

# Datenbanksysteme I

Prof. Dr. E. Rahm

Wintersemester 2016/2017



Universität Leipzig  
Institut für Informatik

<http://dbs.uni-leipzig.de>



## BACHELOR Informatik

Algorithmen und Datenstrukturen 1 (5 LP)	Algorithmen und Datenstrukturen 2 (5 LP)	<b>Datenbanksysteme 1 (5 LP)</b>	Kernmodul 1 (5 LP)	Kernmodul 3 (5 LP)	Kernmodul 5 (5 LP)
Modellierung und Programmierung 1 (5 LP)	Modellierung und Programmierung 2 (5 LP)	Softwaretechnik (5 LP)	SWT-Praktikum (5 LP)	Seminarmodul (5 LP)	Bachelorseminar (5 LP)
Technische Informatik 1 (5 LP)	Technische Informatik 2 inkl. Hardwarepraktikum (5 LP)	Kommunikationssysteme (5 LP)	Kernmodul 2 (5 LP)	Vertiefungsmodul (10 LP)	Bachelorarbeit (10 LP)
Logik (5 LP)	Java-Praktikum (5 LP)	Automaten und Sprachen (5 LP)	Berechenbarkeit (5 LP)		
Analysis 1 (10 LP)	Lineare Algebra 1 (10 LP)	Diskrete Strukturen (5 LP)	Ergänzungsfach 1 (10 LP)	Ergänzungsfach 2 (10 LP)	Fakultätsübergreifende Schlüsselqualifikation (10 LP)
		Wahrscheinlichkeitstheorie (5 LP)			
1. Semester	2. Semester	3. Semester	4. Semester	5. Semester	6. Semester

Legende:

Inf-Pflichtmodul

MI-Modul

Mathematikmodul

Schlüsselqualif. Ergänzungsfach



# Bachelor of Science (B.Sc.) **Wirtschaftsinformatik**

	10 Leistungspunkte		10 Leistungspunkte		10 Leistungspunkte	
1	Wirtschaftswissenschaften	Rechnungswesen	Wirtschaftsinformatik I	Strukturierte Programmierung	Recht für Wirtschaftswissenschaftler	Mathematik für Wirtschaftswissenschaftler
2	Mikroökonomik		Web-Techniken I	Objektorientierte und generische Programmierung		
3	Statistik und Wahrscheinlichkeitsrechnung		Wirtschaftsinformatik II	Software-technik	Datenbanksysteme I	Algorithmen und Datenstrukturen I
4	Externes und internes Rechnungswesen			Entwicklung verteilter Anwendungen	Datenbanksysteme II	Algorithmen und Datenstrukturen II
5	Praktikum		<i>Fakultätsübergreifende Schlüsselqualifikation</i> oder Praktikum		<i>Fachnahe Schlüsselqualifikationen</i> oder Praktikum	
6	Wahlpflichtmodul		Web Techniken II	Web Science	Bachelorarbeit	

<http://wi.wifa.uni-leipzig.de/iwi/iwi/studium/bachelor.html>



## 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 / Seminarmodul
- Bachelorarbeit



# BACHELOR DBS-Profil

Algorithmen und Datenstrukturen 1 (5 LP)	Algorithmen und Datenstrukturen 2 (5 LP)	<b>DBS 1 (5 LP)</b>	<b>DBS 2 (5 LP)</b>	Realisierung von IS (5 LP)	Kernmodul 5 (5 LP)
Modellierung und Programmierung 1 (5 LP)	Modellierung und Programmierung 2 (5 LP)	Softwaretechnik (5 LP)	SWT-Praktikum (5 LP)	Seminarmodul (5 LP)	Bachelorseminar (5 LP)
Technische Informatik 1 (5 LP)	Technische Informatik 2 inkl. Hardwarepraktikum (5 LP)	Kommunikationssysteme (5 LP)	Kernmodul (5 LP)	Vertiefungsmodul (10 LP)	<b>Bachelorarbeit (10 LP)</b>
Logik (5 LP)	Java-Praktikum (5 LP)	Automaten und Sprachen (5 LP)	Berechenbarkeit (5 LP)		
Analysis 1 (10 LP)	Lineare Algebra 1 (10 LP)	Diskrete Strukturen (5 LP)	<b>DB-Praktikum (5 LP)</b>	Ergänzungsfach (10 LP)	Fakultätsübergreifende Schlüsselqualifikation (10 LP)
		Wahrscheinlichkeitstheorie (5 LP)	Ergänzungsfach (5LP)		
1. Semester	2. Semester	3. Semester	4. Semester	5. Semester	6. Semester

