

# Einführung in Datenbanksysteme ++ Datenbanken für Bioinformatiker

H. Schweppe  
FU Berlin, SS 2005  
hs@inf.fu-berlin.de

## Infos

- **Zielgruppen**  
Studierende in den Studiengängen
  - Diplom Informatik ab 4 Semester / typisch 6. Sem.  
*Anrechenbar für Praktische und Anwendungsorientierte Informatik*
  - Bachelor Informatik, 4. Semester
  - Magister, Lehramt, Nebenfach Informatik im Hauptstudium
- **Gemeinsame Veranstaltung Informatik / Bioinformatik** bis Mitte Juni, danach spezielle Themen für Bioinformatiker/innen, z.B. Data Mining, Information Retrieval, ...

## Infos

- **Vorlesung** Di. 14 – 16.00  
Do. 14 – 16.00  
Informatik-Hörsaal
- **Team** **Heinz Schweppe**  
Manuel Scholz  
Yark Schröder  
Ralf Anske (Bioinformatik)  
Patrick Schäfer
- **Übungen**  
Ort und Zeit: siehe Netz  
**Übungsbeginn nächste Woche!**

Unbedingt in  
Übungsgruppen  
eintragen !!  
(Web-Seite der VL)

## Infos

- **Sprechstunden**
  - HS : Mi. 14-15, R 167
  - MS : Mi. 14-15, R 165
- **Email**  
{schweppe,mscholz}@inf.fu-berlin.de
- **Webseite** :  
<http://www.inf.fu-berlin.de/lehre/SS05/19517-V/>
  - Unbedingt **regelmäßig Webseite besuchen**
  - Mailingliste beachten!

## Kursbestandteile

- **Vorlesung:**
  - wie üblich
- **Übungen:**
  - wöchentliche (meist) theoretische Aufgaben
  - Nachbesprechung, keine Korrektur, Musterlösung
- **Kurztests (drei während des Semesters)**
- **Projekt**
  - Begleitend zur Vorlesung
  - Datenbank entwerfen, implementieren, nutzen
  - Thema aus Liste möglicher Projektthemen
  - Teamarbeit (max. 3 in einem Team)

## Scheinkriterien (1)

Studien-gang	Scheine	Kriterien	Benotung
Informatik (Diplom), Informatik (Bachelor)	Leistungsschein für die Veranstaltung (4+2-stündig) 8 Leistungs-punkte (credits)	<u>Für den Vorlesungsschein:</u> - Bestehen von 2 der 3 <b>Kurztests</b> - Bestehen der <b>Klausur</b> - Erfolgreiche Bearbeitung des <b>Projektes</b>	<u>Vorlesungs-schein:</u> 2/3 Klausur, 1/3 Projekt
Informatik (Lehramt), Magister, Nebenfach		- Erfolgreiche Präsentation einer <b>Aufgabenlösung</b> (im Tutorium)	

## Scheinkriterien (2)

Studiengang	Scheine	Kriterien	Benotung
Bioinformatik (Bachelor)	Leistungsschein für die Veranstaltung	<ul style="list-style-type: none"> <li>- Bestehen von 2 der 3 Kurztests</li> <li>- Bestehen der Klausur</li> <li>- Erfolgreiche Bearbeitung des Projektes</li> <li>- Erfolgreiche Präsentation einer Aufgabenlösung (im Tutorium)</li> </ul>	<p><u>Vorlesungsschein:</u> 100% Klausur</p> <p><u>Übungsschein:</u> 2/3 Projekt, 1/3 Tests (2 von 3 mit max. Punktzahl)</p>

## Termine

- Übungen
- Tests: jeweils im Abstand von 3-4 Wochen
- Klausur am Do. 21.7. 14:00 – 17:00

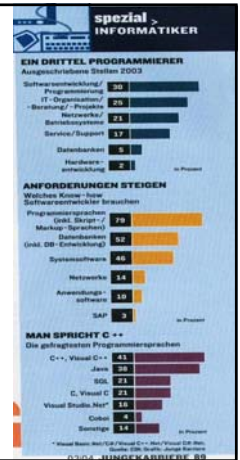
## Offizielle Anmeldung

Alle Studierenden müssen sich in die Teilnehmerliste eintragen.

