



FRIEDRICH-SCHILLER-
UNIVERSITÄT
JENA

Datenbanksysteme I

Prof. Dr. Viktor Leis

Professur für Datenbanken und Informationssysteme

Einführung

- Datenbanksysteme werden durch einen Großteil von Softwareapplikationen eingesetzt
- Deswegen lohnt es sich zu wissen wie man Datenbanksysteme *effektiv einsetzt*
- Außerdem ist sehr von Vorteil zu verstehen, wie Datenbanksysteme *funktionieren* (z.B. um die Performance optimieren zu können)
- Datenbank-Vorlesung ist an den meisten Hochschulen verpflichtend

- Online Vorlesung
- Übungsleitung: Gabriel Haas gabriel.haas@uni-jena.de
- Webseite: <https://dbis1.github.io/courses/ss20/db1/>
- Moodle

- Diplom-Studium der Wirtschaftsinformatik an der FH Regensburg
- Software-Entwickler bei einem Startup
- Masterstudium Informatik, Promotion, Postdoc an der TU München
- Mitwirkung an der Entwicklung und Ausgründung des Datenbanksystems HyPer
- Seit April 2019 Professor an der FSU Jena für Datenbanken und Informationssysteme
- Forschungsschwerpunkt: Entwicklung effizienter Datenbanksysteme auf moderner Hardware

- Forschungsthemen:
 - Anfragebearbeitung
 - Transaktionsverwaltung
 - Daten- und Indexstrukturen, Kompression
 - Anfrageoptimierung
 - Ausnutzung moderner Speichertechnologien (SSD, NVM)
- viktor.leis@uni-jena.de
- Ernst-Abbe-Platz 2, Raum 3234

- Wintersemester:
 - Datenbanksysteme I (4 SWS)
 - Datenbankimplementierung (4 SWS)
- Sommersemester:
 - Datenbanksysteme I (4 SWS)
 - Datenbanksysteme II (4 SWS)
 - Software-Entwicklungsprojekt
 - Seminar Moderne Datenbanksysteme
 - Writing for Computer Science

Datenbanksysteme I ist Basis und Voraussetzung für Datenbanksysteme II, Datenbankimplementierung und Datenbanken Spezialisierung

Alfons Kemper und Andre Eickler
Datenbanksysteme – Eine Einführung
10. Auflage
De Gruyter Verlag, München

[http://www-db.in.tum.de/research/publications/
books/DBMSeinf](http://www-db.in.tum.de/research/publications/books/DBMSeinf)

<https://dbis1.github.io/courses/ss20/db1/>

- A. Kemper, A. Eickler: Datenbanksysteme – Eine Einführung. De Gruyter Verlag, 2015. 10. Auflage.
- A. Kemper, M. Wimmer: Übungsbuch Datenbanksysteme. De Gruyter Verlag, 3. Auflage, 2011.
- A. Silberschatz, H. F. Korth und S. Sudarshan: Database System Concepts
- R. Ramakrishnan, J. Gehrke: Database Management Systems

- Was ist ein Datenbanksystem (DBS)?
Ein System zum Speichern und Verwalten von Daten.
- Warum kein herkömmliches Dateisystem verwenden?
Ausfallsicherheit und Skalierbarkeit nur mit hohem Aufwand erreichbar.

Traditionelle Anwendungsgebiete:

- Geschäftsdaten
- Buchhaltung
- Verwaltung
- ...

Heute sehr viel breiter:

- Wissenschaftliche/Medizinische Daten
- Data Mining
- Informationsintegration
- Websuche
- ...

Beispiele(2)

Indirekt benutzen wir ständig Datenbanken:

- Websuche bei Google, Yahoo, ...
- Anfragen bei Amazon, EBay, ...
- Backend vieler großer Webseiten

Viele Spielarten (DB/IR, zentralisiert/dezentralisiert etc.)

Datenbanken werden fast immer eingesetzt wenn

- die Datenmengen groß sind
- die Daten wertvoll sind

Beispiele(3)

Die großen kommerziellen Datenbanksysteme:

- Oracle
- Microsoft SQL Server
- IBM DB2

Weitverbreitete Open Source Datenbanksysteme:

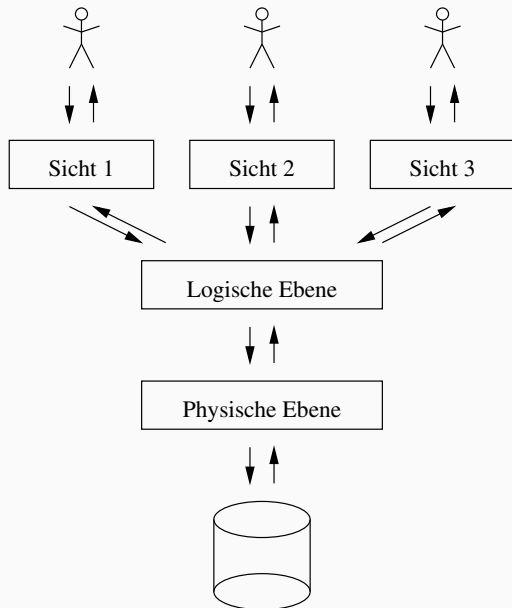
- PostgreSQL
- MySQL
- SQLite

Noch viele weitere, teils stark spezialisierte Systeme.

