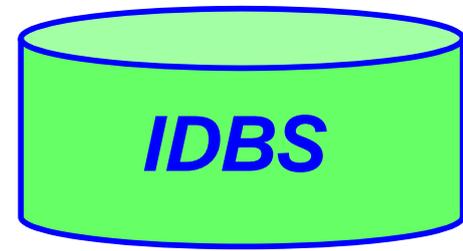


Implementierung von Datenbanksystemen 1 +2 (IDBS1 + IDBS2)



Wintersemester 2018/19

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)
 - Seminar modul
- 10-202-2011 – Masterseminar Informatik
 - Masterarbeit

■ Bachelor-Studium

- 10-201-2211 – Datenbanksysteme 1
- 10-201-2212 – Datenbanksysteme 2
- 10-201-2210 – Datenbankpraktikum
- **10-201-2224 – Realisierung von Informationssystemen**
 - Seminar modul
- 10-201-2010 – Bachelorseminar Informatik
 - Bachelorarbeit



Mapping Module – Lehrveranstaltungen WS18/19

- Master-Module **Moderne Datenbanktechnologie** (5 LP) und Bachelor-Modul **Realisierung von Informationssystemen**: zwei Veranstaltungen aus
 - IDBS1
 - IDBS2
 - Data Mining
- Master-Module **Moderne Datenbanktechnologie** (10 LP)
 - 2 Vorlesungen aus: IDBS1, IDBS2, Data Mining
 - Data Warehouse-Praktikum (ggf. Seminar *Secure Data Processing*)
- **Seminarmodul (Bachelor oder Master)**
 - Problemseminar *Secure Data Processing*– Vorbesprechung 26.10.
- **Bachelorseminar / Masterseminar**
 - Vortrag über laufende Bachelor/Masterarbeit
- überzählige Vorlesung (z.B. IDBS2/Data Mining) kann auch in Modul im SS19 eingebracht werden



Masterstudium Informatik

Sem	5 LP	5LP	10 LP	10 LP
1.	Kernmodul I	Kernmodul II	Vertiefungsmodul I	Ergänzungsfach
2.	Kernmodul III	Seminarmodul	Vertiefungsmodul II	Ergänzungsfach
3.	Vertiefungsmodul III		Vertiefungsmodul IV	Fakultätsinterne Schlüsselqualifikation
4.	Masterseminar	Masterarbeit		

- mögliche Schwerpunkte (Ausweisung im Zeugnis)
 - Big Data
 - Medizinische Informatik
- individuelle Schwerpunktbildung, zB. Datenmanagement/DB



Master Informatik mit DB-Profil

Sem	5 LP	5LP	10 LP	10 LP
1.	Kernmodul I	Kernmodul II	Moderne DB-Technologien	Ergänzungsfach
2.	Vertiefungsmodul III		Anwendungsspezifische DB-Konzepte	Ergänzungsfach
3.	Moderne DB-Technologien	Seminarmodul	Vertiefungsmodul IV	Fakultätsinterne Schlüsselqualifikation
4.	Masterseminar	Masterarbeit		

DB-Modul

- *Moderne DBT* und *Anwendungsspez. DB-Konzepte* können in verschiedenen Semestern in großer und kleiner Ausführung belegt werden
 - falls disjunkte Untersetzung vorliegt

Master Informatik mit Schwerpunkt Big Data

Sem	5 LP	5LP	10 LP	10 LP
1.	Kernmodul I	Kernmodul II	Moderne DB-Technologien	Ergänzungsfach
2.	Kernmodul III	Seminarmodul	Big Data Vertiefung	Ergänzungsfach
3.	Big Data Vertiefung		Vertiefungsmodul zu Data Mining / Visualisierung	Fakultätsinterne Schlüsselqualifikation
4.	Masterseminar	Masterarbeit		

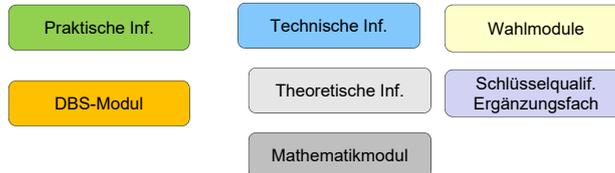
Big Data Modul

- obligatorisches Vertiefungsmodul *Moderne DB-Technologien*
- weitere Kern/Vertiefungsmodule: [Anwend.spez. DBK](#), neuroinspirierte Informationsverarbeitung, statistisches Lernen, Visualisierung, Information Retrieval, Semantic Web, Computational Advertising
- Ergänzungsfächer z.B. aus Bioinformatik oder Digital Humanities

