

Datenbanksysteme II

Prof. Dr. E. Rahm

Sommersemester 2024

Universität Leipzig
Institut für Informatik



<https://dbs.uni-leipzig.de/study/semester/ss-2024/dbs2>



DBS-Module für Bachelor

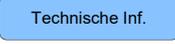
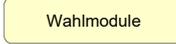
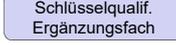
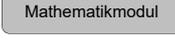
- 10-201-2211 – Datenbanksysteme 1
- 10-201-2212 – **Datenbanksysteme 2**
- 10-201-2210 – Datenbankpraktikum
- *10-201-2224 – Realisierung von Informationssystemen*
- 10-201-2010 – Bachelorseminar Informatik / Seminar modul
- Bachelorarbeit
- **DBS2** und das **DB-Praktikum** können im Bachelorstudiengang Informatik auch als **Ergänzungsfach** eingebracht werden
- Bachelorarbeit / Praktika auch für Studenten der Wirtschaftsinformatik



Bachelor Informatik mit DB-Profil

Sem.	5 LP	5 LP	5 LP	5 LP	5 LP	5 LP
1	Modellierung und Programmierung 1	Algorithmen u. Datenstrukturen 1	Technische Informatik 1	Logik	Analysis	
2	Modellierung und Programmierung 2	Algorithmen u. Datenstrukturen 2	TI 2 + HW-Praktikum	Java-Praktikum	Lineare Algebra	
3	Datenbanksysteme 1	Software-Technik	Kommunikationssysteme	Automaten und Sprachen	Diskrete Strukturen	Wahrscheinlichkeitstheorie
4	Datenbanksysteme 2	Softwaretechnik-Praktikum	Kernmodul 2	Berechenbarkeit	DB-Praktikum	
5	Realisierung v IS	Seminarmodul	Vertiefungsmodul		Ergänzungsfach	
6	Kernmodul 4	Bachelorseminar	Bachelorarbeit		Schlüsselqualifikation	

Legende:

 Praktische Inf.	 Technische Inf.	 Wahlmodule
 DBS-Modul	 Theoretische Inf.	 Schlüsselqualif. Ergänzungsfach
	 Mathematikmodul	



SS24, © Prof. Dr. E. Rahm

DBS-Lehrveranstaltungen

Logo	Name	Typ	SWS	Sem.
	Datenbanksysteme 1	Einführung	2+1	WS
	Datenbanksysteme 2	Einführung	2+1	SS
	Implementierung von DBS 1	Vertiefung	2	WS
	Implementierung von DBS 2	Vertiefung	2	SS
	Mehrrechner-DBS	Vertiefung	2	WS/SS
	Data Warehousing	Vertiefung	2	SS
	Cloud und Big Data Management	Vertiefung	2	SS
	NoSQL-DB	Vertiefung	2	SS
	Data Mining	Vertiefung	2	WS
	Data Preparation & Cleaning	Vertiefung	2+1	SS

Name	Typ	Sem.
DB-Praktikum	Praktikum	SS
Data-Warehouse-Praktikum	Praktikum	WS
Big Data Praktikum	Praktikum	SS
Problemseminar	Seminar	WS
Bachelorseminar	Seminar	SS+WS
Masterseminar	Seminar	SS+WS

LV im SS2024



SS24, © Prof. Dr. E. Rahm

DBS2 Leistungsbewertung

- u.a. für Bachelor Informatik / Wirtschaftsinformatik (4. Semester)
 - Modul Nr 10-201-2212
 - 5 Leistungspunkte (Credits)
- einheitliche Leistungsbewertung
 - **Abschlussklausur** im Juli in Präsenz (60 Minuten)
 - Zulassungsvoraussetzung: erfolgreiche **Zwischenklausur** (60 Minuten, online in Moodle)
 - erfolgreiche Klausurbewältigung erfordert Kenntnisse und Fertigkeiten aus der Vorlesung sowie der intensiven Bearbeitung von Übungen
 - Klausurtermine siehe DBS-Webseite <https://dbs.uni-leipzig.de/study/exams>