Legende:

Inf-Pflichtmodul

MI-Modul

**DBS-Modul**

Mathematikmodul

Schlüsselqualif. Ergänzungsfach



## 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
	Data Warehousing	Vertiefung	2	SS
	Datenintegration	Vertiefung	2	SS
	Bio Data Management	Vertiefung	2	WS
	Ontologie-Management	Vertiefung	2	SS
	Cloud Data Management	Vertiefung	2	WS
	NoSQL-DB	Vertiefung	2	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



# DBS1 Leistungsbewertung

- Geregelt in Modulbeschreibung Nr 10-201-2211
  - 5 Leistungspunkte für DBS1
- Anmeldung über Almaweb
- einheitliche Verfahrensweise für Bachelor Informatik / Wirtschaftsinformatik / Sonstige
  - benotete Prüfungsleistung über **Abschlussklausur** (60 Minuten)
  - Zulassungsvoraussetzung: erfolgreiche **Zwischenklausur** (60 Minuten)
  - keine mündliche Prüfung
- Klausurerfolg erfordert
  - Wissen über die Vorlesungsinhalte
  - Kenntnisse und Fertigkeiten zur Anwendung des Wissens
  - > Vorlesungsteilnahme, Vorlesungsnachbearbeitung anhand Folien/Mitschriften sowie Begleitliteratur, intensive Bearbeitung der Übungen

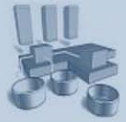


## Übungsbetrieb

- DBS1 umfaßt 2 SWS Vorlesung + 1 SWS Übungen
- **Übungsblätter**
  - ca. alle 2 Wochen auf [dbs.uni-leipzig.de](http://dbs.uni-leipzig.de) (ab 17.10.)
  - Besprechung jeweils ab **1 Woche später** in den Übungen
  - **Übungsleiter:** Victor Christen / Martin Junghanns
  - Bearbeitung erforderlich, jedoch **keine Abgabe von Lösungen**
  - zusätzlich Online-Aufgaben mit **LOTS** (Leipzig Online Test System)
- **Übungsgruppen** (ca. 2-wöchentliche Termine, genauer Zeitplan im Web)

Gruppen (A/B-Wochen)	Ort
Montags, A-Woche, 17:15 Uhr, ab 24.10.	HS 19
Dienstags, A-Woche, 11:15 Uhr, ab 25.10.	Felix-Klein-HS
Montags, B-Woche, 17:15 Uhr, <i>erst ab Blatt2: 15.11.</i>	HS 19
Dienstags, B-Woche, 11:15 Uhr, ab 3.11.	Felix-Klein-HS



### Inhalte

- ▶ Mitarbeiter
- ▶ Forschung
- ▼ Studium
  - ▶ Klausurtermine
  - ▶ SS 2012
  - ▶ SS 2013
  - ▶ SS 2014
  - ▶ SS 2015
  - ▶ SS 2016
  - ▶ WS 2011/12
  - ▶ WS 2012/13
  - ▶ WS 2013/14
  - ▶ WS 2014/15
  - ▶ WS 2015/16
  - ▶ WS 2016/17

### Studium

Bachelor/Master-Module zu Datenbanken mit Angebot/Zuordnungen im WS 2016/17

WS16/17: Datenbanksysteme 1, Implementierung von DBS 1, Cloud Data Management, Data Warehouse Praktikum, Seminar: Big Data Streaming, Oberseminar

Abschlussarbeiten (Bachelor/Master) – Anfrage, offene Themen, fertiggestellte Arbeiten

Klausurtermine, Hinweise und Ergebnisse

Gezieltere Suche/Navigation zu Klausurergebnissen über die Kategoriensuche

LOTS , Hinweise zu LOTS

Bisherige Oberseminare / Zingst-Seminare


Bisherige Problemseminare

Hinweise zu schriftlichen Seminararbeiten



## Online-Übungen

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



# Leipzig Online-Test-System

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

Home | **Registrierung** | Impressum

### Login

Username:

Password:

[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.





