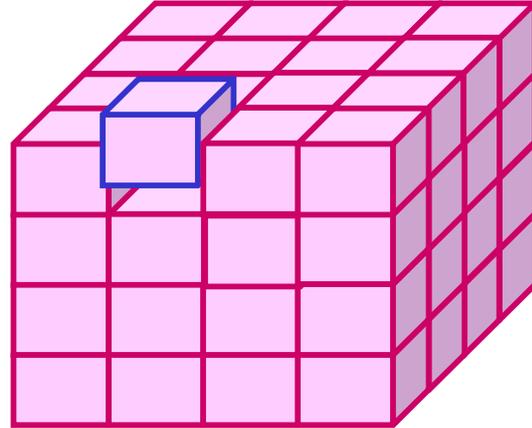


Data Warehousing

Sommersemester 2018

Prof. Dr. E. 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 SS18

- Master-Module “Anwendungsbezogene Datenbankkonzepte” (zwei bzw. drei Veranstaltungen für kleines bzw. großes Modul):
 - Data Warehousing
 - NoSQL-Datenbanken
 - Big-Data-Praktikum
 - ggf. DBS2
- Bachelor-Modul “Realisierung von Informationssystemen”
 - Data Warehousing
 - NoSQL-Datenbanken
- Bachelorseminar / Masterseminar
 - Vortrag über laufende Bachelor/Masterarbeit
 - extern an Uni-Außenstelle in Zingst (22.-25. Mai 2018)



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:

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

SS18, © Prof. Dr. E. Rahm



Master Wirtschaftsinformatik

Sem.	10 LP	10 LP	10 LP
1	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre	Integration und Architektur von Anwendungssystemen	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
2	Anwendungssysteme I Modellierung und Management von Geschäftsprozessen	<p>Zusammenstellung aus:</p> <p>a) den Modulen des Masterstudiengang Informatik: Neuroinspirierte Informationsverarbeitung, Ausgewählte Verfahren des Mobile Advertising, Grundlagen komplexer Systeme, Textdatenbanken</p> <p>b) den Wahlpflichtmodulen der Masterstudiengänge BWL/VWL</p> <p>c) den Modulen der Wirtschaftsinformatik: Architekturen und Systeme für Geschäftsprozesse, Softwaresystemfamilien und -produktlinien, Business Innovation</p>	Advanced Software Engineering
3	Anwendungssysteme II Überbetriebliche Anwendungssysteme	<p>Vertiefungsmodul aus</p> <ul style="list-style-type: none"> > Service Science, > Computational Advertising > Mobile Peer-to Peer Systeme, > Wissensrepräsentation 	Wahlpflichtmodule aus den Masterstudiengängen Betriebswirtschaftslehre bzw. Volkswirtschaftslehre
4	<p>Vertiefungsmodul aus</p> <ul style="list-style-type: none"> > Anwendungsbezogene Datenbankkonzepte, > Betriebliche Informationssysteme oder > ein Wahlpflichtmodul aus 2. Semester a) oder c) 	Masterarbeit	



Wirtschaftsinformatik

SS18, © 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
	Data Warehousing	Vertiefung	2	SS
	Datenintegration	Vertiefung	2	WS
	Cloud Data Management	Vertiefung	2	SS
	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

LV im SS2018



Leistungsbewertung

- Prüfungsklausur von 60 Minuten
 - Juli 2018
 - überprüft konzeptionelles Wissen + Anwendungsfälle

- Klausurerfolg durch
 - Vorlesungsteilnahme und –nachbearbeitung
 - Online-Übungen
 - Literatur



Lernziele der Vorlesung

- Überblicks- und Detailkenntnisse zum Aufbau, Betrieb, Funktionsweise und Einsatz von Data Warehouses
- Kenntnis von Architekturalternativen
- Modellierung von Data Warehouses
- Überblick zu Analyseverfahren (OLAP, Data Mining)
- Probleme der Datenintegration / Datenbereinigung
- Implementierungstechniken zur Sicherstellung einer hohen Performanz
- sachkundige Beurteilung von kommerziell verfügbaren Data-Warehouse-Lösungen