- Redundanz und Inkonsistenz
- Verschiedene Datenformate
- Probleme beim Mehrbenutzerbetrieb
- Verlust von Daten
- Integritätsverletzungen
- Sicherheitsprobleme
- Hohe Entwicklungskosten für Anwendungsprogramme

- Datenunabhängigkeit
- Deklarative Anfragesprachen
- Mehrbenutzersynchronisation
- Fehlerbehandlung
- Sicherstellung der Datenintegrität
- Effizienz und Skalierbarkeit

Datenunabhängigkeit



- Sicht: beschreibt wie ein Benutzer die Daten sieht
- Logische Ebene: beschreibt wie die Daten strukturiert sind
- Physische Ebene: beschreibt wie die Daten gespeichert werden

- DBS entkoppelt Anwendungen von der eigentlichen Struktur und Speicherung der Daten
- Logische Datenunabhängigkeit
 - Änderungen auf der logischen Ebene haben keinen Einfluss auf Anwendungen
- Physische Datenunabhängigkeit
 - Änderungen auf der physischen Ebene haben keinen Einfluss auf Anwendungen
 - Wird in fast allen modernen DBS durchgesetzt

- Benutzer sagt DBS *was* für Daten geholt werden sollen ...
- ...und nicht *wie* die Daten geholt werden sollen
- Weniger fehleranfällig (beim Formulieren von Anfragen/Entwickeln von Anwendungen), da kein Wissen über die tieferen Schichten des DBS nötig sind
- Mit etwas Erfahrung kann man sehr schnell komplexe Anfragen formulieren

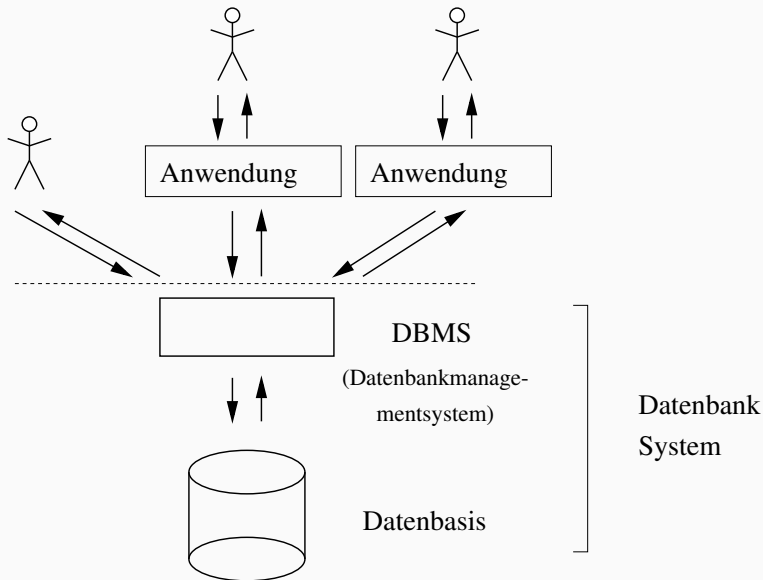
- Wenn mehrere Benutzer ohne jegliche Kontrolle gleichzeitig Daten ändern können, gibt es große Probleme
- DBS erlaubt gleichzeitigen Zugriff und verhindert schlimme Seiteneffekte

- DBS kann Zustand zum Zeitpunkt eines Absturzes rekonstruieren
- Dafür werden Logdateien vom DBS angelegt und verwaltet

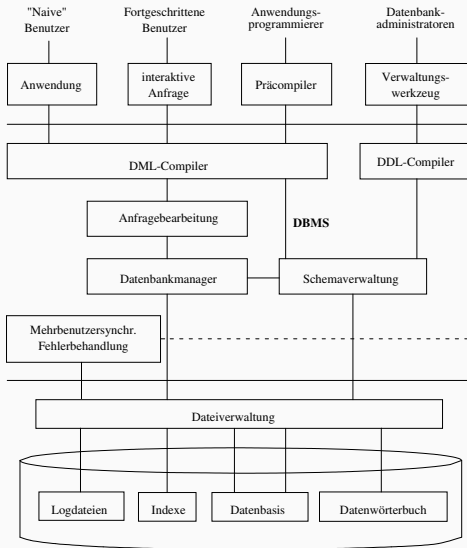
- Datenverarbeitung in einer Anwendung läuft nicht völlig zufällig ab, sondern folgt gewissen Prinzipien
- DBS befolgt (angegebene) Prinzipien und schützt so vor:
 - Benutzerfehlern
 - Programmfehlern

