

Datenmanagement

Wechseln zu: [Navigation](#), [Suche](#)

Dieser Artikel erfüllt die [GlossarWiki-Qualitätsanforderungen](#) nur teilweise:

Korrektheit: 3 (zu größeren Teilen überprüft)	Umfang: 2 (wichtige Fakten fehlen)	Quellenangaben : 2 (wichtige Quellen fehlen)	Quellenarten: 5 (ausgezeichnet)	Konformität: 5 (ausgezeichnet)
---	--	---	---	--

Inhaltsverzeichnis

- 1 Definition (Kowarschick)^[1]
- 2 Bedeutung des Datenmanagements (vgl. Bayer (1996))^[2]
- 3 Spezielle Begriffe des Datenmanagements (Kowarschick^{[1][9]})
- 4 Anforderungen an ein Datenbankmanagementsystem
- 5 Übersicht über verschiedene Arten von Datenmanagement
- 6 Quellen

1 Definition (Kowarschick)^[1]

Unter Datenmanagement versteht man die Verwaltung von **Daten**, d. h. die Konstruktion und Modifikation eines **Datenmodells** sowie die Konstruktion, Modifikation, Selektion, Präsentation und Destruktion von zugehörigen Daten.

2 Bedeutung des Datenmanagements (vgl. Bayer (1996))^[2]

Der Mensch kann die Welt/Realität nicht direkt erfassen. Er bildet letztlich immer (zumindest auf geistiger Ebene) **Modelle**^[3] der Realität. Neue Erkenntnisse kann er nur gewinnen, wenn er neue **Fakten** in seine Modellwelt einordnen oder die Modellwelt geeignet abändern kann.

Die Abbildung eines Realitätsausschnittes auf Fakten eines Modells oder stellt immer einen so genannten *Abstraktionsschritt* dar. Abstraktion ermöglicht einerseits ein besseres Verständnis, bedeutet andererseits aber einen Präzisionsverlust, d. h., (hoffentlich) unwesentliche Teile der Realität werden vernachlässigt. Der Präzisionsverlust ist allerdings immer erheblich. Man denke nur an den Unterschied zwischen einer Landkarte und dem realen Gebiet, das damit modellhaft dargestellt wird.

Realitätsausschnitt	Datenmodell	Daten
Musik	Notensystem	Eine kleine Nachtmusik, Hänschen klein
Chemie	Chemische Formeln	\$H_2O\$, \$CO_2\$
Städte, Länder ...	Landkarten	Augsburger Stadtplan, Weltatlas
	Sprache	Bücher, Nachrichten, E-Mails

In der Mathematik und der Informatik wird die Realität ebenfalls auf abstrakte Modelle abgebildet, allerdings wesentlich formaler als dies mit natürlichen Sprachen möglich ist.

Fakten werden in der Informatik normalerweise in Form von sogenannten **Daten**, d.h. als **Zeichenketten** oder Bitfolgen modelliert. Für gegebene Realitätsausschnitte können, um die Bedeutung von Daten festzulegen, so genannte **Datenmodelle** oder **Datenschemata** entwickelt werden.

Damit ist es möglich, Wissen über die reale Welt dauerhaft auf **digitalen Datenträgern** zu speichern. Das Datenmanagement hat nun zur Aufgabe, gespeichertes Wissen bei Bedarf abrufen, neuen Gegebenheiten anpassen und – sobald es veraltet ist – auch wieder löschen zu können. Es gibt also vier fundamentale Operationen, die sogenannten **CRUD-Operationen**:

Create (Insert): Neue Daten erstellen.

Read (Select, Retrieve, Search): Bestehende Daten selektieren und zur weiteren Verarbeitung bereit stellen.

Update (Modify): Bestehende Daten aktualisieren.

Delete (Destroy): Veraltete Daten löschen.

In der Datenbankwelt wird eine Sprache, die diese Operationen unterstützt **Data Manipulation Language (DML)** genannt. Bevor man jedoch DML-Operationen verwenden kann, muss man zunächst ein passendes Datenmodell definieren. Dies ist die Aufgabe der **Data Definition Language (DDL)**. Da sich auch das Datenschema in Laufe der Zeit ändern kann – man spricht hier von **Schemaevolution** – benötigt man auch in der DDL CRUD-Operationen zum Erstellen, Aktualisieren und Löschen von Schemainformationen. Die Selektion und Ausgabe von bestehenden Schemainformationen wird dagegen seltener benötigt und ist häufig nur systemspezifisch (proprietär) implementiert. ^{[4][5][6][7][8]}

3 Spezielle Begriffe des Datenmanagements (Kowarschick^{[1][9]})

Datenmanagement erfolgt im Allgemeinen mit Hilfe von **Datenbankmanagementsystemen**, **Content-Management-Systemen**. Allerdings ist Datenmanagement auch mit Hilfe von **Aktenordner**, **Karteikästen** und anderen nicht-informationstechnischen Hilfsmitteln möglich.

Im Zusammenhang mit Datenbanken gibt es fünf wesentliche Begriffe, die man nicht verwechseln sollte:

Datenbankschema/Datenmodell

Ein **Datenmodell** oder **Datenbankschema** legt fest, wie bestimmte Aspekte der „realen“ oder einer „abstrakten Welt“ durch **Daten** repräsentiert werden. Es legt fest, welche Arten von Daten es geben kann, in welchen Beziehungen diese stehen können und welche Bedingungen (**Constraints**) sie erfüllen müssen.

Daten

Daten repräsentieren Elemente der „realen“ oder einer „abstrakten Welt“.

Datenbank (DB)

Eine **Datenbank** ist ein Datenmodell zusammen mit einer **Menge** von zugehörigen Daten, d.h. von Daten, die die im Datenmodell formulierten Bedingungen erfüllen.

Datenbankmanagementsystem (DBMS)

Ein **Datenbankmanagementsystem** ist ein informationstechnisches System zur Verwaltung von Datenbanken (im Sinne des Datenmanagements).

Datenbanksystem (DBS)

Ein aktives Datenbankmanagementsystem zusammen mit beliebig vielen (evtl. auch null) zugehörigen Datenbanken wird **Datenbanksystem** genannt.

Für spezielle Datenbanksysteme wie **Relationale DBS**, **Objektorientiertes DBS**, **Objektrelationales DBS** etc. kann man bei jedem Begriff den jeweiligen den Typ des DBS ergänzen: Relationales Datenmodell, Objektrelationale Datenbank, Objektorientiertes Datenbankmanagementsystem etc. Für **Content-Management-Systeme** gibt es analoge Definitionen.

4 Anforderungen an ein Datenbankmanagementsystem

Welche Anforderungen muss ein Datenbankmanagementsystem erfüllen? Um diese Frage beantworten zu können, muss man sich die Randbedingungen ansehen, unter denen ein DBMS üblicherweise eingesetzt werden. Folgendes lässt ich dabei beobachten:

Daten sind langlebig. Wichtige Daten sollen oft über viele Jahre oder Jahrzehnte hinweg zugänglich sein. Die Hard- und auch die Software, die zur Verwaltung dieser Daten eingesetzt wird, ist meist deutlich kurzlebiger.

Im Laufe der Zeit ändert sich die Welt, die modelliert wird. Und damit ändert sich das zugehörige Datenmodell.

Aufgrund der Langlebigkeit der Daten sind auch Fehler im Datenbestand langlebig.

Bei häufigen Modifikationen des Datenbestands akkumulieren sich die Fehler.

Auf wichtige Daten wird von vielen Personen und Programmen zugegriffen. Man denke nur an bekannte Suchmaschinen, Video- und Foto-Portale, Online-Enzyklopädien etc.

Benutzer warten nicht gerne. Sie erwarten beispielsweise, dass eine Suchmaschine sofort antwortet, auch wenn viele Milliarden Seiten durchsucht werden müssen.

Je mehr Zugriffen auf einen Datenbestand erfolgen, desto wahrscheinlicher ist es, dass mehrere Zugriffe gleichzeitig stattfinden.

Datenverluste können gravierende Folgen haben, wie z. B. den Konkurs eines Unternehmens.

Unerlaubter Zugriff auf Daten kann ebenso gravierende Folgen haben, wenn z. B. Millionen Kreditkartendaten in falsche Hände geraten.

Aus diesen Beobachtungen lassen sich nun einige Anforderungen ableiten, die ein gutes DBMS erfüllen sollte^{[4][5][6][7][8]}:

Es muss **CRUD**-Operationen für das Datenmodell zur Verfügung stellen.

Es muss **CRUD**-Operationen für die Daten zur Verfügung stellen.

Es muss Daten dauerhaft speichern (Dauerhaftigkeit/**Persistenz**).

Es muss mehrbenutzerfähig sein (**Parallelität**, **Transaktionen**).

Es muss den unerlaubten Zugriff auf Daten unterbinden (**Datenschutz: Benutzer- und Rechteverwaltung**).

