ versteht, kann man an Stelle des Begriffes „Datenbank“ auch den Begriff des „**Datenbanksystems**“ verwenden. Beide Begriffe werden aber synonym verwendet (sind also austauschbar).

Szenario:

Die Firma "Hochbau" besteht aus mehreren **Abteilungen**. Die Abteilungen haben jeweils eine **eindeutige Abteilungsnummer** und einen **beliebigen Namen**. Alle **Mitarbeiter** haben eine **eindeutige Mitarbeiter-Nummer**. Die **Mitarbeiter gehören genau einer Abteilung an**. In jeder Abteilung können mehrere Mitarbeiter sein. Für jeden Mitarbeiter ist zu speichern, ob er über eine **Maschinenberechtigung** verfügt. Zu jedem Mitarbeiter müssen der **Name** und die **Postleitzahl des Wohnorts** gespeichert werden.

Das Unternehmen arbeitet auf verschiedenen **Baustellen**. Die Baustellen haben eine **eindeutige Baustellennummer** und einen beliebigen **Baustellennamen**. Die **Mitarbeiter können auf mehreren Baustellen tätig sein**. Auf jeder Baustelle können mehrere **Mitarbeiter tätig sein**. Für jeden Mitarbeiter soll erfasst werden, **wieviel Stunden er auf welcher Baustelle gearbeitet hat**.

Verwenden Sie folgende Relation als Ausgangspunkt:

| Baustellen-nummer | Baustellen-name | Baustellen-Stunden | Abteilungs-nummer ✓ | Abteilungs-name ✓ | Maschinen-berechtigung | MA-Nummer ✓ | MA-Name ✓ | MA-PLZ ✓ |
|----------------------|---------------------------------|--------------------|---------------------|-------------------|------------------------|-------------|-----------|----------|
| B021 B112 | MIDL Kaufstadt | 12 23 | 12 | Ausbau | J - | M010 | Stein | 04838 |
| B253 | GaleriaX | 37 | 9 | Hochbau | N - | M009 | Örtel | 04105 |
| B056 B112 B253 | Brutto Kaufstadt GaleriaX | 21 24 34 | 10 | Haustechnik | J | M021 | Hahn | 04509 |
| B056 B253 | Brutto GaleriaX | 8 24 | 9 | Hochbau | N | M024 | Holzer | 04119 |

Erstellen Sie ein Entity Relationship Model als Konzept für eine relationale Datenbank für die Speicherung der Daten aus dem beschriebenen Szenario.

1. ERM
↓
2. RM
↓
3. DB

Entität / Entitätstyp

- Mittels Daten werden Informationen angesprochen, die „**Etwas**“ beschreiben.
- Dieses „Etwas“ kann (z.B.) eine Person, ein Gegenstand, ein Ereignis ... sein.
- Anstelle von einem „Etwas“ werden wir zukünftig von einer „**Entität**“ sprechen.
- Eine Datenbank wird üblicherweise Daten zu mehreren Entitäten beinhalten.

Betrachten wir z.B. eine Datenbank, in der wir Informationen zu mehreren Personen und mehreren Autos vorfinden:

- Jede einzelne, konkrete Person (und jedes einzelne Auto) stellt eine eigenständige Entität dar.
- Aber allen Personen ist gemeinsam, dass sie zum selben **Entitätstyp** „Person“ gezählt werden können.
(so wie allen Autos gemeinsam ist, dass diese unter dem **Entitätstyp** „Auto“ zusammengefasst werden können)

Attribut

- Daten einer Datenbank beschreiben also **Eigenschaften** von Entitäten.
- Anstelle von „Eigenschaften“ werden wir zukünftig von „**Attributen**“ sprechen.
- Mehrere Entitäten, die zum selben Entitätstypen gehören, werden auch durch die selben Attribute beschrieben. Daher können wir diese Attribute dem gemeinsamen Entitätstypen zuordnen (und müssen das nicht für jede Entität wiederholen), ...