- DBSe sind für groß angelegte Anwendungen konzipiert
- In DBSen sind Techniken integriert, die mit großen Datenvolumen umgehen können

Grobe Architektur



Detailiertere Übersicht



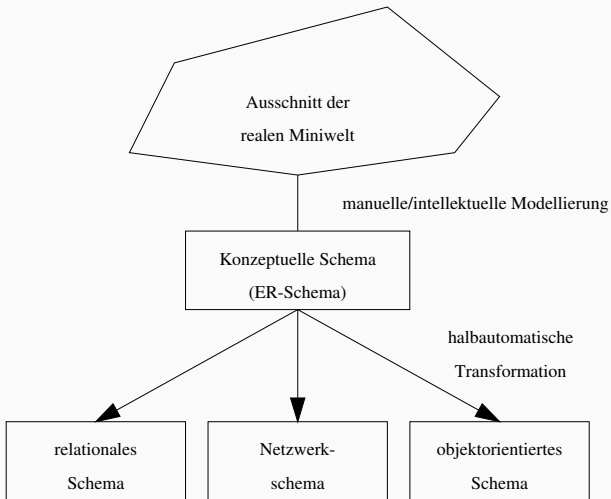
Hintergrundspeicher

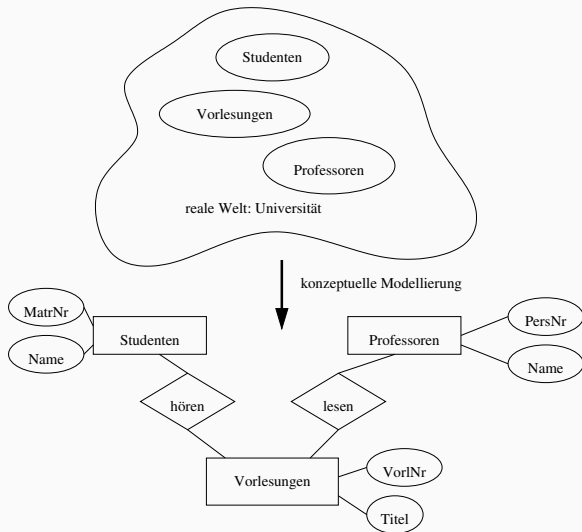
- Auch wenn DBS vieles kann, es kann nicht alles!
- Ein Benutzer muss immer noch die Anforderungen einer Anwendung ...
- ...und die Art von Daten die gespeichert werden sollen spezifizieren
- Zwei wichtige Konzepte beim Entwurf:
 - Datenmodell: legt fest, welche Konstrukte zum Beschreiben der Daten existieren
 - Schema: eine konkrete Beschreibung einer bestimmten Datensammlung (unter Verwendung eines Datenmodells)

- Konzeptuelle Modelle:
 - Entity-Relationship-Modell (ER-Modell)
 - Unified Modeling Language (UML)
- Logische Modelle¹:
 - Hierarchisches Modell
 - Netzwerkmodell
 - Relationales Modell
 - Objekt-orientiertes Modell
 - Objekt-relationales Modell
 - Semistrukturiertes Modell
 - Graphstrukturiertes Modell

¹*What Goes Around Comes Around*, Stonebraker und Hellerstein, 2005

Schritte der Modellierung





Relationales Schema

Studenten		hören	
MatrNr	Name	MatrNr	VorlNr
26120	Fichte	25403	5022
25403	Jonas	26120	5001
...

Vorlesungen	
VorlNr	Titel
5001	Grundzüge
5022	Glaube und Wissen
...	...

- Datenbanksystem (DBS) = Datenbankmanagementsystem (DBMS)
- Datenbank = Datenbasis
- Datenbank \neq Datenbanksystem
- Oft wird der Begriff “Datenbank” als Synonym für “Datenbankmanagementsystem” verwendet, dies ist aber ungenau

- Viele Anwendungen haben ähnliche Erfordernisse was die Datenverwaltung angeht
- Ein DBS bildet eine gemeinsame Basis und bietet eine Infrastruktur die diese Anwendungen unterstützt
- Für eine Übersicht über das was geboten wird siehe Folie “Gründe für den DBS-Einsatz”

- Wir versuchen zwei Seiten abzudecken:
 - Wie benutzt man ein DBMS?
 - Was passiert hinter den Kulissen?

Übersicht über Vorlesung(2)

- Inhalt:
 - Datenbankentwurf (Kapitel 2)
 - Relationales Modell (Kapitel 3)
 - SQL (Kapitel 4)
 - Relationale Entwurfstheorie (Kapitel 5)
 - Physische Datenorganisation (Kapitel 6)
 - Anfragebearbeitung (Kapitel 7)
 - Transaktionsverwaltung (Kapitel 8)