Es sollte Fehler im Datenbestand nach Möglichkeit vermeiden (**Datenkonsistenz: Integrität, Transaktionen, Trigger**)

Es sollte im Fehlerfall wieder einen konsistenten Zustand herstellen können (**Datensicherheit: Backup und Recovery**).

Es sollte die Daten in einem Standardformat ex- und importieren können.

5 Übersicht über verschiedene Arten von Datenmanagement

Datenmanagement	Atomare Typen	Aggregation unterschiedlicher Datentypen	Aggregation gleichartiger Datentypen	Datenmodell
Relationales Datenbankmanagement	Zahlen, Zeichenketten, Boolean, Zeit, ...	Tupel	Relation/Tabelle	Entity-Relationship-Modell
Objektorientiertes Datenmanagement	Zahlen, Zeichenketten, Boolean ...	Objekte	Klassen (genauer: Klassenextensionen)	UML: Klassendiagramm etc.
Content-Management	Zahlen, Zeichenketten, Boolean, Zeit, ..., (digitale) Medien	Hypermedium		CMS-spezifisch
JSON	Zahlen, Zeichenketten, Boolean	Objekte	Arrays	JSON Schema, JSON Hyper-Schema
XML	Zahlen, Zeichenketten, Boolean, Zeit	Attribute oder alternativ XML-Elemente	XML-Elemente	XML Schema
Mengenlehre	Boolean, leere Menge sowie evtl. Zahlen, Punkte, Linien, ... (abh. von Theorie)	geordnete Paare/Tupel, Mengen und Klassen	Relationen, Funktionen	Formale Logik plus Axiome

Es gibt noch diverses weitere Datenmanagement-Technologien: Netzwerk-Datenmanagement (z.B. Codasyl), Objektrelationales Datenmanagement, Graphenbasiertes Datenmanagement, NoSQL, ...

6 Quellen

- Kowarschick (MMDB):** Wolfgang Kowarschick; Vorlesung „Multimedia-Datenbanksysteme“; Hochschule: Hochschule Augsburg; Adresse: Augsburg; Web-Link; 2016; Quellengüte: 3 (Vorlesung)
- Bayer (1996):** Rudolf Bayer; Vorlesung „Datenbanksysteme“; Hochschule: Technische Universität München; Adresse: München; 1996; Quellengüte: 3 (Vorlesung)
- Stachowiak (1973):** Herbert Stachowiak; Allgemeine Modelltheorie; Verlag: Springer Verlag; Adresse: Wien, New York; Web-Link; 1973; Quellengüte: 5 (Buch), S. 131-133
- Ullman (1988):** Jeffrey D. Ullman; Principles of Database and Knowledge-Base Systems – Volume I: Classical Database Systems; Verlag: Computer Science Press; Adresse: New York, Oxford; ISBN: 0-7167-8158-1; Web-Link; 1988; Quellengüte: 5 (Buch)
- Date, Darwen (1993):** Christopher J. Date und Hugh Darwen; A Guide to the SQL Standard – A user's guid to the standard relational language SQL; Auflage: 3; Verlag: Addison-Wesley; Adresse: Reading, Massachusetts, USA; 1993; Quellengüte: 5 (Buch)
- Conolly, Begg (2002):** Thomas Connolly und Carolyn Begg; Database Systems; Auflage: 3; Verlag: Addison-Wesley; ISBN: 0-201-70857-4; 2002; Quellengüte: 5 (Buch)
- Garcia-Molina, Ullman, Widom (2002):** Hector Garcia-Molina, Jeffrey D. Ullman und Jennifer Widom; Database Systems: The Complete Book; Verlag: Prentice Hall; Adresse: New Jersey, Upper Saddle River; ISBN: 0-13-031995-3; Web-Link; 2002; Quellengüte: 5 (Buch)

Elmasri, Navathe (2011): Ramez Elmasri und Shamkant B. Navathe; Fundamentals of Database Systems; Auflage: 3; Verlag: [Pearson Studium](#); ISBN: 978-0-136-08620-8; 2011; Quellengüte: 5 (Buch)

vgl. **Kowarschick (MMDB-Skript)**: Wolfgang Kowarschick; Vorlesung Multimedia-Datenbanksysteme – Sommersemester 2018; Hochschule: [Hochschule Augsburg](#); Adresse: [Augsburg](#); [Web-Link](#); 2018; Quellengüte: 4 (Skript)

Kategorien:

[Datenmanagement](#)

[Glossar](#)

Diese Seite wurde zuletzt am 31. Juli 2019 um 18:56 Uhr bearbeitet.

Inhalt verfügbar unter [CC BY-SA 4.0](#).