Beispiel:

Wir könnten in einer Datenbank (z.B.) für jede konkrete Personen-Entität mitteilen, dass wir uns jedes mal für die Attribute „Vorname“ und „Geburtsdatum“ interessieren.

*Weil dies aber umständlich wäre, werden wir darauf verzichten und einmalig mitteilen, dass sich diese Attribute auf den **Entitätstyp** „Person“ beziehen (und damit dann eben auch auf jede konkrete Entität vom Typ Person).*

Datensatz

- Einem Entitätstypen werden üblicherweise mehrere Attribute zugeordnet.
- Dies bedeutet dann aber, dass für die Beschreibung einer konkreten Entität (dieses Entitätstyps) entsprechend auch mehrere Daten benötigt werden, ...

Beispiel:

*Angenommen in einer Datenbank wurde zunächst der Entitätstyp **Auto** (zusammen mit seinen Attributen **Auto_ID**, **Marke**, und **Baujahr**) eingeführt, um anschließend die Entität **(1, „Audi“, 1998)** abspeichern zu können, dann gilt:*

Diese drei Daten bilden zusammen eine „Gruppe von Daten“ und gehören in sofern zusammen, als dass alle drei Daten die selbe Entität beschreiben.

- Eine solche Gruppe wird in der Fachsprache als **„Datensatz“** bezeichnet.

Beziehung (Relation)

- Verschiedene Entitätstypen (bzw. deren Entitäten) können in einer Beziehung zu anderen Entitätstypen (bzw. deren Entitäten) stehen, ...

Beispiel:

In einer Datenbank wurden die Entitätstypen „Person“ und „Auto“ eingeführt. Wenn man nun (z.B.) von jeder Auto-Entität wissen möchte, welche Person-Entität der Eigentümer dieses Autos ist, so bedeutet dies, dass man sich für eine Beziehung zwischen den Entitätstypen Auto und Person interessiert.

- Anstelle von „Beziehung“ kann man (in diesem Fall *) auch von „**Relation**“ sprechen ...

Hinweis (*):

Leider gilt dies nicht durchgehend. Tatsächlich werden wir in einigen Tagen einen Themenbereich kennenlernen, in dem der Begriff „Relation“ anders gebraucht wird.

ERM

(Entity-Relationship-Modell)

ERD

(Entity-Relationship-Diagramm)

ERM/ERD -> Definition

- Ein Entity-Relationship-*Modell* ist eine **gedankliche Zusammenfassung aller Elemente**, die für den Aufbau einer geplanten Datenbank von Bedeutung sein werden.
- Ein Entity-Relationship-*Diagramm* ist die **graphische Darstellung dieser Elemente**.
- Der Name dieses Modells macht deutlich, dass die zu erstellenden Diagramme im wesentlichen aus der Darstellung von **Entitätstypen** und **Relationen** bestehen.
- ER-Diagramme können als **(grober) „Bauplan“** einer **leeren** Datenbank betrachtet werden. (Entsprechend werden zu dieser Planungsphase **noch keine Daten** berücksichtigt.)
- ER-Diagramme können auf unterschiedliche Weise notiert werden. Wir werden in diesem Kurs die klassische (und von der IHK erwartete) **„Chen-Notation“** verwenden.

ERM/ERD -> Motivation

- Mit Hilfe von ER-Diagrammen können Datenbanken **übersichtlich dargestellt** werden.
- Zudem präsentierte ihr Erfinder (Peter Chen) mit seiner gleichnamigen Notation ein erstes Beispiel für eine **graphische Softwareentwicklung**.
- Darüber hinaus nutzen wir diese Modellierung (bzw. deren Diagramme) als ersten Schritt eines **Datenbankentwurfs-Prozesses** (näheres: siehe folgende Folie).

