

Vorlesung "Informationssysteme"
Dr. Ralf Schenkel, Prof. Dr. Gerhard Weikum
Universität des Saarlandes, Sommersemester 2004

Inhalt und Lernziel:

Die Vorlesung vermittelt grundlegende Kenntnisse über Konzepte und Schnittstellen von Datenbanksystemen und anderen Arten von Informationsdienstsoftware (z.B. Search-Engines oder Data-Mining-Tools) sowie der Anwendungsentwicklungswerkzeuge zur Realisierung von Informationssystemen mit strukturierten, semistrukturierten und unstrukturierten Daten. Schwerpunktthemen sind Anfragesprachen für Datenbanksysteme, Daten- und Prozeßmodellierung sowie Relevanzmodelle für Information-Retrieval und Data-Mining. Die entsprechenden Grundlagen aus der Logik und Stochastik werden in der Vorlesung eingeführt.

Organisatorisches:

Vorlesungs- und Übungsunterlagen

Unterlagen zur Vorlesung und zu den Übungen werden unter der folgenden URL zur Verfügung gestellt: <http://www.mpi-sb.mpg.de/units/ag5/teaching/ss04/is.htm> .

Termine und Betreuung

Vorlesung: Dienstag 9-11 und Donnerstag 9-11 Uhr in 27/001

Übungen: Mi und Fr

Die Vorlesung beginnt am 20. April; die erste Übungsstunde ist in der zweiten Vorlesungswoche.

Die Übungskoordinatoren sind

Dipl.-Math. Sergej Sizov (sizov@mpi-sb.mpg.de) für die praktischen Übungen und

Dipl.-Inform. Christian Zimmer (czimmer@mpi-sb.mpg.de) für die Papierübungen.

Sprechstunde von Dr. Schenkel ist ... oder n.V. (schenkel@mpi-sb.mpg.de)

Sprechstunde von Prof. Weikum ist Dienstag 14-15 Uhr in 46/401 oder n.V. (weikum@mpi-sb.mpg.de).

Leistungsprüfung

Es werden 9 benotete Leistungspunkte vergeben, wenn folgende Voraussetzungen erfüllt sind:

1. a) erfolgreiche Teilnahme an zwei Teilklausuren in der Mitte und am Ende des Semesters
oder
b) erfolgreiche Teilnahme an einer Teilklausur und der Wiederholungsklausur Anfang Oktober,
2. mindestens einmaliges Präsentieren von Lösungen in der Übungsgruppe,
3. erfolgreiche Bearbeitung der praktischen Übungen (Teamarbeit in Dreiergruppen möglich).

Die Note wird aus den Ergebnissen der zwei bestandenen (Teil-) Klausuren berechnet.

Geplante Gliederung:

1. Einführung und Überblick: Anwendungen, Systeme, Prinzipien

Teil I: Suchmaschinen

2. Vektorraummodell für Suchmaschinen
3. Automatische Klassifikation von Dokumenten
4. Linkanalyse für Autoritäts-Ranking

Teil II: Datenbankschnittstellen

5. Relationenmodell und algebraorientierte Anfragesprachen
6. Logikorientierte Anfragesprachen
7. Datenbanksprache SQL
8. Anwendungsentwicklung mit SQL und JDBC
9. Integritätssicherung mit SQL
10. Objektorientierte und objekt-relationale Datenmodelle

Teil III: Datenbank- und Anwendungsentwurf

11. Relationale Entwurfstheorie
12. Datenbankentwurf mit UML
13. Prozessmodellierung mit Statecharts

Teil IV: Implementierungskonzepte von Datenbanksystemen

14. Datenspeicherung, Indexstrukturen und Anfrageauswertung
15. Transaktionsverwaltung: Concurrency-Control
16. Transaktionsverwaltung: Recovery

Teil V: Informationsdienste

17. OLAP und Data-Warehouses
18. Daten-Mining: Klassifikation, Assoziationsregeln
19. Semistrukturierte Daten und XML-Anfragesprachen

Literatur:

Zu Teil I: Suchmaschinen

- Ricardo Baeza-Yates, Berthier Ribeiro-Neto: Modern Information Retrieval, Addison-Wesley, 1999
- Christopher D. Manning, Hinrich Schütze: Foundations of Statistical Natural Language Processing, MIT Press, 1999
- Soumen Chakrabarti: Mining the Web, Morgan Kaufmann, 2002
- Pierre Baldi, Paolo Frasconi, Padhraic Smyth: Modeling the Internet and the Web, Wiley&Sons, 2003
- Ian Witten, Alistair Moffat, Timothy C. Bell: Managing Gigabytes, Academic Press, 1999

Zu Teil II: Datenbankschnittstellen

- Jeffrey D. Ullman, Jennifer Widom: First Course in Database Systems, Prentice Hall, 1997
- Patrick O'Neil, Elizabeth O'Neil: Database Principles, Programming, and Performance, Morgan Kaufmann, 2001
- Joachim Biskup: Grundlagen von Informationssystemen, Vieweg-Verlag, 1995
- Andreas Heuer, Gunter Saake: Datenbanken - Konzepte und Sprachen, International Thomson Publishing, 2000
- Alfons Kemper, Andre Eickler: Datenbanksysteme - eine Einführung, 5. Auflage, Oldenbourg, 2004
- Gottfried Vossen: Datenbankmodelle, Datenbanksprachen und Datenbankmanagement-Systeme, Oldenbourg, 1999
- Can Türker: SQL:1999 & SQL:2003, dpunkt-Verlag, 2003

Zu Teil III: Datenbank- und Anwendungsentwurf

- Carlo Batini, Stefano Ceri, Shamkant Navathe: Conceptual Database Design - An Entity-Relationship Approach, Addison-Wesley, 1991
- Toby J. Teorey: Database Modeling & Design, Morgan Kaufmann, 1998
- Robert J. Muller: Database Design for Smarties - Using UML for Data Modeling, Morgan Kaufmann, 1999
- Martin Fowler, Kendall Scott, Grady Booch: UML Distilled, Addison-Wesley, 1999
- David Harel, Michal Politi: Modeling Reactive Systems - the Statemate Approach, McGraw Hill, 1998
- Edmund M. Clarke, Orna Grumberg, Doron A. Peled: Model Checking, MIT Press, 2000

Zu Teil IV: Implementierungskonzepte von Datenbanksystemen

Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom: Database System Implementation, Prentice Hall, 1999

Theo Härder, Erhard Rahm: Datenbanksysteme - Konzepte und Techniken der Implementierung, Springer, 2001

Raghu Ramakrishnan, Johannes Gehrke: Database Management Systems, McGraw Hill, 2000

Gerhard Weikum, Gottfried Vossen: Transactional Information Systems - Theory, Algorithms, and the Practice of Concurrency Control and Recovery, Morgan Kaufmann, 2001

Zu Teil V: Informationsdienste

Jiawei Han, Micheline Kamber: Data Mining - Concepts and Techniques, Morgan Kaufmann, 2001

Martin Ester, Jörg Sander: Knowledge Discovery in Databases – Techniken und Anwendungen, Springer-Verlag 2000

Harald Schöning: XML und Datenbanken, Hanser-Verlag, 2003

Henk Blanken, Torsten Grabs, Hans-Jörg Schek, Ralf Schenkel, Gerhard Weikum (Editors): Intelligent Search on XML Data, Springer-Verlag, 2003