Übungsbetrieb

- DBS2 umfasst 2 SWS Vorlesung + **1 SWS Übungen**
- 4 Übungsgruppen (Di 15:15 Uhr, HS19, A+B-Woche;
Do 15:15 Uhr, Felix-Klein-HS, A+B-Woche)
- ca. 6 Übungsblätter
 - ca. alle 14 Tage im Moodle-Kurs von DBS2 , Blatt 1 am 15. April
 - **zeitversetzte Besprechung der Übungsblätter**
 - Termine siehe Webseite (Ersttermin ab 23.4.)
- Übungsleiter: Lucas Lange & Benjamin Uhrich
 - Email: lange bzw. uhrich [at] informatik.uni-leipzig.de
- zusätzliche Online-Aufgaben mit LOTS (Leipzig Online Test System)



Online-Übungen

- **LOTS (Leipzig Online Test System)**, <http://lots.uni-leipzig.de>
 - Kennung zum Beitritt in Übungsgruppe:

The screenshot shows the homepage of the Leipzig Online-Test-System. At the top, there is a blue header with the LOTS logo on the left, the text 'Leipzig Online-Test-System' in the center, and 'UNIVERSITÄT LEIPZIG' on the right, along with the faculty and department information: 'Fakultät für Mathematik und Informatik', 'Institut für Informatik', and 'Abteilung Datenbanken'. Below the header is a navigation bar with links for 'Home', 'Registrierung', and 'Impressum'. The main content area is divided into several sections: a 'Login' section with input fields for 'Username:' and 'Passwort:', a 'Gast Login' section with explanatory text and a 'Viel Spass!' message, a 'System Info' section showing '# Benutzer: 31' and a clock icon, and a 'News' section with a message: 'Bitte loggen Sie sich ein, um personalisierte News lesen zu können.' There are also buttons for 'login' and 'als Gast einloggen'.



Vorlesungsziele

- vertiefte Kenntnisse und Fertigkeiten in der Nutzung von Datenbanksystemen, insbesondere
 - DB-Programmierung
 - Web-Anbindung von Datenbanken
 - Erweiterungen des Relationenmodells und SQL, objekt-relationale DBS
 - Dokumenten-DB: XML-Datenbanken / XQuery, JSON
 - Big Data / NoSQL (Einstieg)
- Voraussetzung für Übernahme von Tätigkeiten
 - Entwicklung von datenbankgestützten Anwendungen
 - Nutzung von relationalen oder alternativen Datenbanken
 - Systemverantwortlicher für Datenbanken, insbesondere Datenbankadministrator
 - Forschung im Bereich Datenmanagement / Big Data



Vorläufiges Inhaltsverzeichnis

1. DB-Anwendungsprogrammierung: Teil 1

- Überblick
- embedded SQL, Cursor-Konzept, Dynamic SQL
- gespeicherte Prozeduren (Stored Procedures)

2. DB-Anwendungsprogrammierung: Teil 2

- Java und Datenbanken: JDBC , SQLJ
- Web-Anbindung: CGI, Java Server Pages
- PHP
- SQL Injections

3. Von relationalen zu objekt-relationalen DBS

- Beschränkungen des relationalen Datenmodells
- NF2-Ansatz
- OODBS Grundkonzepte
- OODBS vs. ORDBS

4. Objekt-relationale DBS

- Typkonstruktoren
- benutzerdefinierte Typen und Funktionen (UDTs, UDFs)
- Typhierarchien / Tabellenhierarchien (Subtypen, Subtabellen)



Vorläufiges Inhaltsverzeichnis (2)

5. Fortgeschrittene SQL-Themen

- rekursive Anfragen
- fortgeschrittene Datenanalysen (mehrdimensionale Gruppierungen, Window-Anfragen)
- temporales SQL

6. Datenanalyse mit Pandas

- Einführung / Dataframes
- Übersetzung von SQL nach Pandas
- erweiterte Konzepte

7. Semistrukturierte Daten (XML, JSON)

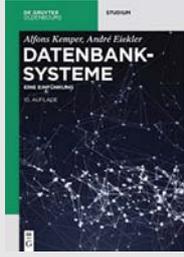
- XML-Dokumente und Schemas
- XML Anfragesprachen (Xpath, Xquery)
- XML-Datenbanken, SQL/XML
- JSON / SQL JSON

8. Big Data / NoSQL

- Big Data: Merkmale und Architekturen
- NoSQL-Datenbanken



Lehrbücher (Auswahl)

Autoren	Titel	Cover	Verlag	Auflage	Jahr
Kemper, A.; Eickler, A.	Datenbanksysteme (880 S., 50 Euro)		De Gruyter / Oldenbourg	10	2015
Saake, G.; Sattler, K.; Heuer, A.	Datenbanken: Konzepte und Sprachen (800 S., 40 Euro)		mitp	6	2018



2023

RA

Was sollte man im Informatikstudium zu Datenbanksystemen lernen?



