

Lernzettel

Informationssysteme und Datenanalyse

Universität: Technische Universität Berlin
Kurs/Modul: Informationssysteme und Datenanalyse
Erstellungsdatum: September 19, 2025



Zielorientierte Lerninhalte, kostenlos!
Entdecke zugeschnittene Materialien für deine Kurse:

<https://study.AllWeCanLearn.com>

Informationssysteme und Datenanalyse

Lernzettel: Informationssysteme und Datenanalyse

(1) Architektur von Informationssystemen. Information Systems Architecture umfasst Struktur, Interaktion und Governance von Quellen, Speichern, Verarbeitung, Präsentation und Steuerung.

- Schichten: Datenebene, Anwendungslogik, Präsentation.
- Typische Komponenten: Datenbanken, Middleware, Analyse-Tools, Dashboards, Sicherheit.
- Ziel: konsistente, zuverlässige und sichere Datenflüsse für Entscheidungen.

(2) Relationales Modell. Daten werden als Relationen $R(A_1, \dots, A_k)$ beschrieben, mit Tupeln $t = (a_1, \dots, a_k)$ und Domänen D_i für jedes Attribut A_i .

$$R(A_1, \dots, A_k), \quad a_i \in D_i$$

(3) Relationale Algebra. Zentrale Operatoren:

$$\sigma_\theta(R) \quad (\text{Selektion})$$

$$\pi_{A_1, \dots, A_m}(R) \quad (\text{Projektion})$$

$$R \bowtie S \quad (\text{Join})$$

$$R \cup S, \quad R \setminus S, \quad R \cap S$$

(4) SQL als Anfragesprache. SQL dient als deklarative Sprache für Abfragen, Transformationen und Persistenz.

Datenbeschreibung (DDL):

```
CREATE TABLE Customers (id INT PRIMARY KEY, name VARCHAR(100), email VARCHAR(100));
```

Datenmanipulation (DML):

```
SELECT name, email FROM Customers WHERE status = 'aktiv';
```

Update-Beispiel:

```
UPDATE Inventory SET stock = stock - 1 WHERE product_id = 123; UPDATE Inventory SET stock =
```

(5) Transaktionssysteme. Transaktionssysteme sichern konsistente Änderungen an Datenbanken.

Wichtige Konzepte:

- ACID-Eigenschaften: Atomicity, Consistency, Isolation, Durability.
- Transaktionen: BEGIN, COMMIT, ROLLBACK.

- Gleichzeitigkeit, Sperren (Locks), Deadlocks, Serialisierbarkeit.

(6) Data Warehousing. Zentralisierte, integrierte Datenspeicher für Entscheidungsunterstützung.

Wichtige Konzepte:

- Star-Schema: Faktentabellen (Fakten) plus Dimensionstabellen (Zeit, Ort, Produkt, Kunde, ...).
- ETL: Extraction, Transformation, Loading.
- OLAP: Mehrdimensionale Analysen, Drill-Down/Roll-Up.
- Data Marts als Teilmengen des Data Warehouse.
- Historisierung und Slowly Changing Dimensions.

(7) Verhältnis von Datenbank-Management zu Datenstrom-Management. DBMS verwalten persistente Daten, Transaktionen und Abfragen; DSMS fokussiert auf kontinuierliche Datenströme und zeitabhängige Abfragen.

Typische DSMS-Konzepte:

- Window-Operatoren, kontinuierliche Abfragen, Streaming-Clients.
- Anwendungen: Finanzströme, Sensoren, Logdaten.

$$\text{Window}(t) = \{x_i \mid t_i \in [t - W, t)\}$$

(8) Data Science Einführung. Data Science verbindet Datenmanagement, Statistik und maschinelles Lernen.

- Pipeline: Datenerfassung → Vorverarbeitung → Merkmalsextraktion → Modellierung → Bewertung → Deployment.
- Ziel: Erkenntnisse und Vorhersagen zur Unterstützung von Entscheidungen.

(9) Datenanalyse-Algorithmen: Klassifikation und Clustering.

- Klassifikation (überwacht): Modelle wie Entscheidungsbäume, k-NN, SVM; Vorhersage

$$\hat{y} = \arg \max_{c \in C} p(y = c \mid x)$$

aus Merkmalen x .

- Clustering (unüberwacht): Gruppenbildung ohne labeln; Beispiele: K-Means, DBSCAN.

$$\min_{\{\mu_j\}} \sum_{i=1}^n \|x_i - \mu_{z_i}\|^2$$

mit Zuweisung z_i zu einem Cluster.

(10) Zusammenfassung und Ausblick. Kernkonzepte: Architektur, relationales Modell, Abfragesprachen (SQL), Transaktionen, Data Warehousing, Datenströme, Data Science und typische Algorithmen. Diese Bausteine unterstützen datengetriebene Entscheidungen in Unternehmen.