ERM/ERD -> Kurzer Exkurs zum Datenbankentwurfs-Prozess

- Wer den Auftrag erhält, eine Datenbank zu entwickeln, der wird in der Regel die Erfahrung machen, dass der Auftraggeber eine **zunächst nur unspezifische Vorstellung** vom Aufbau der gewünschten Datenbank hat.
- Die Entwicklung einer Datenbank sollte daher (mindestens) in die folgenden **3 Teilschritte** zerlegt werden:
 - **Konzeptioneller Datenbankentwurf** (Ermittlung aller Entitätstypen, Attribute und Relationen -> Darstellung mittels **ERD**)
 - **Logischer Datenbankentwurf** (Übergang zum „Relationalen Modell“ -> Näheres dazu in den nächsten Tagen)
 - **Physischer Datenbankentwurf** (Implementierung des Relationalen Modells -> wird mittels SQL geschehen)
→ Umsetzung

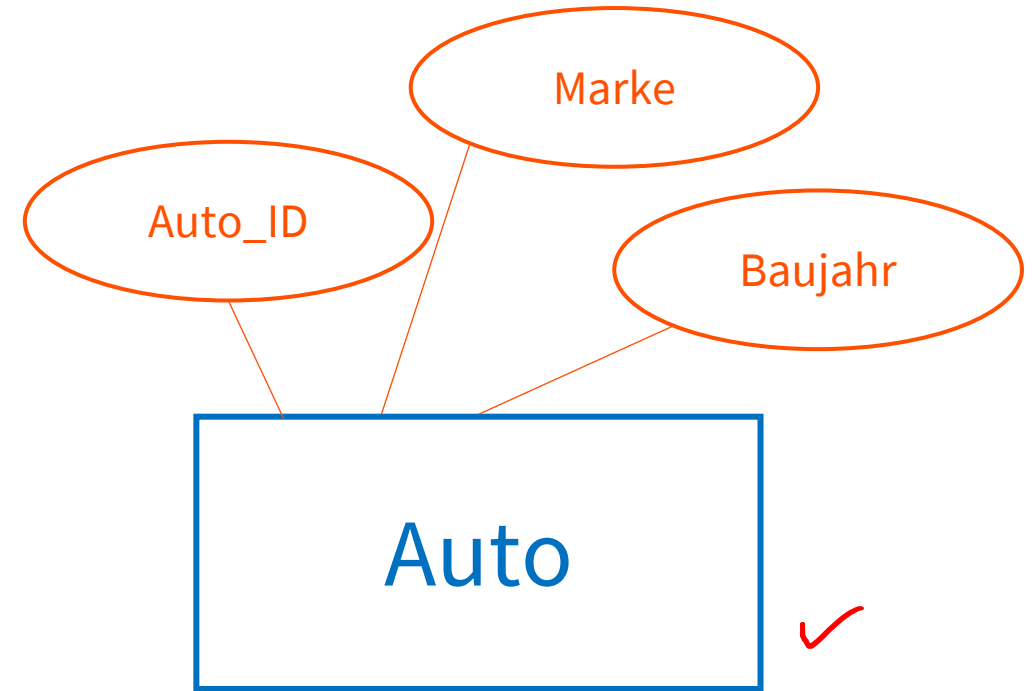
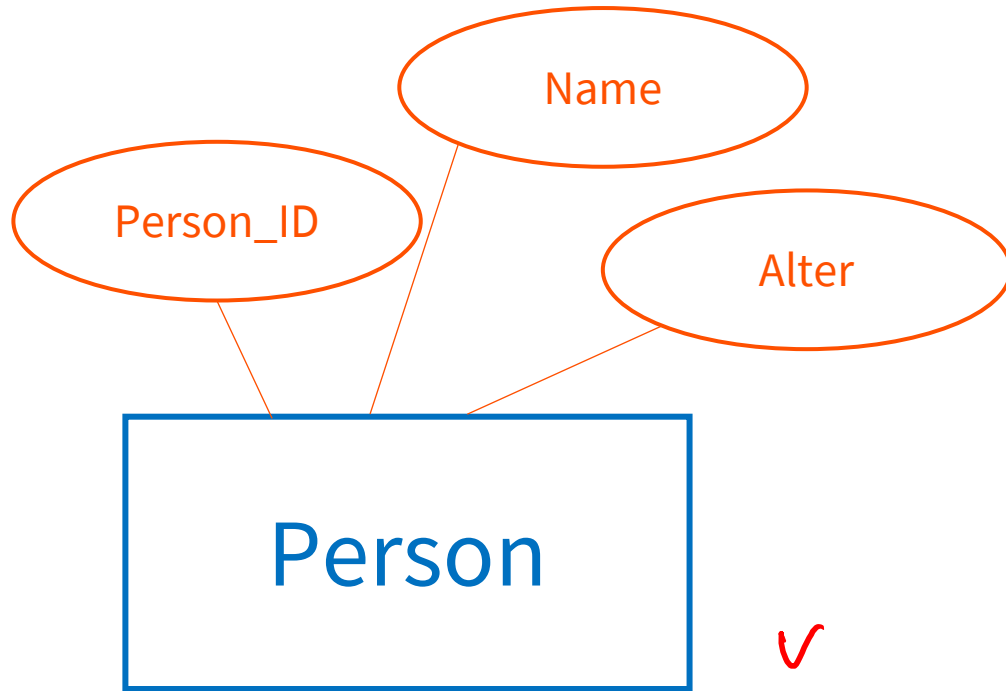
ERD-Symbole