### Tutorial

- 1 Einleitung
- 2 Datenbankmodellierung und Relationenmodell
- 3 SQL
- 4 Einfache SQL-Anfragen
- 5 Verbund-Anfragen
  - 5.1 Vorbemerkungen
  - 5.2 Join-Anfragen
  - 5.3 Verbundausdrücke
- 6 Unterabfragen
- 7 Aggregatfunktionen
- 8 Partitionierung in Gruppen und Auswahl
- 9 Suchbedingungen
- 10 Mengentheoretische Operationen

[Zurück](#) [Weiter](#) [Hoch](#) | [zurück zum SQL-Anfrageformular](#)

## 5.2 Join-Anfragen

Wenn Attributwerte aus mehreren Relationen abgefragt werden sollen und diese Attribute in einem inneren Zusammenhang stehen oder gestellt werden sollen, sind Join-Anfragen zu entwickeln. In der ersten Möglichkeit, dem Verbund über eine gemeinsame Spalte, sind zunächst in der FROM-Klausel alle beteiligten Relationen anzugeben. Die WHERE-Klausel enthält Selektionsbedingungen und zusätzlich die Join-Bedingung. Die Join-Bedingung gibt an, über welche Attribute (gemeinsame Spalte) die Beziehung zwischen den Relationen hergestellt ist.

Folgendes einfache Anfrage-Beispiel dient der Auflistung aller in Berliner Verlagen erschienenen Bücher:

**Beispiel:**

BNF: [select-ausdruck](#)

[diese Anfrage ausführen](#)

```
SELECT b.titel
FROM verlag AS v, buch AS b
WHERE v.ort = 'Berlin' AND v.verlagsid = b.verlagsid
```



## Vorlesungsziele

### ■ Kenntnisse, Fähigkeiten und Fertigkeiten

- zur Nutzung von Informations- und Datenmodellen, insbesondere
  - Entity/Relationship-Modell und Erweiterungen, UML-Klassendiagramme
  - Relationenmodell und SQL
  - *Weitere Modelle (OO, XML, NoSQL) -> Vorlesung DBS2*
- zur Modellierung von anwendungsbezogenen Realitätsausschnitten (Miniwelten, Diskursbereiche)
- im Entwerfen, Aufbauen und Warten von Datenbanken
- *zur Programmierung von DB-Anwendungen -> Vorl. DBS2; DB-Praktikum*



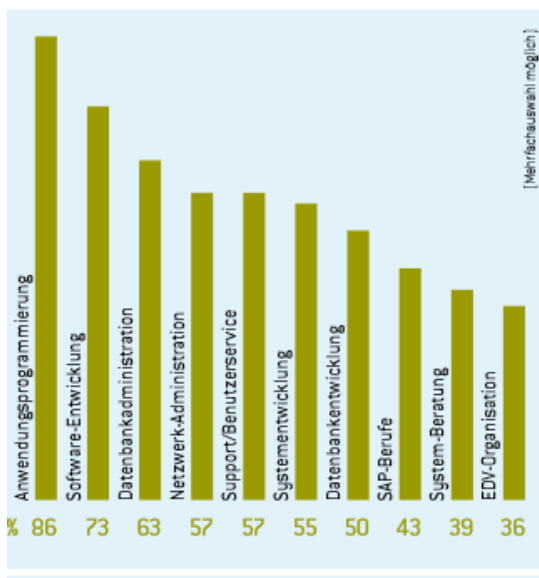


# Warum ist die Vorlesung relevant?

- DBS-Grundkenntnisse sind in fast allen IT-Berufen erforderlich und werden erwartet
- Voraussetzung für Übernahme von Tätigkeiten:
  - Entwicklung von datenbankgestützten Anwendungen
  - Nutzung von Datenbanken unter Verwendung von (interaktiven) Datenbanksprachen
  - Systemverantwortlicher für Datenbanksysteme, insbesondere Datenbank-, Datensicherungs-, Anwendungs- und Unternehmensadministrator



## Umfrage: Wo setzen Unternehmen Informatiker ein?



Informatiker starten bei den Unternehmen am häufigsten in der **Anwendungsprogrammierung** (86 Prozent) und der **Software-Entwicklung** (73 Prozent). Ebenfalls wichtige Einsatzfelder für IT-Spezialisten in den befragten Unternehmen sind die Datenbank- und Netzwerk-Administration sowie der Bereich Support und Benutzerservice. Der Bereich Multimedia spielt als Einsatzfeld für Informatiker eine deutlich geringere Rolle.

stufenziel  
**JOB**  
**TRENDS**  
**2015**



# Vorläufiges Inhaltsverzeichnis DBS1

## 1. Einführung / Grundlagen von DBS

- DBS vs. Dateisysteme
- Eigenschaften von DBS
- Datenmodelle
- Transaktionskonzept (ACID)
- Aufbau von DBS
- Einsatzformen