Im Informatikstudium sollte man im Bereich Datenbanksysteme grundlegende Konzepte und Prinzipien erlernen. Hier sind einige wichtige Themen, die man in diesem Bereich lernen sollte:

1. Datenmodellierung: Die Fähigkeit, Datenmodelle zu erstellen, die das Wesen der Daten, die gespeichert werden sollen, genau widerspiegeln. Das Modell kann aus Entitäten, Attributen und Beziehungen zwischen den Entitäten bestehen.
2. Datenbanksprachen: SQL ist die am häufigsten verwendete Sprache für die Arbeit mit Datenbanken. Man sollte SQL-Kenntnisse erwerben, um Daten zu extrahieren, einzufügen, zu aktualisieren und zu löschen.
3. Datenbankdesign: Die Fähigkeit, Datenbanken zu entwerfen und zu implementieren, indem man das Datenmodell in eine physische Datenbankstruktur übersetzt. Hier sollten grundlegende Konzepte wie Normalisierung und Indexierung erlernt werden.
4. Datenintegrität: Die Fähigkeit, die Datenkonsistenz zu wahren, indem man Dateneinschränkungen und -prüfungen implementiert. Hier sollte man die Konzepte von Primär- und Fremdschlüsseln sowie anderen Einschränkungen wie NOT NULL oder UNIQUE kennen.
5. Transaktionsmanagement: Die Fähigkeit, Transaktionen zu verwalten, um Datenkonsistenz zu gewährleisten und Datenverluste zu vermeiden. Hier sollten Konzepte wie ACID (Atomicity, Consistency, Isolation, Durability) verstanden werden.
6. Performance-Optimierung: Die Fähigkeit, Datenbanken zu optimieren, um eine optimale Leistung zu erzielen. Hier sollte man Kenntnisse in den Bereichen Indexierung, Abfrageoptimierung und Datenbanktuning erwerben.
7. Datenbanksystemarchitektur: Die Fähigkeit, Datenbanken zu verstehen, indem man sich mit der Architektur von Datenbanksystemen wie Client-Server- oder verteilten Datenbanksystemen auseinandersetzt.

Diese Themen sollten ein solides Verständnis von Datenbanken im Rahmen eines Informatikstudiums vermitteln. Es ist jedoch wichtig zu beachten, dass Datenbanktechnologie ein breites Spektrum an Anwendungen und Anwendungsfällen hat, und es gibt viele weitere fortgeschrittene Themen, die man im Rahmen eines spezialisierten Kurses erlernen kann, wie z.B. Big Data-Systeme, NoSQL-Datenbanken oder Data Warehousing.



You
Was sollte man im Informatikstudium über Datenbanksysteme lernen?

ChatGPT
Im Informatikstudium gibt es verschiedene Aspekte, die man über Datenbanksysteme lernen sollte. Hier sind einige wichtige Themen und Konzepte:

1. **Datenbankgrundlagen:**
 - Verständnis der grundlegenden Datenbankkonzepte wie Datenbanken, Tabellen, Datentypen, Zeilen und Spalten.
2. **Relationale Datenbanken:**
 - Konzepte relationaler Datenbanken wie Schlüssel, Fremdschlüssel, Normalisierung und den relationalen Datenbankmodell.
3. **SQL (Structured Query Language):**
 - Gründliches Verständnis von SQL, der Sprache zur Abfrage und Manipulation von Datenbanken.
 - Erstellen, Lesen, Aktualisieren und Löschen von Daten (CRUD-Operationen).
 - Aggregatfunktionen (z.B. SUM, AVG, COUNT) und Gruppierungsoperationen (GROUP BY).
 - Join-Operationen zum Verknüpfen von Daten aus verschiedenen Tabellen.
4. **Datenbankdesign:**
 - Entwurf von Datenbanken unter Berücksichtigung von Normalisierung, Effizienz und Datenintegrität.
 - Verwendung von Tools und Techniken zum Datenbankdesign wie ER-Diagramme (Entity-Relationship-Diagramme).
5. **Datenbankverwaltungssysteme (DBMS):**
 - Kenntnisse über verschiedene DBMS wie MySQL, PostgreSQL, Oracle, SQL Server usw.
 - Installation, Konfiguration und Verwaltung von DBMS.
 - Sicherheitsaspekte wie Benutzerzugriffsrechte, Rollen und Berechtigungen.
6. **Datenbankadministration und -optimierung:**
 - Überwachung und Wartung von Datenbanken, Backups und Wiederherstellung.
 - Optimierung von Abfragen und Datenbankleistung durch Indexierung, Abfrageoptimierung und Datenbanktuning.
7. **Transaktionsverarbeitung und ACID-Prinzip:**
 - Verständnis von Transaktionen, deren Eigenschaften (ACID-Prinzip: Atomicity, Consistency, Isolation, Durability) und deren Verwaltung.
8. **NoSQL-Datenbanken:**
 - Grundlagen von NoSQL-Datenbanken wie dokumenten-, Spaltenorientierte, Key-Value- und Graphdatenbanken.
9. **Datenbankanwendungen und -entwicklung:**
 - Entwicklung von Anwendungen, die auf Datenbanken zugreifen und mit ihnen interagieren.
 - Nutzung von Programmierschnittstellen wie JDBC (Java Database Connectivity) oder ORM-Frameworks (Object-Relational Mapping) wie Hibernate.
10. **Datenbanksicherheit:**
 - Sicherheitskonzepte und -mechanismen zum Schutz von Datenbanken vor unbefugtem Zugriff, Datenverlust und -beschädigung.
11. **Big Data und Datenbanken:**
 - Grundlagen von Big-Data-Technologien und -Konzepten wie Hadoop, Spark, NoSQL-Datenbanken für die Verarbeitung großer Datenmengen.
12. **Cloud-Datenbanken:**
 - Konzepte und Technologien für Datenbanken in der Cloud, wie z.B. Amazon RDS, Microsoft Azure SQL Database, Google Cloud SQL.

Diese Themen bilden eine solide Grundlage für das Verständnis von Datenbanksystemen im Rahmen eines Informatikstudiums. Fortgeschrittene Themen können sich auf spezialisierte Bereiche wie Data Warehousing, Datenbankmigration, Big-Data-Analyse und Data Mining konzentrieren. Es ist auch wichtig, sich über aktuelle Entwicklungen in der Datenbanktechnologie auf dem Laufenden zu halten, da sich die Technologielandschaft kontinuierlich weiterentwickelt.



You
ist ein Studium Data Science empfehlenswert?

ChatGPT
Ein Studium im Bereich Data Science kann eine ausgezeichnete Wahl sein, insbesondere in einer Zeit, in der Daten eine immer wichtigere Rolle in Wirtschaft, Wissenschaft und Gesellschaft spielen. Hier sind einige Gründe, warum ein Studium in Data Science empfehlenswert sein kann:

1. **Hohe Nachfrage nach Data-Science-Fachleuten:** Unternehmen und Organisationen jeder Größe sammeln und analysieren zunehmend große Datenmengen. Es besteht eine hohe Nachfrage nach Fachleuten, die in der Lage sind, diese Daten zu verstehen, zu analysieren und daraus Erkenntnisse zu gewinnen.
2. **Breites Anwendungsspektrum:** Data Science wird in verschiedenen Branchen und Bereichen eingesetzt, darunter Wirtschaft, Gesundheitswesen, Marketing, Finanzen, Technologie, Regierung und mehr. Dadurch eröffnen sich vielfältige Karrieremöglichkeiten.
3. **Interdisziplinäre Fähigkeiten:** Data Science vereint Fähigkeiten aus den Bereichen Statistik, Informatik, Mathematik und Domänenwissen. Ein Studium in Data Science bietet die Möglichkeit, diese interdisziplinären Fähigkeiten zu entwickeln, die in vielen Arbeitsumgebungen gefragt sind.
4. **Datengestützte Entscheidungsfindung:** Durch die Analyse von Daten können fundierte Entscheidungen getroffen und Geschäftsprozesse optimiert werden. Data-Science-Fachleute spielen eine Schlüsselrolle bei der Identifizierung von Trends, Mustern und Zusammenhängen in Daten.
5. **Innovationspotenzial:** Data Science ermöglicht es, neue Erkenntnisse zu gewinnen, Prognosen zu erstellen, maschinelles Lernen und künstliche Intelligenz zu nutzen sowie datengetriebene Produkte und Dienstleistungen zu entwickeln.
6. **Flexibilität und Remote-Arbeit:** Viele Data-Science-Jobs bieten Flexibilität bei der Arbeitsplatzwahl, einschließlich Remote-Arbeitsoptionen, was für viele Fachkräfte attraktiv ist.
7. **Wachsende Technologien und Werkzeuge:** Mit dem Fortschritt in Bereichen wie Big Data, maschinelles Lernen, künstliche Intelligenz und Cloud-Computing gibt es ständig neue Technologien und Tools, die Data-Science-Fachleute nutzen können, um effektive Lösungen zu entwickeln.