ERD-Symbole -> Entitätstypen

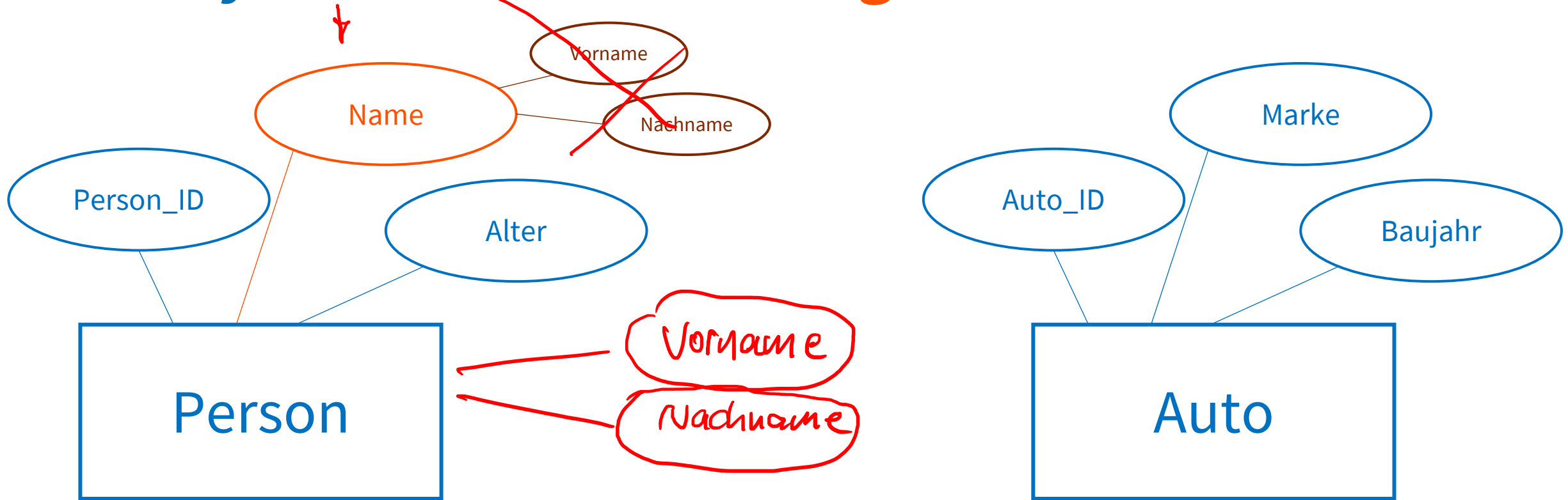
Person

Auto

ERD-Symbole -> Attribute



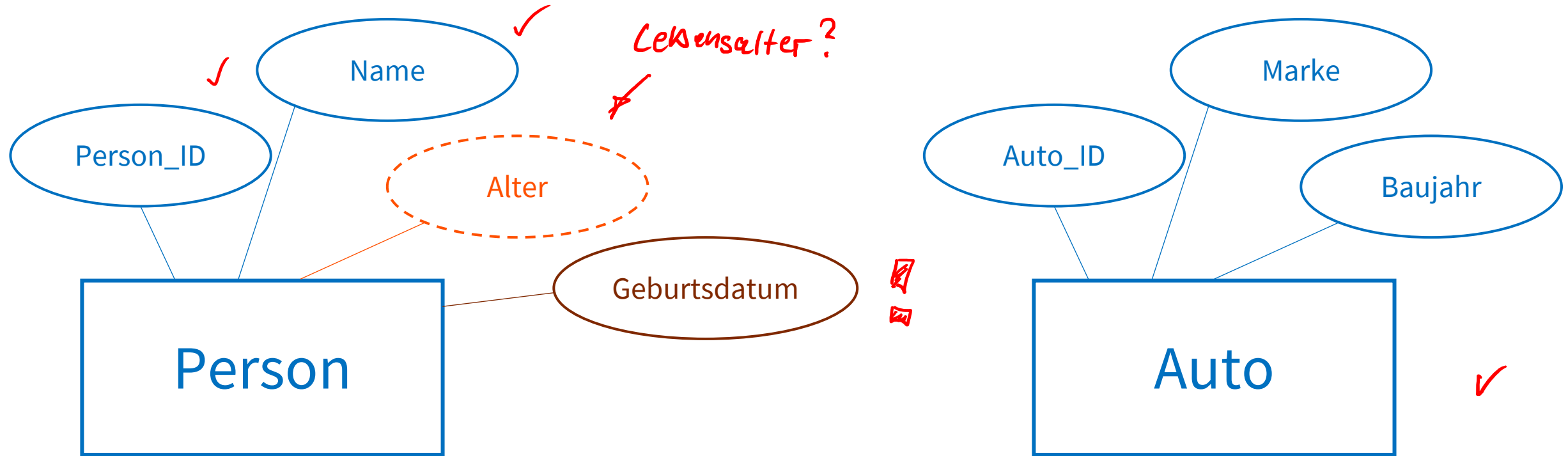
ERD-Symbole -> zusammengesetzte Attribute



Wir werden später erfahren, dass man im Rahmen des **Logischen Datenbankentwurfes üblicherweise** fordert, dass das zusammengesetzte Attribut „Name“ in die Attribute „Vorname“ und „Nachname“ zerlegt wird.