Vorläufige Vorlesungsübersicht

1. Einführung

2. Architektur von Data Warehouse-Systemen

3. Modellierung / Operationen

- MOLAP, ROLAP, hybride Ansätze
- SQL-Erweiterungen, Operatoren

4. Metadaten und ETL (Extraktion, Transformation, Laden)

- Metadaten-Integration, Schema Matching
- Datenbereinigung (Data Cleaning) / Instanz-Matching

5. Performance-Techniken

- Indexstrukturen
- Materialisierte Sichten
- Nutzung paralleler Datenbanken

6. Data Mining-Verfahren

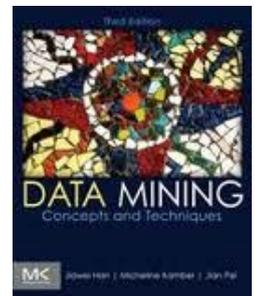
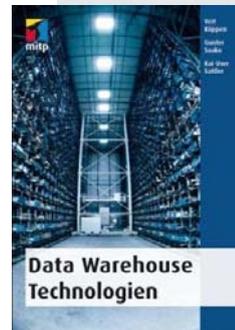
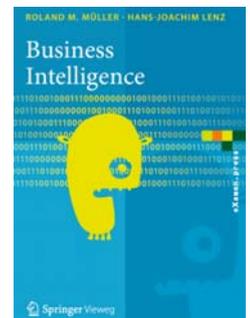
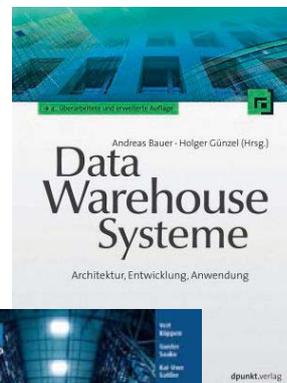
- Clusterung
- Assoziationsregeln
- Deep Learning



Literatur

■ Buchauswahl

- Bauer/Günzel (Hrsg.): *Data Warehouse Systeme*. 4. Auflage, dpunkt 2013
- Köppen/Saake/Sattler: *Data Warehouse Technologien*. 2. Auflage, Mitp 2014
- Müller/Lenz: *Business Intelligence*. Springer 2013
- Han / Kamber / Pei: *Data Mining: Concepts and Techniques*. 3rd edition, Morgan Kaufmann 2011



■ Eigene Forschungsarbeiten: <http://dbs.uni-leipzig.de>

- Schema Matching (COMA++, GOMMA)
- Data Cleaning / Instanz-Matching (DEDOOP, FAMER)



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:

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.





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



Abschlussarbeiten / SHK-Jobs

- Studenten mit guten DB-Kenntnissen für
 - Bachelor- oder Masterarbeiten
 - SHK/WHK-Tätigkeiten
- Mitarbeit an aktuellen Forschungsthemen
 - des Lehrstuhls bzw
 - des BigData-Zentrums ScaDS



Open House Day

@SAP HANA Student Campus

Where?

SAP Headquarters Walldorf/Germany

When?

May 15th, 2018 09:30 - 17:00

Who?

Students, PhD Students, Faculty Members

About What?

Database Research and Development

More Information and Registration:

tinyurl.com/hcod18

Internships,
Theses

Career
Opportunities

Free Lunch
and Snacks ☺



SAP

4th International Summer School for Big Data and Machine Learning

Jun. 30th - Jul. 6th

■ „Streamline“ Hackathon

- Weekend: : June 30-July 1st
- participants learn building complex real-time data analytics pipelines, e.g. using Apache Flink or Apache Spark

■ Summer school: July 2nd-6th

- international experts from academia and industry

■ 5-10 sponsored admissions for qualified students