Es ist jedoch wichtig zu beachten, dass ein Studium in Data Science auch anspruchsvoll sein kann und ein solides Verständnis von Mathematik, Statistik, Programmierung und Datenanalyse erfordert. Darüber hinaus sollten Studierende auch praktische Erfahrungen durch Projekte, Praktika oder Forschungsarbeiten sammeln, um ihre Fähigkeiten zu entwickeln und sich für den Arbeitsmarkt vorzubereiten. Insgesamt kann ein Studium in Data Science jedoch eine lohnende Investition in die Zukunft sein, besonders für diejenigen, die an datengetriebenen Herausforderungen und Innovationen interessiert sind.

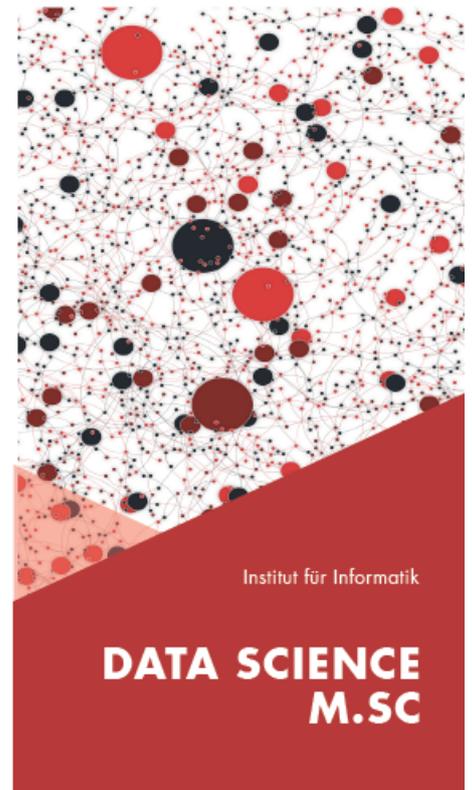


Master Data Science



UNIVERSITÄT
LEIPZIG

- seit SS2020
- Start im Winter- oder Sommersemester
 - Bachelorabschluss Informatik/Wirtsch.informatik erfüllt fachliche Zugangsvoraussetzungen
- Inhalte
 - skalierbare Datenmanagement (“Big Data“)
 - Datenanalyse / maschinelles Lernen
 - Praktika
- breites Modulangebot
 - <https://dbs.uni-leipzig.de/study/mds/modules>



<https://www.uni-leipzig.de/studium/vor-dem-studium/studienangebot/studiengang/course/show/data-science-m-sc/>

SS24, © Prof. Dr. E. Rahm

Deutsche KI-Zentren

- KI-Strategie des Bundes beinhaltet Einrichtung von 5 Zentren für Künstliche Intelligenz (neben DFKI)
 - Berlin (BIFOLD)
 - Dortmund / Bonn (ML2R)
 - **Dresden / Leipzig** (ScaDS.AI)
 - München (MCML)
 - Tübingen (tuebingen.ai)



SS24, © Prof. Dr. E. Rahm

ScaDS.AI Leipzig

- ca 100 Mitarbeitende am Standort Leipzig
 - Graduiertenschule mit ca. 50 Doktorand:innen
- 25+ Principal Investigators aus 5 Fakultäten und 5 außeruniv. Einrichtungen
 - 6 assoziierte Mitglieder (Post-Docs / Profs)
- 4-6 neue KI-Professuren
 - besetzt: Buchmann, Humboldt-Prof. S. Mukherjee
 - in Verhandlung/Ausschreibung: Multi-modal Machine Learning, Computer Vision, Advanced ML
- 3 Nachwuchsgruppenleiter



Forschungsthemen



Research Areas



Topic Areas

- Life Science & Medicine
- Environment & Earth Sciences
- Software Engineering
- Physics / Chemistry
- Engineering / Business
- Understanding Language
- Methods and Hardware for Neuro-Inspired Computing
- Graph-based Artificial Intelligence
- Knowledge Representation & Engineering
- Scalable Visual Computing
- Federated, Efficient Learning
- Math Foundations & Statistical Learning
- Big Data Analytics
- Open Data & Open Models
- Data Quality & Data Integration

Crosscutting Topics