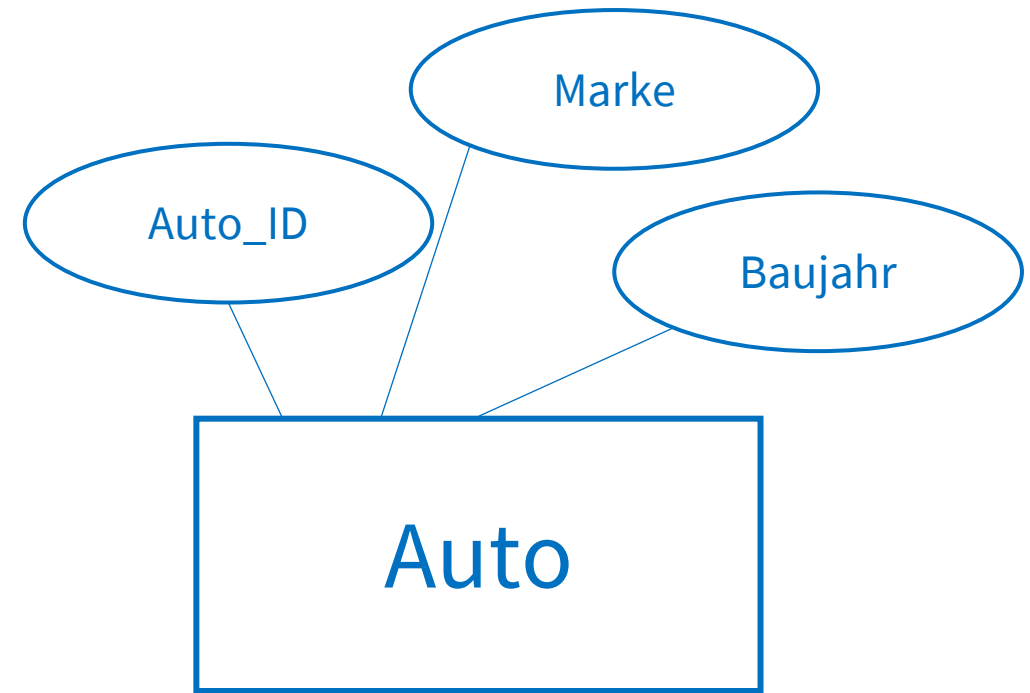
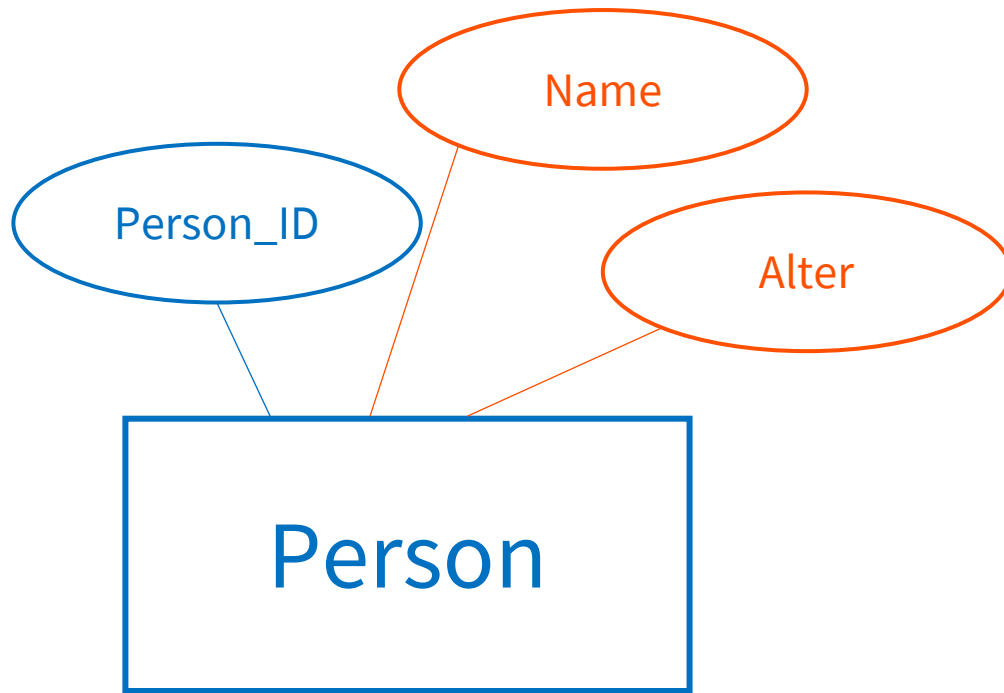
Während des **Konzeptionellen Datenbankentwurfs** wird dies jedoch **nicht zwingend** verlangt.

ERD-Symbole -> **abgeleitete Attribute**



Wir werden später beim **Logischen Datenbankentwurf** darauf achten, dass wir veränderliche Größen von konstanten Größen ableiten:
Das sich jedes Jahr ändernde „Alter“ können wir aus dem konstanten „Geburtsdatum“ ableiten.