## 2. Informationsmodellierung: Entity-Relationship-Modell / UML

- Stufen des DB-Entwurfs
- Grundkonzepte des ER-Modells
- Beziehungstypen, Kardinalitätsrestriktionen
- Generalisierung und Aggregation
- UML (Klassendiagramme)

## 3. Grundlagen des Relationalen Datenmodells

- Relationale Invarianten
- Relationenalgebra



# Vorläufiges Inhaltsverzeichnis DBS1 (2)

## 4. Einführung in die Standardsprache SQL

- Befehlsübersicht
- Anfragemöglichkeiten (SELECT)
- SQL-Änderungsoperationen (Insert, Update, Delete)
- Vergleich SQL - Relationenalgebra

## 5. Normalisierung relationaler Schemas

- Funktionale Abhängigkeiten
- Normalformenlehre: 1NF, 2NF, 3NF, BCNF

## 6. Datendefinition in SQL

- Datendefinitionen, Schemaevolution
- Sichtkonzept (Views)

## 7. Datenkontrolle

- Integritätsbedingungen und Trigger
- Zugriffskontrolle
- Synchronisation und Recovery





# Lehrbücher (Auswahl)

Autoren	Titel	Cover	Verlag	Auflage	Jahr
Kemper, A.; Eickler, A.	Datenbanksysteme (880 S., 50 Euro)		Oldenbourg	10	2015
Saake, G.; Sattler, K.; Heuer, A.	Datenbanken: Konzepte und Sprachen		mitp	5	2013



## Lehrstuhl Datenbanksysteme

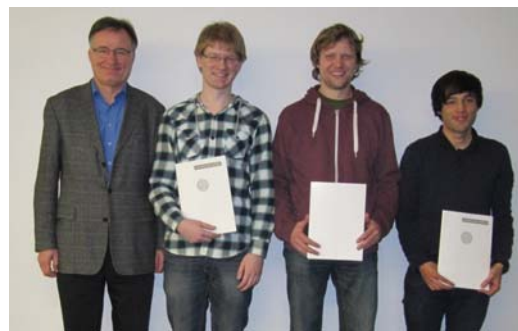
- seit 1994 am Institut für Informatik
- umfangreiches Lehrangebot
  - Vorlesungen, Praktika, Seminare
  - Online-Übungssystem LOTS
  - Eigene Lehrbücher: MRDBS, IDBS, ...



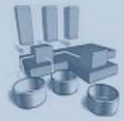
# Oberseminare an Uni-Außenstelle Zingst – seit 2001



# Auszeichnung von Top-Student(inn)en - seit 2008







### Inhalte

- ▶ Mitarbeiter
- ▼ **Forschung**
  - Publikationen
  - ▶ Projekte
  - Prototypes
  - Jahresberichte
  - Kooperationen
  - ▶ Promotionen
  - Colloquia
  - ▶ Conferences
- ▶ Studium
- ▶ Service

### Neue Publikationen

- **Holistic Entity Clustering for Linked Data**

## Forschung

[Publications \(2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, ...\)](#)

[Big Data Center ScaDS Dresden/Leipzig](#)

[Graph-based data analysis \(GRADOOP, BIIIG\)](#)

[Entity Matching for Big Data \(Dedoop\)](#)

[LOD Link Discovery](#)

[Evolution of Semantic Annotations \(ELISA\)](#)

[Schema and Ontology Matching \(COMA++, GOMMA\), Ontology Merging \(ATOM\)](#)

[Semantic ontology matching \(STROMA mapping enrichment, SemRep repository\)](#)

[Evolution of ontologies and mappings \(Schema Evolution bibliography\)](#)

[Web Data Integration \(WDI\) Lab](#)

[Object Matching / Entity Resolution](#)



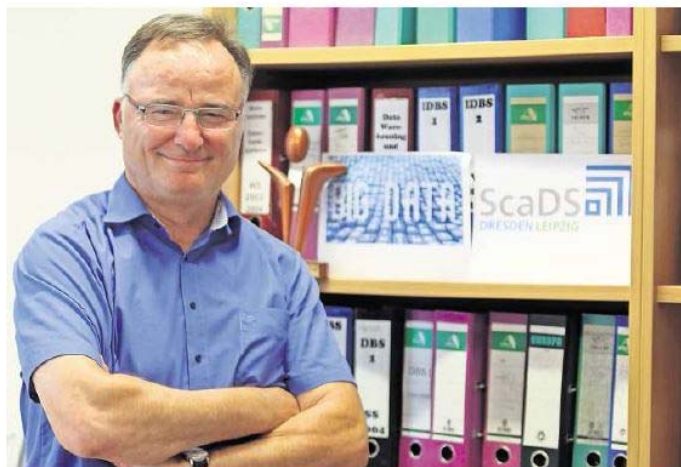
# Galaxy und Sirius: Big-Data-Zentrum an der Uni mit mehr Computerpower