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:



Master Wirtschaftsinformatik

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 und/oder Computergrafik und/oder Datenbankpraktikum und/oder Textdatenbanken und/oder Linguistische Informatik und/oder Information Retrieval und/oder Grundlagen komplexer Systeme	Softwareengineering in frühen Phasen
3	Anwendungssysteme II	Moderne Datenbanktechnologien oder Integration und Architektur von Informationssystemen oder Softwaresystemfamilien	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
4	Anwendungssysteme III oder Anwendungsbezogene Datenbankkonzepte	Masterarbeit	



Wirtschaftsinformatik

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
	Cloud Data Management	Vertiefung	2	SS
	NoSQL-DB	Vertiefung	2	SS
	Data Mining	Vertiefung	2	WS

Name	Typ	Sem.
DB-Praktikum	Praktikum	SS
Data-Warehouse-Praktikum	Praktikum	WS
Big Data Praktikum	Praktikum	SS
Seminarmodul	Seminar	WS
Bachelor-seminar	Seminar	SS+WS
Master-seminar	Seminar	SS+WS

LV im WS2018/19



WS18/19, © Prof. Dr. E. Rahm

Leistungsbewertung

- Prüfungsklausur IDBS1/IDBS2 von je 60 Minuten
 - IDBS1: wählbarer Termin Dez. oder Jan. oder Feb.
 - IDBS2: Feb 2019
 - überprüft konzeptionelles Wissen + Anwendungsfälle
- Klausurerfolg durch
 - Vorlesungsteilnahme und –nachbearbeitung
 - Online-Übungen
 - Literatur

WS18/19, © Prof. Dr. E. Rahm

0-10

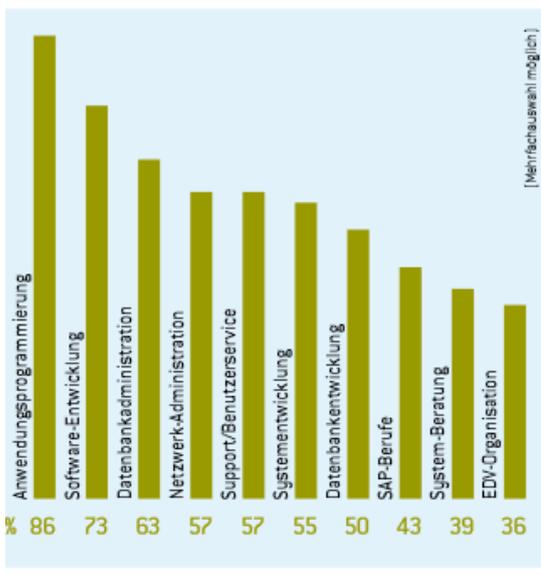


Lernziele der Vorlesung IDBS

- 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 (-> Big Data, Web-/ Applikations-Server, Datei-Management,...)



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



Vorlesungsübersicht IDBS1

1. Einführung

- Architektur von DBS: Schichtenmodelle
- Tuning von DBS

2. E/A-Architekturen und Speicherhierarchien

- Speichertechnologien
- Disk-Arrays
- nicht-flüchtige Halbleiterspeicher

3. DBS-Pufferverwaltung

- Ersetzungsverfahren

4. Satzverwaltung

- Zuordnung Sätze – Seiten, Clusterung, BLOBs
- Satzadressierung
- Column Stores

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

Vorlesungsübersicht IDBS2

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

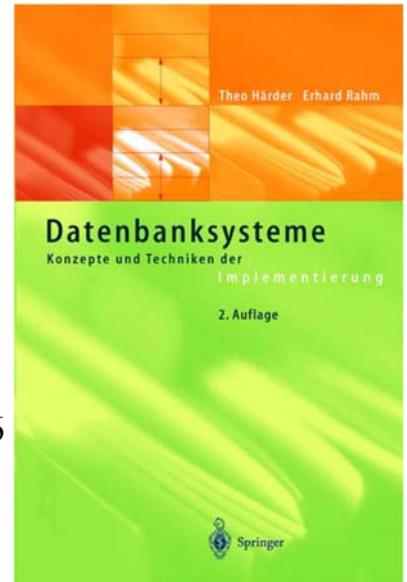
- Geschachtelte Transaktionen
- Transaktionsketten (Sagas)