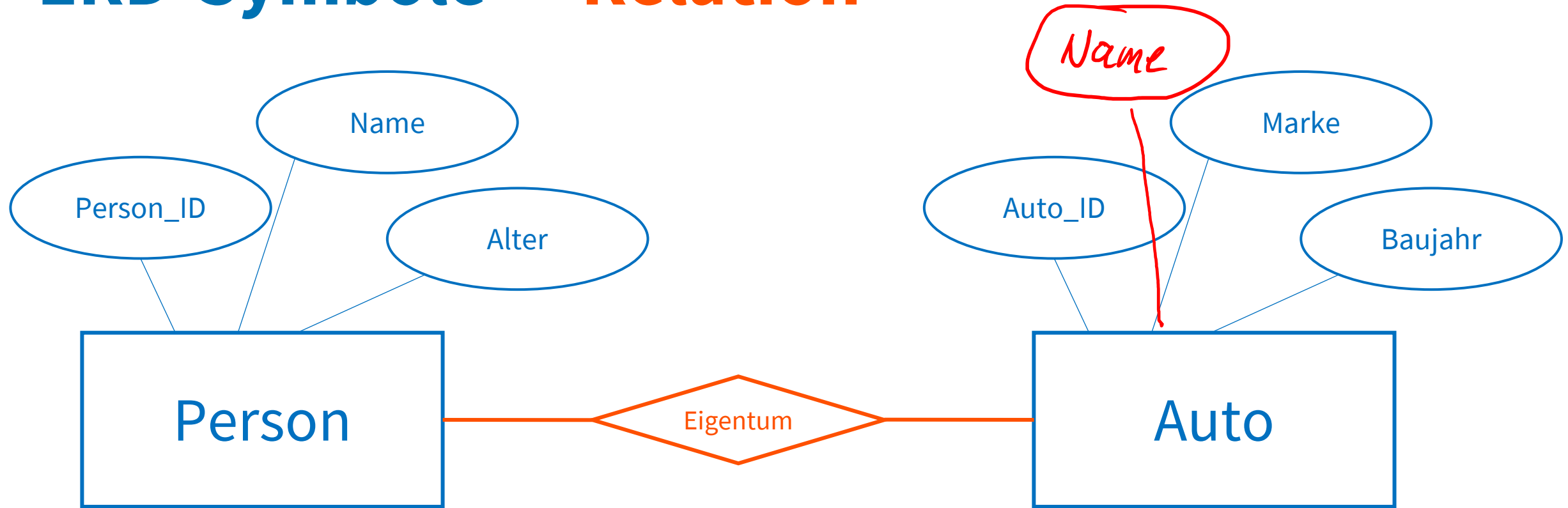
ERD-Symbole -> zusammengesetzte oder abgeleitete Attribute



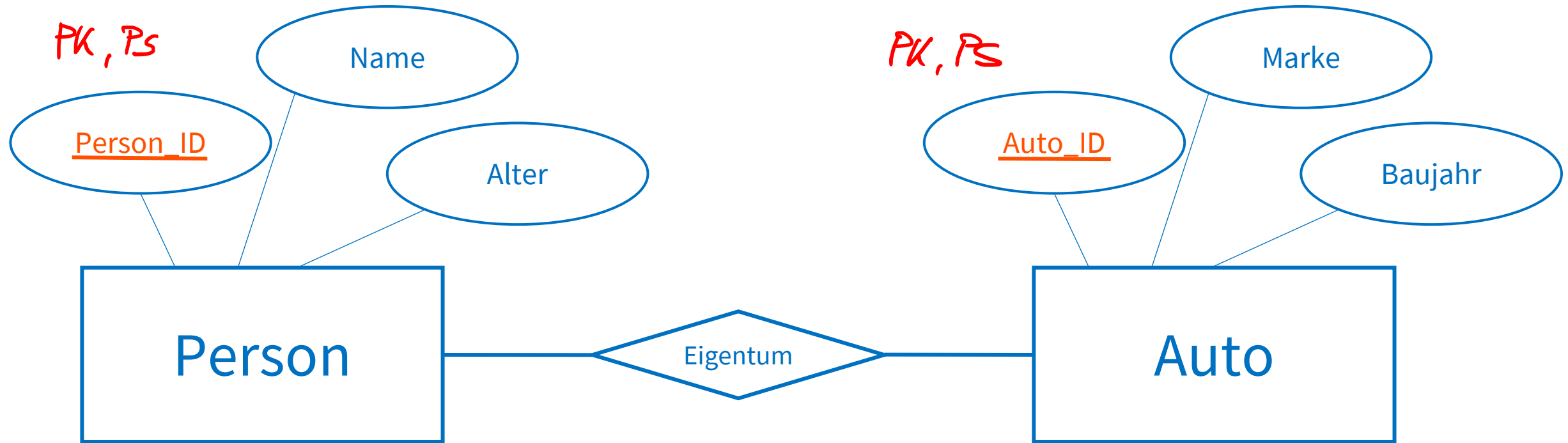
Während des **Konzeptionellen Datenbankentwurfs** dürfen diese beiden Aspekte aber noch **unberücksichtigt bleiben, weil ...**

