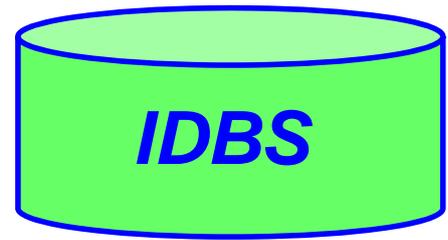


# Implementierung von Datenbanksystemen 1 (IDBS1)



Wintersemester 2012/13  
**Prof. Dr. Erhard Rahm**  
Universität Leipzig  
Institut für Informatik  
<http://dbs.uni-leipzig.de>



## DBS-Module

### ■ Master-Studium

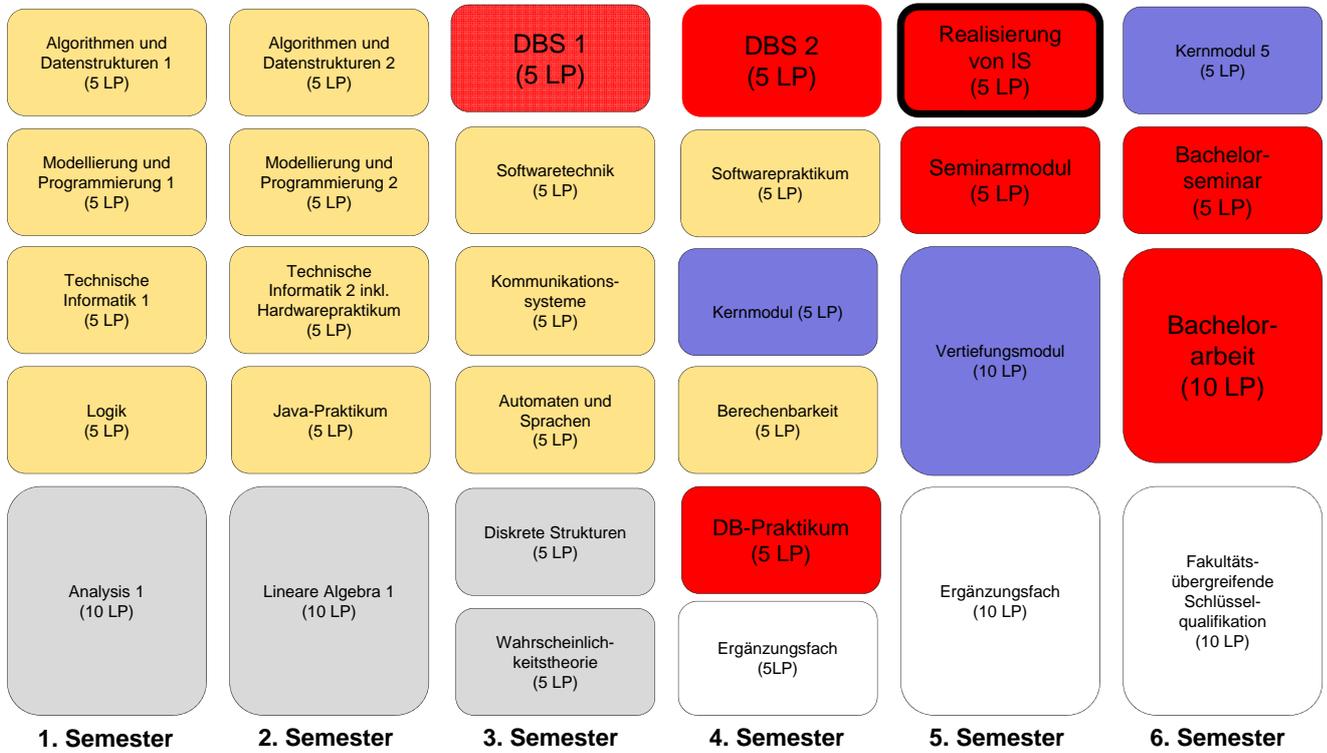
- 10-202-2215 – **Moderne Datenbanktechnologien (Kleines Modul)**
- 10-202-2216 – **Moderne Datenbanktechnologien (Großes Modul)**
- 10-202-2213 – **Anwendungsbezogene Datenbankkonzepte (Kleines Modul)**
- 10-202-2214 – **Anwendungsbezogene Datenbankkonzepte (Großes Modul)**
- 10-202-2011 – **Masterseminar Informatik / Seminar modul**

### ■ Bachelor-Studium

- 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**



# BACHELOR DBS-Profil



Legende:

Inf-Pflichtmodul

MI-Modul

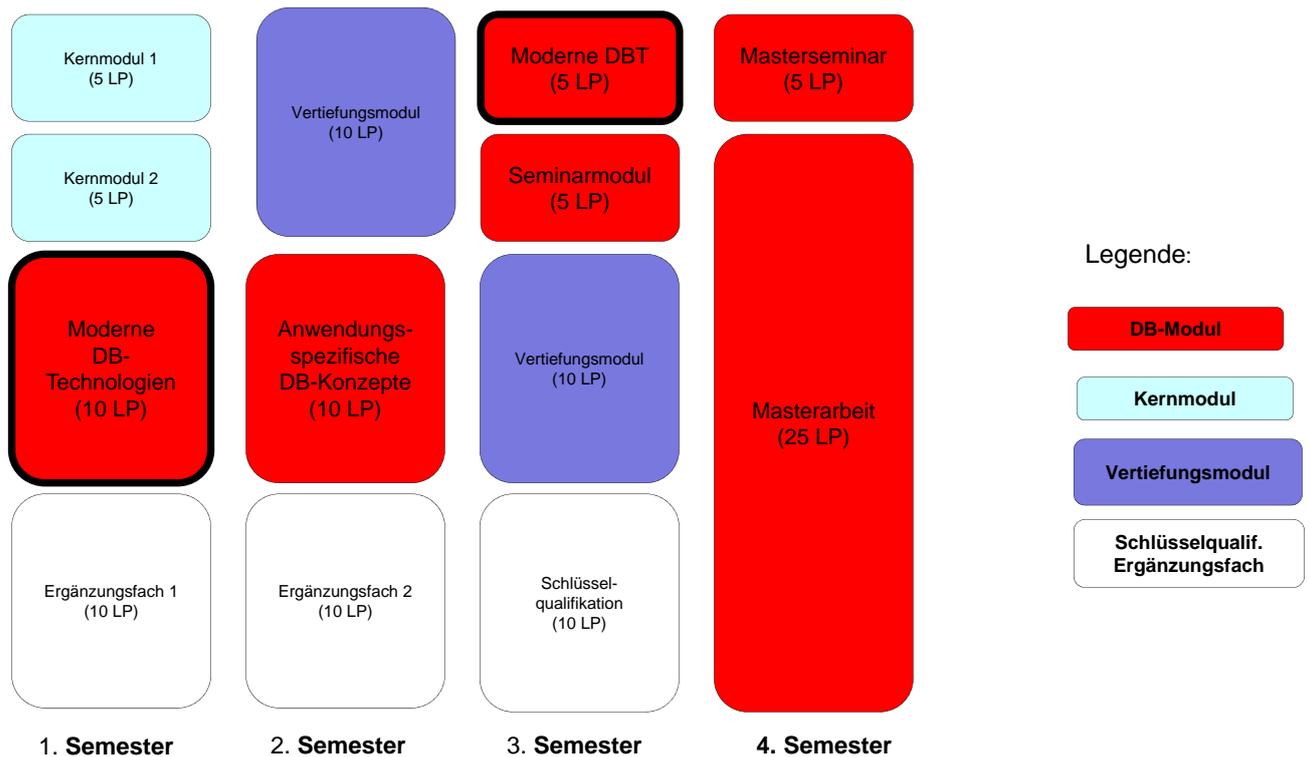
**DBS-Modul**

Mathematikmodul

Schlüsselqualif. Ergänzungsfach



## Informatik Masterstudium DBS-Profil



Legende:

**DB-Modul**

Kernmodul

Vertiefungsmodul

Schlüsselqualif. Ergänzungsfach



# Master of Science (M.Sc.) Wirtschaftsinformatik



**Wirtschaftsinformatik**



UNIVERSITÄT LEIPZIG

Flyer

	10 Leistungspunkte	10 Leistungspunkte	10 Leistungspunkte
1	ein Wahlpflichtmodul aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre	Integration und Architektur von Anwendungssystemen	ein Wahlpflichtmodul aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
2	Anwendungssysteme I - Modellierung und Management von Geschäftsprozessen	Zusammenstellung aus a) den Modulen des IWI Architekturen und Systeme für Geschäftsprozesse (5 LP), Softwaresystemfamilien und -produktlinien (10 LP) b) <b>den Modulen (10 LP) des Masterstudiengangs Informatik</b> c) den Wahlpflichtmodulen der Masterstudiengänge Betriebswirtschaftslehre bzw. Volkswirtschaftslehre	Software-Engineering in frühen Phasen
3	Anwendungssysteme II - Überbetriebliche Anwendungssysteme	ein Vertiefungsmodul aus Rechnernetze (10 LP), Service Science (10 LP), Intelligente Systeme (10 LP)	ein Wahlpflichtmodul aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
4	ein Vertiefungsmodul aus <b>Anwendungsbezogene Datenbankkonzepte (10 LP)</b> , Betriebliche Informationssysteme (10 LP) oder ein Wahlpflichtmodul aus 2. Semester a) oder b)	<b>Masterarbeit</b>	

<http://www.iwi.uni-leipzig.de/lehre/master>



## Master Wirtschaftsinformatik



**Wirtschaftsinformatik**

[http://db.uni-leipzig.de/studieren/downloadfile.php?abschluss\\_id=12213](http://db.uni-leipzig.de/studieren/downloadfile.php?abschluss_id=12213)

Sem.	10 LP	10 LP	10 LP
1	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre	Operations Research	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
2	Anwendungssysteme I	Wissensbasierte Systeme <b>und/oder</b> Computergrafik <b>und/oder</b> Datenbankpraktikum <b>und/oder</b> Textdatenbanken <b>und/oder</b> Linguistische Informatik <b>und/oder</b> Information Retrieval <b>und/oder</b> Grundlagen komplexer Systeme	Softwareengineering in frühen Phasen
3	Anwendungssysteme II	<b>Moderne Datenbanktechnologien</b> <b>oder</b> Integration und Architektur von Informationssystemen <b>oder</b> Softwaresystemfamilien	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
4	Anwendungssysteme III <b>oder</b> <b>Anwendungsbezogene Datenbankkonzepte</b>	Masterarbeit	



# 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	WS
	Bio- Datenbanken	Vertiefung	2	SS
	Ontologie-Management	Vertiefung	2	WS
	Cloud Data Management	Vertiefung	2	SS

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

LV im WS2012/13



WS12/13, © Prof. Dr. E. Rahm

## Mapping Module – Lehrveranstaltungen WS12/13

- Master-Module “Moderne Datenbanktechnologie” (zwei bzw. drei Veranstaltungen für kleines bzw. großes Modul):
  - Implementierung von DBS 1
  - Ontologie-Management
  - Data-Warehouse-Praktikum
  - Problemseminar Large-Scale Data Analytics
- Bachelor-Modul “Realisierung von Informationssystemen”
  - Implementierung von DBS 1
  - Ontologie-Management (evtl. Seminar)
- Seminar modul: Large Scale Data Analytics
- Bachelorseminar / Masterseminar
  - Vortrag über laufende Bachelor/Masterarbeit

WS12/13, © Prof. Dr. E. Rahm



# Leistungsbewertung

- Prüfungsklausur von 60 Minuten
  - Febr. 2013
  - überprüft konzeptionelles Wissen + Anwendungsfälle
- Klausurerfolg durch
  - Vorlesungsteilnahme und –nachbearbeitung
  - Durcharbeiten Lehrbuch
  - Online-Übungen



## Lernziele der Vorlesung IDBS

- fundierte Kenntnisse der Funktionsweise von DBS
- Implementierungstechniken u.a. zur
  - Sicherstellung einer hohen Performanz
  - 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
- sachkundige Beurteilung von kommerziell verfügbaren DBS
- Verfahren nicht nur für Datenbanksysteme relevant (-> Web-/Applikations-Server, Datei-Management,...)



# Vorläufige Vorlesungsübersicht

## 1. Einführung

- Architektur von DBS: Schichtenmodelle
- Tuning von DBS

## 2. E/A-Architekturen und Speicherhierarchien

- Speichertechnologien
- Disk-Arrays
- Nicht-flüchtige Halbleiterspeicher

## 3. DBS-Pufferverwaltung

- Speicherzuordnung
- Ersetzungsverfahren

## 4. Satzverwaltung

- Satzadressierung
- Satzabbildung
- lange Felder



# Vorlesungsübersicht (2)

## 5. Indexstrukturen

- Mehrweg-Bäume
- externe Hash-Verfahren
- Mehrdimensionale Zugriffspfade
- Bitlisten-Indizes
- Text-Indexierung

## 6. Algorithmen zur Implementierung relationaler Operationen

- Selektion
- Verbund / Join
- Sortierung

## 7. Anfrageoptimierung

- Übersetzung von DB-Anweisungen
- Optimierung
- EXPLAIN



# Literatur

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

<http://dbs.uni-leipzig.de/buecher/DBSI-Buch/inhalt.html>

- Weitere Lehrbücher

- Garcia-Molina, H., Ullman, J.D., Widom, J.: *Database System Implementation*. Prentice Hall, 2000
- Saake, Sattler, Heuer: *Datenbanken: Implementierungstechniken*, MITP-Verlag, 2011

