

Normalisierung von Datenbanken

Normalisierung

Unter **Normalisierung** eines relationalen Datenschemas versteht man die Aufteilung von Attributen in mehrere Relationen gemäß den Normalisierungsregeln. Ziel ist ein Datenschema in einer Form, die

- Datenredundanzen vermeidet
- logische Widersprüche in der Datenbasis verhindert
- höchstmögliche Flexibilität und einen schnellen Zugriff gewährleistet

Frei von Redundanzen bedeutet, dass sämtliche Daten **nur einmal erfasst** werden. Sprich man muss bei einer Änderung die Daten auch **nur einmal ändern**.

Sollten redundante Informationen in einer Tabelle vorhanden sein, kann es zu **Anomalien** kommen. Grundsätzlich werden Anomalien in **3 verschiedene Arten** unterschieden. Gehen wir dafür von einem Beispiel aus, bei dem wir alle 3 vorfinden:

Name	A-Nr	Adresse	Abteilung
Bucky Barnes	1	Musterstraße 2; 12345 Musterhausen	1; Sicherheit
Steve Rogers	2	Testallee 3; 54321 Teststadt	3; Personal
Tony Stark	3	Beispielweg 6; 67890 Beispieldorf	2; IT
Peter Parker	4	Musterburg 1; 12345 Musterhausen	3; Personal

①

-Anomalie: Datensatz kann nur werden kann, wenn ein anderer Datensatz auch hinzugefügt wird. Z.B. eine nur hinzugefügt werden, wenn auch ein jeweiliger Mitarbeiter hinzugefügt wird.

-Anomalie: Datensätze, die aus Versehen werden. So würden, wenn man den Mitarbeiter "Bucky Barnes" löschen würde, alle Informationen über die Abteilung "Sicherheit" ebenfalls gehen.

-Anomalie: bei einer Änderung müssen immer Datensätze geändert werden. → oder übersehen von Datensätzen
Wenn zum Beispiel die Abteilung "Personal" ihren Namen ändern würde, müsste dies in geändert werden.