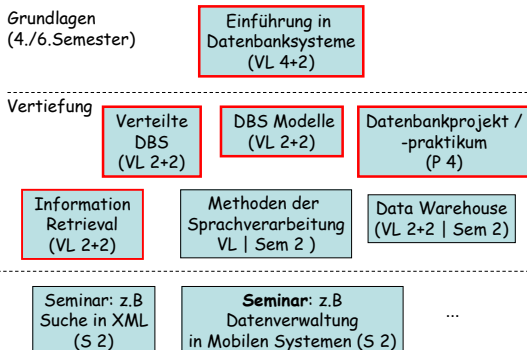
- Studierende in Bachelorstudiengängen, die **keine Prüfung** machen wollen, müssen sich **bis zum 30.4.2005 abmelden**
  - ⇒ Lehrveranstaltung gilt als nicht belegt.
- Sonst
  - ⇒ Prüfungsleistung

## Warum DBS?

- **Bachelor:**
  - Pflichtveranstaltung
- **Diplom:**
  - Eines der zentralen Gebiete in beruflicher Praxis
  - Grundlage für Vertiefung in Datenbanken
    - Anrechenbar für mündliche Abschlussprüfung in
      - Anwendungsorientierte Informatik oder
      - Praktische Informatik
- **Bioinformatik:**
  - Wahlpflichtveranstaltung: 2 aus 3 (numerische Mathematik, Datenbanken, multivar. Statistik),
  - Grundlage für algorithmische Bioinformatik



## Studieneinordnung



## Literatur

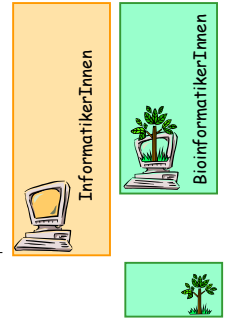
- **Bücher:**
  - Kemper/Eickler: Datenbanksysteme – Eine Einführung, Oldenbourg-Verlag, München, 3. Auflage, 1999
  - Elmasri/Navathe: Grundlagen von Datenbanksystemen Addison-Wesley / Pearson Studium, 3. Auflage, 2000
  - Garcia-Molina, H., Ullman, J., Widom, J.: *Database Systems – the Complete Book*, Prentice Hall, 2002
  - O’Neil, P., O’Neil, E.: *Database - Principles, Programming, Performance*, 2<sup>nd</sup> ed., Morgan Kaufmann, SanMatteo, 1999
  - Liste siehe Web-page (<http://www.inf.fu-berlin.de/lehre/SS02/DBS/literatur.html>)
- **Ausgewählte wissenschaftliche Literatur:**
  - Historisch oder technisch relevant
  - Im **Reader**: Link auf Web-page, Zugriff nur von 160.45.x.x Pflichtlektüre gekennzeichnet
  - Kopien der Folien:
  - Als pdf-Datei (2/6 Folien pro Seite) auf der Web-page
  - Folien in Englisch

## Begriffe

- **Datenbank**
  - Strukturierte Datenmenge, meist verwaltet von DBMS
  - Repräsentiert Schnappschuss eines Realwelt-Ausschnittes
  - Beispiele: Bankkonten und Kunden; Studenten, Kurse und Angestellte einer Universität; Genom, Proteine
- **Datenbank-System (DBS oder DBMS):** Software zum Erzeugen und Verwalten einer Datenbank
  - Definition der DB: Datentypen und Strukturen
  - Erzeugen der DB: Abspeichern der Daten
  - Operationen auf der DB: Anfragen und Updates
  - Beispiele: Oracle, DB2, SQLServer, Sybase, SAP DB, Postgres, MySQL, O2, ...

## Schwerpunkte

1. Datenmodellierung:  
systematischer Entwurf von DB  
Schwerpunkt: Relationale DB
2. Datenbanknutzung:  
Zugriff auf die Daten mit SQL  
(Structured Query Language),  
interaktiv oder mittels  
Applikationsprogrammen
3. Datenbankimplementierung



Spezielle Aspekte von  
Datenbanken für  
Bioinformatiker/innen

## Lernziele (1)

- **Datenbankentwurf**
  - relationale Datenbanken für verschiedene Anwendungen **entwerfen**
- **Datenbanknutzung**
  - SQL-Anfragen verstehen und entwerfen
  - Anwendungsprogrammierung (besonders Java) beherrschen
  - Konzepte Objektorientierter / relationaler DBS verstehen
  - **Wichtige Datenverwaltungstechniken für molekularbiologische Anwendungen kennenlernen**

## Lernziele(2)

- **Datenbanksystem-Technik**
  - Datenorganisation in DBS
  - Grundlagen der **Anfragebearbeitung**
  - Prinzipien und Techniken **transaktionaler Verarbeitung**
  - **Synchronisations-** und **Recovery-**Techniken
  - Mehrschichtarchitekturen kennen (z.B. Browser - Webserver – DB)
- **Praktische Erfahrungen** mit einem Datenbanksystem sammeln