... es in dieser ersten Entwurfsphase lediglich darum geht, alle vom Auftraggeber gewünschten **Elemente des Datenbankaufbaus** zu ermitteln. Fragen zur **technischen Umsetzung** dieser Elemente können daher zu diesem Zeitpunkt noch zurückgestellt werden.

ERD-Symbole -> Relation

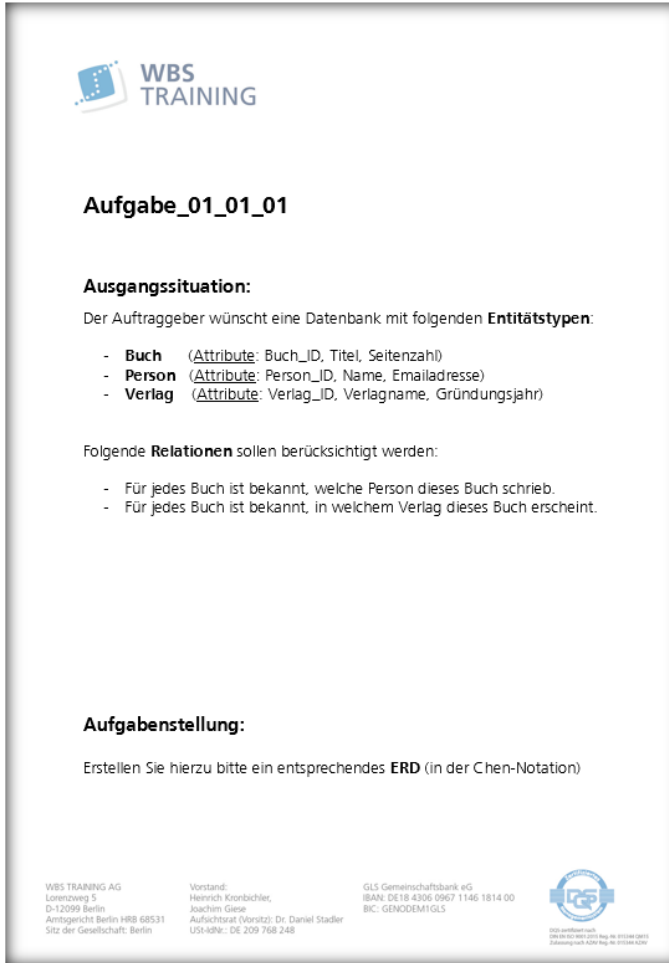


ERD-Symbole -> ID („Schlüssel“)



Wir wollen uns von Beginn an daran gewöhnen, jedem Entitätstypen ein **ID-Attribut** zuzuordnen (das mittels Unterstreichung kenntlich gemacht wird) ...
... die **zentrale Bedeutung** dieser Attribute für unser gesamtes Themenfeld werden wir aber erst in den nächsten Tagen verstehen lernen.

Verwendung von DfA („Live-Coding“) -> A_01_01_01



Vielen Dank für Ihre Aufmerksamkeit!