7. DBS- Benchmarks

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

- Dittrich, J.: *Patterns in Data Management: A Flipped Textbook*. 2016
Website: Datenbankenlernen.de
- Saake, Sattler, Heuer: *Datenbanken: Implementierungstechniken*, MITP-Verlag, 2011
- 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: 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 HauptspeichergroÙe 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

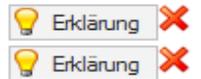


Aufgabe 4 (Hash-Join)

0 von 2

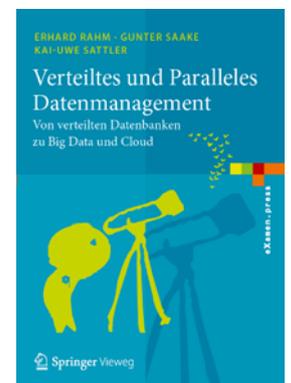
Markieren Sie die zutreffenden Aussagen.

- 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

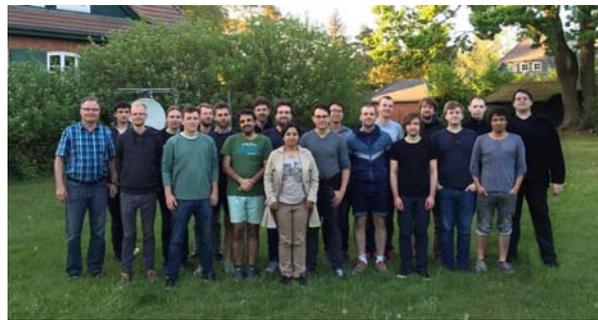


Lehrstuhl Datenbanksysteme

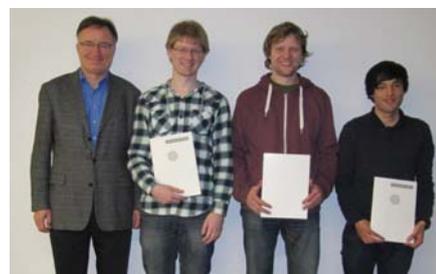
- seit 1994 am Institut für Informatik
- umfangreiches Lehrangebot
 - Vorlesungen, Praktika, Seminare
 - Online-Übungssystem LOTS
 - Eigene Lehrbücher: MRDBS, IDBS, ...



Oberseminare an Uni-Außenstelle Zingst – seit 2001



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



Forschung



Abteilung Datenbanken Leipzig

am Institut für Informatik

 Suchen

UNIVERSITÄT LEIPZIG

Hilfe | Registrieren



Startseite

login

Inhalte

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

Forschung

Publications (2018, 2017, 2016, 2015, 2014, 2013, 2012, 2011, 2010, 2009, 2008, 2007, 2006, 2005, ...)

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

FAst Multi-source Entity Resolution system (FAMER)

Graph-based data analysis (GRADOOP, BIIIG)

Privacy-Preserving Record Linkage for Big Data

Entity Matching for Big Data (Dedoop), LOD Link Discovery

Semantic annotations: ELISA project, LHA Annotation Linking

Schema and Ontology Matching (COMA++, GOMMA, STROMA), SemRep repository, Ontology Merging (ATOM)

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

Web Data Integration (WDI) Lab [↗](#)

Object Matching / Entity Resolution

Bibliometric Analysis

Current prototypes and implementations



WS18/19, © Prof. Dr. E. Rahm

Big Data Center ScaDS – Phase 2 (Competence Center for Scalable Data Services and Solutions)



Application Sciences
and Business
Applications

Service
Center

Visual Analytics

Scalable Visual Analytics

Immersive Visual Interaction

Big Data Integration & Analytics

Big Data Integration

Data Analytics

Scalable and Secure Data Platforms

Scalable Architectures

Hardware-based Data Security



E. Rahm



M. Bogdan



G. Scheuermann



P. Stadler



M. Potthast



G. Heyer



B. Franczyk

WS18/19, © Prof. Dr. E. Rahm