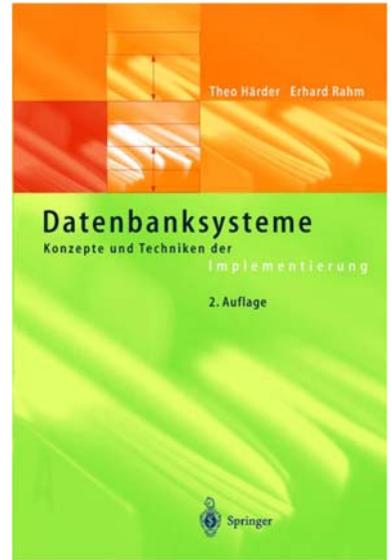
- Forschungsergebnisse

- Tagungsbände: VLDB (jährliche Konferenz "Very Large Data Bases"), SIGMOD (Konferenz der ACM Special Interest Group on Management of Data), IEEE Data Engineering, EDBT, BTW ...
- Zeitschriften: VLDB Journal (Very Large Data Bases), ACM TODS (Transactions on Database Systems), Datenbank-Spektrum ...

- DBLP-Portal: <http://dblp.uni-trier.de>

(>1 Million Referenzen, viele Links auf Volltexte, Homepages etc.)

- Google Scholar <http://scholar.google.com/> (Volltexte, 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.



# LOTS-Online-Übungen

## Aufgabe 3 (Nested-Block-Join)

0 von 2

Für den Gleichverbund zwischen R und S (je 100.000 Sätze, Blockungsfaktor 100) soll ein Nested-Block-Join genutzt werden. Welche Kombinationen zwischen verfügbarer Hauptspeichergröße M und Anzahl erreichbarer Plattenzugriffe (ohne Schreiben des Resultats) treffen zu?

- M=1001; 1 Million Plattenzugriffe
- M=5001; 2000 Plattenzugriffe
- M=1001; 2000 Plattenzugriffe
- M=501; 3000 Plattenzugriffe



- Erklärung ✗
- Erklärung ✗
- Erklärung ✗

## Aufgabe 4 (Hash-Join)

0 von 2

Markieren Sie die zutreffenden Aussagen.

- Durch die Nutzung von Bitvektoren lässt sich der Umfang der Hash-Tabellen kleiner halten
- Hash-Joins sollten nur eingesetzt werden, wenn die kleinere Eingabetabelle im Hauptspeicher gespeichert werden kann
- Hash-Joins können effektiv zur Realisierung von Mehr-Wege-Joins genutzt werden
- Hash-Join-Verfahren eignen sich zur Beantwortung von Equi-Join-Anfragen

- Erklärung ✗
- Erklärung ✗



## Lehrstuhl Datenbanksysteme

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



# Forschung



**Abteilung Datenbanken Leipzig**  
am Institut für Informatik

UNIVERSITÄT LEIPZIG  
Hilfe | Registrieren

Suchen

Startseite

Inhalte

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

**Forschung**

Publications (2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, 2004, ...)

Entity Matching for Big Data

Evolution of ontologies and mappings (Schema Evolution bibliography)

Mashup-based data integration (WETSUIT)

Schema and Ontology Matching (COMA++, GOMMA)

Bibliometric Analysis

wdilab

WS12/13, © Prof. Dr. E. Rahm

IDBS

## Main Research Areas

### ■ Data Integration

- Schema and Ontology Matching (COMA, GOMMA)
- Ontology Merging (ATOM)
- Object Matching / Entity Resolution
- Mashup Framework (WETSUIT)
- DI for Life Sciences

### ■ Cloud Data Management / Big Data:

- Load Balancing
- Dedoop: Deduplication based on Hadoop

### ■ Evolution of Ontologies and Mappings (DFG-funded)

### ■ Bibliometric Analyses

- Citation and affiliation analysis

■ EU project LINKED DESIGN

- Data integration platform for manufacturing design processes
- 12 partners: SAP, VW, AKER, ...



■ DFG project: Learning-based Link Discovery

- Entity resolution for Linked Data

■ Semi-automatic discovery of product offers for faked products in web shops

## Abschlussarbeiten / SHK-Jobs

- Wir suchen ständig interessierte Studenten für SHK/WHK-Tätigkeiten, Bachelor- und Masterarbeiten
- Mitarbeit an aktuellen Forschungsthemen des Lehrstuhls
- Kooperationen mit regionalen Unternehmen, z.B. eigenen Spinoffs
- Optimale Betreuung
- Anfragen per Formular oder persönlich