

# Implementierung von Datenbanksystemen 2 (IDBS2)

Sommersemester 2021

**Prof. Dr. Erhard Rahm**

Universität Leipzig

Institut für Informatik

<http://dbs.uni-leipzig.de/stud/2021ss/idbs2>



## DBS-Module 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
- Bachelorarbeit



# DBS-Module Master

## ■ Master Data Science

- 10-INF-DS01 – Skalierbare Datenbanktechnologien 1 (Pflichtmodul, 10 LP)
- **10-INF-DS101 – Skalierbare Datenbanktechnologien 2 (5 LP)**
- 10-INF-DS102 – Big Data Praktikum (5 LP)
- 10-INF-DS103 – Praktikum Data Warehousing und Data Mining (5 LP)
- **10-INF-DS301 – Aktuelle Trends in Data Science (5 LP)**
- 10-INF-DS02 – Masterseminar Data Science (5 LP)

## ■ Master Informatik

- 10-202-2215 – Moderne Datenbanktechnologien (kleines Modul)
- 10-202-2216 – *Moderne Datenbanktechnologien (ausgelaufen)*
- **10-202-2213 – Anwendungsbezogene Datenbankkonzepte (kleines Modul)**
- 10-202-2214 – *Anwendungsbezogene Datenbankkonzepte (ausgelaufen)*
- 10-202-2011 – Masterseminar Informatik
  - Masterarbeit



## Mapping Module – Lehrveranstaltungen SS21

- Bachelor-Modul *Realisierung von Informationssystemen*, Master-Module *Skalierbare Datenbanktechnologien 2* und *Moderne Datenbanktechnologie (5 LP)*:
  - **IDBS1**
  - NoSQL
- Aktuelle Trends in Data Science
  - VL/Praktikum *Data Wrangling* (Gastprofessor Peter Christen)
  - beschränkt auf Master Data Science
- Bachelorseminar / Masterseminar
  - Vortrag über laufende Bachelor/Masterarbeit im Rahmen des DB-Oberseminars



# Masterstudium Data Science



- neuer Studiengang seit SS20
- Hauptinhalte
  - skalierbares Datenmanagement („Big Data“), mind. 20 LP
  - Datenanalyse / Machine Learning, mind. 20 LP
  - Ergänzungs-/Anwendungsmodule
  - viele Praktika möglich

**1. Semester                      2. Semester                      3. Semester                      4. Semester**

<b>Skalierbare Datenbanktechnologien 1</b>		Skalierbares Datenmanagement	Skalierbares Datenmanagement	Vertiefung	Mastersem. Data Science (5 LP)
Datenanalyse		Datenanalyse		Vertiefung	Masterarbeit (25 LP)
Ergänzung	Ergänzung	Ergänzung	Anwendung/ Ergänzung	Anwendung/ Ergänzung	



## Beispielbelegungen Data Science

Wintersemester		Sommersemester		Wintersemester	4. Semester
Skalierbare Datenbanktechnologien 1		SDBT2	Big Data Praktikum	Fortgeschr. Methoden Information Retrieval	Masterseminar (5 LP)
Künstl. neuronale Netze u. maschin. Lernen		Multivariate Statistik und Data Mining		Wissens- und Content Management	Masterarbeit (25 LP)
IT-Sicherheit	Prakt. Data Wareh./ Mining	Textdatenbanken	Aktuelle Trends in DS	Verfahren und Anwendungen in den Digital Humanities	

Skalierbare Datenbanktechnologien 1		SDBT2	Big Data Praktikum	Text Mining	Masterseminar (5 LP)
Statistisches Lernen		Multivariate Statistik und Data Mining		Künstl. neuronale Netze u. maschin. Lernen	Masterarbeit (25 LP)
Sequenzanalyse und Genomik		Grundl. komplexer Systeme	Aktuelle Trends in DS	Visualisierung	



# Leistungsbewertung

## ■ Prüfungsklausur 60 Minuten

- vstl. Juli 2021
- Präsenzklausur (benotet) oder Online-Klausur (unbenotet)
- überprüft konzeptionelles Wissen + Anwendungsfälle

## ■ Klausurerfolg durch

- intensives Vorlesungsstudium (Live/Video, Skript)
- Online-Übungen (LOTS)
- Literatur



## Lernziele der Vorlesung IDBS

- fundierte Kenntnisse der Funktionsweise von DBS und zu *Big Data Engineering*
- Implementierungstechniken u.a. zur
  - Sicherstellung einer hohen Performanz/Skalierbarkeit auf großen Datenmengen
  - Datensicherheit
- IDBS1: Verfahren zur Externspeicher-Nutzung, Verwaltung von Pufferspeichern, Indexstrukturen, Anfrageoptimierung ...
- IDBS2: Verfahren zur Transaktionsverwaltung: Synchronisation (Concurrency Control), Logging/Archivierung, Recovery
- tiefgehende Kenntnisse wichtig für DB-Administration sowie generell für anspruchsvolle DB-Nutzung



# Vorläufige Vorlesungsübersicht

## 1. Einführung: Transaktionsverwaltung, Integritätskontrolle

## 2. Synchronisation: Grundlagen, Sperrverfahren

- Mehrbenutzer-Anomalien
- Serialisierbarkeit
- Sperrverfahren: 2PL, Hierarchische Sperrverfahren
- Konsistenzstufen
- Deadlock-Behandlung

## 3. Synchronisation: Weitere Verfahren, Leistungsbewertung

- Optimistische Verfahren
- Zeitstempel- und Mehrversionen-Verfahren
- Spezialverfahren für B\*-Bäume und „High Traffic“-Elemente
- Leistungsanalyse und Lastkontrolle