1,3 Millionen Euro für neue Server plus Zubehör / Professor Rahm: „Viel Performance durch Parallelverarbeitung“

VON MARIO BECK

Das technische Rückgrat des von der Leipziger Universität zusammen mit der TU Dresden betriebenen Big-Data-Kompetenzzentrums (Scads) wird weiter gestärkt. Zur bisher installierten Hardware kommt jetzt ein als Galaxy-Cluster bezeichneter Komplex mit 90 Servern. 60 davon sind an der Alma mater schon installiert und gehen nach und nach ans Netz. 30 Server sind für Dresden bestimmt, werden aber auch vom hiesigen Uni-Rechenzentrum (URZ) aus gesteuert. Über exakt 1080 Prozessoren verfügt das Galaxy-Gesamtsystem, die Festspeicherkapazität erreicht 2,16 Petabyte – ein Petabyte entspricht einer Million Gigabyte. „Das ist schon eine richtige Hausnummer“, sagt Erhard Rahm, der an der Uni die Professur für Datenbanken innehat und am Scads einer der zwei Koordinatoren ist.

Für dieses Jahr wird am URZ noch die Lieferung des Sirius-Clusters mit zwei Memory-Servern erwartet, die alleine in ihren Hauptspeichern je sechs Terabyte Daten bunkern können. Solche Mengen fallen beispielsweise bei der Erforschung von Genomen oder Krankheitsmarkern an. Beides sind Felder, in die sich die im Scads vereinten Experten unter anderem einbringen. Vor zwei Jahren war das Big-Data-Zentrum als Gemeinschaftsprojekt



Professor Erhard Rahm leitet seitens der Leipziger Uni das gemeinsam mit der TU Dresden betriebene Big-Data-Kompetenzzentrum.

Foto: André Kempner

der Unis in Leipzig und Dresden mit einem Fördervolumen von 5,6 Millionen Euro an den Start gegangen. Galaxy, Sirius und Zubehör schlagen jetzt mit rund 1,3 Millionen Euro zu Buche.

Zum Scads-Konsortium gehören noch das Leibniz-Institut für ökologische Raumentwicklung und das Max-Planck-Institut

für Zellbiologie und Genetik, die in Dresden ansässig sind. Hinzu gesellen sich eine Vielzahl assoziierter Partner aus Wissenschaft und Wirtschaft, darunter die Leipziger Data Virtuality GmbH und das Logistik-Netzwerk Leipzig-Halle.

Die neuen Galaxy-Server, die sowohl für die Grundlagen- als auch die ange-

wandte Forschung eingespannt werden, sind in einer Shared-Nothing-Architektur konfiguriert. Rahm: „Im Verbund liegt ihre Leistungsstärke. Die Performance kommt aus der Parallelverarbeitung.“ Aufgesetzt ist dabei meist eine Software-Technik namens Hadoop, auf deren Basis Rahm mit seinem Team ein Vorhaben mit der Kennung Gradoop vorantreibt. Es zielt auf die schnelle Analyse vernetzter Daten, die – wie bei Facebook – von hoher Beziehungsdynamik geprägt sind. Mit den Themen Datenintegration, Wissensextraktion und Visualisierung hat das Scads drei feste Standbeine und strahlt stark aus: In die Lebens- und Umweltwissenschaften, die Werkstoffkunde oder die digitalen Geisteswissenschaften. Der gute Ruf des Zentrums hat sich schon international herumgesprochen und sorgt für Echo.

Zur jüngsten Big-Data-Sommerschule an der Uni war der Andrang aus aller Welt groß. Rund 120 Teilnehmer kamen – und Rahm ließ wegen seiner Verpflichtungen bei der Summerschool eine Einladung in die USA sausen. Er war zum 25. Geburtstag des Microsoft-Research-Centers (MRC) gebeten worden, um dort den Outstanding Collaborator Award entgegenzunehmen. 32 Auserwählte aus aller Herren Länder, die in der Vergangenheit als Gastforscher am MRC gewirkt und dabei die nachhaltigsten Spuren hinterlassen hatten, bekamen die Ehrung.