## Vorlesungsübersicht (2)

## 4. Logging und Recovery: Grundlagen

- Begriffe und Annahmen, Fehlermodell
- Logging-Verfahren
- Klassifikation von Recovery-Strategien

## 5. Crash- und Medien-Recovery

- Crash-Recovery
- Platten-Recovery

## 6. Transaktionskonzept: Weiterentwicklungen

- Geschachtelte Transaktionen
- Transaktionsketten (Sagas)



# Literatur

- Härder, T., Rahm, E.: *Datenbanksysteme - Konzepte und Techniken der Implementierung*. Springer-Verlag, 2. Auflage 2001 (Kap. 1 und 13 online)



- weitere Lehrbücher

- Saake, G., Heuer, A.: *Datenbanken: Implementierungstechniken*, MITP-Verlag, 2019
- Garcia-Molina, H., Ullman, J.D., Widom, J.: *Database System Implementation*. Prentice Hall, 2000
- 

- Google Scholar

- <http://scholar.google.com/> (Volltexte von Publikationen, Zitierungsangaben)



# Online-Übungen

- LOTS (Leipzig Online Test System), <http://lots.uni-leipzig.de>

**LOTS** Leipzig Online-Test-System

UNIVERSITÄT LEIPZIG  
Fakultät für Mathematik und Informatik  
Institut für Informatik  
Abteilung Datenbanken

Home Registrierung Impressum

**Login**

Username:

Passwort:

[Passwort vergessen?](#)

**Gast Login**

Sie können sich als Gast einloggen, um LOTS ohne vorherige Anmeldung zu testen. Der Gast Account ist auf 30 min Benutzung und im Funktionsumfang beschränkt. Bei weiterem Interesse sollten Sie sich als Gast [registrieren](#).

Viel Spass!

Ihr LOTS Team

**System Info**

# Benutzer: 31

05.02.2007 12:15:21

**News**

Bitte loggen Sie sich ein, um personalisierte News lesen zu können.



# 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)



## Künstliche Intelligenz für Leipzig

Mit einer langfristigen Finanzierung von Bund und Land entsteht ein neuer Forschungsstandort

Von Leven Wortmann

Am Institut für Informatik der Uni Leipzig gibt es schon seit langem einen Schwerpunkt: Big Data. Mit einer Anfang November beschlossenen Finanzierung soll ein neuer Schwerpunkt hinzukommen: Künstliche Intelligenz (KI). An der Universität Leipzig, in Kooperation mit der TU Dresden, wird das ScaDS.AI, eines von fünf ausgewählten KI-Kompetenzzentren in Deutschland, langfristig finanziert werden.

### Thema KI wird Teil der Hochschulbildung

Im Rahmen der Förderung soll nicht nur das KI-Kompetenzzentrum aufgebaut, sondern auch das Thema KI in die Hochschulbildung aufgenommen werden, wofür der Bund

und das Land Sachsen im Jahr zwischen 7,5 und 12 Millionen Euro zur Verfügung stellen. Dafür werden an der Uni Leipzig vier neue Professuren am Institut für Informatik eingerichtet. Ebenso wurde dieses Jahr der neue Masterstudiengang „Data Science“ eingeführt, in dem die Studierenden lernen sollen, wie große Datenmengen verarbeitet und analysiert werden können.

Professor Erhard Rahm, Ko-Direktor des KI-Zentrums und Studiengangverantwortlicher vom Institut für Informatik der Universität Leipzig, erklärt: „Die neuen Professuren werden vor allem für den



Professor Erhard Rahm

Masterstudiengang Data Science gebraucht, damit man hier breit aufgestellt ist und mehr Kapazitäten für Studierende schaffen kann.“

### Leipzig soll für KI-Forschung attraktiver werden

Mit der praxisnahen Ausrichtung des Kompetenzzentrums und des Studiengangs ist einerseits eine Zusammenarbeit mit anderen Instituten der Universität denkbar, um auch andere Studiengänge mit den Themen Big Data und KI zu erweitern, aber andererseits auch Zusammenarbeit mit Unternehmen aus der Region: „Damit können wir erreichen, dass Forschungsergebnisse von unserem KI-Zentrum in die Unternehmen kommen und damit die Unternehmen im Wettbewerb gestärkt werden – also man

mit neuesten Methoden eine hohe Konkurrenzfähigkeit etabliert“, erklärt Rahm.

Um den Standort Leipzig für die KI-Forschung noch attraktiver zu machen, ist auch ein neues Rechenzentrum geplant, um die nötige Infrastruktur für die rechenintensive Forschung und Entwicklung von KIs zu gewährleisten. Die Kosten für dieses Vorhaben sollen über das Strukturstärkungsgesetz Kohleregion getragen werden. Die Projektbeteiligten hoffen, damit eine zukünftige Infrastruktur für viele Arbeitsplätze schaffen zu können. Rahm ist zuversichtlich: „Im Bereich der KI sind wir dabei, mit der Berufung der neuen Professuren an der Spitze mitzumachen, und wir haben hier die Chance, sehr guten Nachwuchs an Studierenden zu gewinnen.“

# KI-Zentrum ScaDS.AI Dresden/Leipzig

- Center for Scalable Data AnalyticS and Artificial Intelligence [scads.ai](https://scads.ai)
  - eines der vom Bund geförderten nationalen KI-Zentren
  - Erweiterung des seit 2014 bestehenden Big-Data-Kompetenzzentren
  - Ausbau u.a. mit 4 neuen KI-Professuren sowie einer Graduiertenschule
- Forschungsthemen:

