Datenbanksysteme II: Implementation of Database Systems
• Slides in English, Vortrag auf Deutsch
• Much input from
  – Prof. J-C Freytag, HU Berlin
  – Prof. K-U Sattler, TU Illmenau
  – Prof. A Kemper, Dr. Eickler, TU München
• Goya
• Prof Freytag / Prof Leser
• DBS-1 formale Voraussetzung?
• Was studieren Sie?
My first contact with database

• Father: “I need to track my customers”
  - Name, address, profession, prior contracts

• Young Ulf: “No problem”
  - ~1984: Turbo Pascal 4.0, Schneider CPC646, 512 KB main memory
  - Load customers from disk into memory
  - Repeat until “Q”
    • (S)earch and list customers
    • (E)dit customer
    • (D)elete customer
    • (I)nsert customer
    • (Q)uit
  - Write customers to disk
  - Invoice: 250.- DM
Story part 2

- Father: “I need to track my offers”
  - Customers have projects and call for bids, father makes offers
  - Many customers have many projects over time
- Young naïve Ulf: “No problem”
  - Reuse existing architecture
  - Load offers from disk into memory
  - Repeat until “Q”
    - (S)earch and list offers
    - (E)dit offer
    - (D)elete offer
    - (I)nsert offer
    - (Q)uit
  - Write offers to disk
  - Invoice: 400.- DM
Part 3

• Disaster!
  – No dynamic arrays at this time
  – Max. 500 customers
  – Max. 40 offers
  – Max. 2 KB per offer
  – Gives 20,000 times 2,000 = 40 MEGABYTE!
    • App. size of hard disc
  – No way to load and hold all data at startup

• Wrong architecture
Solution 1

• Repeat until “Q”
  – (E)dit offer
    • Open file
    • Read and write until record found
    • Read, edit, and write record
    • Read and write until EOF
  – (D)elete offer
    • ...
  – (I)nsert offer
    • ...
  – (Q)uit

• Very slow
Scalability

Abbildung 6.3: Entwicklung der Laufzeiten im Vergleich zu anderen Algorithmen

Quelle: G. Witterstein, Dissertation, 2005
Solution 2

- Someone should take care of finding, loading, storing data
- I want to concentrate on what to store
- Buy dBase IV
Lessons learned

- **Scalability** is an issue
  - Works fine for small datasets, breaks down for large datasets
- **Data is Information is Knowledge is essential**
  - If offers-file corrupted – father goes out of business
    - Which will affect me too
- When is comes to large volumes of data – get a database and a **good database programmer**

- **Think before you program**
  - Project 100% over budget (dBase IV license)
  - Project 300% over time (6 months instead of 2)
Today’s Database Systems

• Databases are essential parts of enterprise infrastructures
  – More important than OS
  – Long-running, expensive, essential investment
  – Holds the most important business asset: information

• Database administrator is a profession
  – Many programmers, but few good database developers/ admins
  – One of the skills most demanded in industry

• RDMS became an often “invisible” piece of software
  – “So nützlich wie fließendes Wasser” (G. Weikum, MPI Saarbrücken)
Importance

- **DBs are behind almost every non-trivial piece of software**
  - Data needs to be stored
    - Disk access is the main performance bottleneck
    - Hence: **Optimize disk access** in the RDBMS
  - Data is manipulated from many clients
    - Concurrent access quickly screws up data
    - Hence: **Synchronize access** through the RDBMS
  - Data is used by many applications with different requirements
    - No good to design app-specific optimal data structures
    - Hence: Use **application independent** languages
  - Systems crash
    - Crashes cannot be avoided
    - Hence: **Secure most valuable asset** in a centralized manner
History

• Prior to relational databases
  - Network databases, hierarchical databases, COBOL

• Waves and hypes
  - Relational databases since 1980th
  - Object-oriented databases early 1990th
  - XML databases early 2000\textsuperscript{th}
  - “Big data” since 2008

• Few „big“ vendors survived
  - Oracle, IBM DB2, Microsoft SQL Server

• Some niche players
  - Open Source: MySQL, PostgreSQL, Ingres, SAP DB
  - Data Warehousing: Teradata, Tandem
  - Mobile databases: Sybase
Complexity

- A RDBMS is a very, very complex piece of software
- One code base
  - Oracle9i Database Release 2 Documentation for Microsoft Windows
  - Oracle9i Database Release 2 Documentation for 64-bit Windows
  - Oracle9i Database Release 2 Documentation for Solaris Operating System (SPARC)
  - Oracle9i Database Release 2 Documentation for AIX-Based Systems
  - Oracle9i Database Release 2 Documentation for Compaq Tru64 UNIX
  - Oracle9i Database Release 2 Documentation for HP 9000 Series HP-UX
  - Oracle9i Database Release 2 Documentation for HP-UX Itanium
  - Oracle9i Database Release 2 Documentation for hp OpenVMS Alpha
  - Oracle9i Database Release 2 Documentation for IBM OS-390 (MVS)
  - Oracle9i Database Release 2 Documentation for Linux x86-64
  - Oracle9i Database Release 2 Documentation for Linux x86
  - Oracle9i Database Release 2 Documentation for Linux Itanium
  - Oracle9i Database Release 2 Documentation for IBM Power based Linux
  - Oracle9i Database Release 2 Documentation for Linux/390
  - Oracle9i Database Release 2 Documentation for Fujitsu Siemens Computers BS2000/OSD
  - Oracle9i Database Release 2 Documentation for Stratus Continuum 600 and 1200 Series
## List of Books 1

### Oracle8i Database List of Books (Release 2 (9.2)) - Mozilla Firefox

<table>
<thead>
<tr>
<th>Title</th>
<th>Contents</th>
<th>Search</th>
<th>Index</th>
<th>PDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Access Manager for AS/400 Installation and User's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrator's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrator's Guide for Windows</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Administrator's Reference for UNIX Systems: AIX-Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, Linux Intel, and Sun Solaris</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Replication</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Advanced Security Administrator's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Developer's Guide - Advanced Queuing</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Developer's Guide - Fundamentals</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Developer's Guide - Large Objects (LOBs)</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Developer's Guide - Object-Relational Features</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Application Developer's Guide - Workspace Manager</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backup and Recovery Concepts</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Backup and Recovery Documentation Online Roadmap</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client Installation Guide for 32-Bit Windows</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Client Release Notes for 32-Bit Windows</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COM Automation Feature Developer's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Concepts</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Cartridge Developer's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Guard Brokerson</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Guard Concepts and Administration</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Mining Administrator's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Mining Concepts</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data Warehousing Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Database Tuning with the Oracle Tuning Pack</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Directory Service Integration and Deployment Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dynamic Services User's and Administrator's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Edition Installation Guide for OS/390</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Edition Messages Guide for OS/390</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Edition Release Notes for OS/390</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Edition System Administration Guide for OS/390</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Manager Administrator's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Manager Concepts Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enterprise Manager User's Guide</td>
<td>Contents</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
List of Books 2

<table>
<thead>
<tr>
<th>Book Title</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Oracle8 Database List of Books (Release 2 9.2)) - Mozilla Firefox</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Enterprise Manager Configuration Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Enterprise Manager Event Test Reference Manual</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Enterprise Manager Messages Manual</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Getting Started for Windows</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Getting Started with Oracle Change Management Pack</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Getting Started with Oracle Management Pack for Oracle Applications</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Getting Started with the Oracle Diagnostics Pack</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Getting Started with the Oracle Standard Management Pack</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Globalization Support Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Heterogeneous Connectivity Administrator’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Installation Guide for 32-Bit Windows</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Installation Guide for UNIX Systems: AIX-Based Systems, Compaq Tru64 UNIX, HP 9000 Series HP-UX, Linux Intel and Sun Solaris</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Intelligent Agent User’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>interMedia Annotator User’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>interMedia Java Classes User’s Guide and Reference</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Internet Directory Administrator’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Internet Directory Application Developer’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Java Developer’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Java Stored Procedures Developer’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>JDBC Developer’s Guide and Reference</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>JPublisher User’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Label Security Administrator’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Master Glossary</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Master Index</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Holders</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Workbench Frequently Asked Questions (FAQ)</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Workbench Reference Guide for IBM DB2/AS400 V4R5 Migrations</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Workbench Reference Guide for Informix Dynamic Server 7.3 Migrations</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Workbench Reference Guide for Microsoft SQL Server and Sybase Adaptive Server Migrations</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Workbench Reference Guide for MySQL 3.22, 3.23 Migrations</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Migration Workbench Release Notes</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Net Services Administrator’s Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
<tr>
<td>Net Services Reference Guide</td>
<td>Contents / Search / Index / PDF</td>
</tr>
</tbody>
</table>
List of Books 3

<table>
<thead>
<tr>
<th>Title</th>
<th>Publisher</th>
<th>URL</th>
</tr>
</thead>
<tbody>
<tr>
<td>New Features</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLAP Developer's Guide to the OLAP API</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLAP Developer's Guide to the OLAP DML</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OLAP User's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle C++ Call Interface Programmer's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle Call Interface Programmer's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle Call Interface Programmer's Guide Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle Call Interface Programmer's Guide Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle Objects for OLE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle Objects for OLE C++ Class Library</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Oracle Transparent Gateway for IBM DB2/400 Installation and User's Guide for IBM AS/400</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance Planning</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Performance Tuning Guide and Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL/SQL User's Guide and Reference</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro*C/C++ Precompiler Getting Started for Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro*C/C++ Precompiler Programmer's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro*COBOL Precompiler Getting Started for Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro*COBOL Precompiler Programmer's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro*FORTRAN Supplement to the Oracle Precompilers Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro*PL/I Supplement to the Oracle Precompilers Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway and Tools for IBM MQSeries for MS Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway and Tools for IBM MQSeries for UNIX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway for APPC Installation and Configuration Guide for UNIX</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway for APPC Installation and Configuration Guide for Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway for APPC Messages Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway for APPC User's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Procedural Gateway for APPC User's Guide Windows</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Programmer's Guide to the Oracle Precompilers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Provider for OLE DB Developer's Guide</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Application Clusters Concepts</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Application Clusters Deployment and Performance</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Application Clusters Documentation Online Roadmap</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Application Clusters Real Application Clusters Guard I Concepts and Administration</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Real Application Clusters Real Applications Clusters Guard I Configuration Guide Release 2 (9.2.0.1.0) for UNIX Systems: AIX-Based</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Done
List of Books 4
## List of Books 5

<table>
<thead>
<tr>
<th>Title</th>
<th>Contents</th>
<th>Search</th>
<th>Index</th>
<th>PDF</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transparent Gateway for DB2 Installation and User’s Guide for IBM C9/390</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Transparent Gateway for IBM DRDA Installation and User’s Guide for UNIX</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Transparent Gateway for IBM DRDA Installation and User’s Guide for Windows</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Ultra Search Online Documentation</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Universal Installer Concepts Guide</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>User-Managed Backup and Recovery Guide</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Utilities</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Workflow Client Installation Notes (release 2.6.2)</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>Workflow Server Installation Notes (release 2.6.2)</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>XML API Reference - XDK and Oracle XML DB</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>XML Database Developer’s Guide - Oracle XML DB</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
<tr>
<td>XML Developer’s Kits Guide - XDK</td>
<td>Contents</td>
<td>Search</td>
<td>Index</td>
<td>PDF</td>
</tr>
</tbody>
</table>
3 Tier

- **Servlets/EJB**
  - Upwards: OO Interface
  - Application Server
- **Downwards: SQL**
  - DBMS
  - Database 1
  - Database 2

**Presentation**

**Application logic**

“Business processes”

“State”

Storage, Backup & Recovery

Synchronization
SAP R/3: Enterprise Resource Planning

ONE central database server
(Possibly clustered)

LAN

WAN (Internet)
Making the difference

• Databases are **not self-**
  – Self-administrating, -adapting, -developing, -healing
• Work mostly in the background – no GUI or dialogs
• Good database development requires **intensive knowledge**
  about the why & what & how
  – Which index to use and where to place it?
  – Which synchronization level?
  – When use parallel servers?
  – Optimization possibilities?
  – Where to look if something goes wrong?
DBS2: Implementation of Database Systems

- **Lecture**  
  - Tuesday, 15 – 17, RUD26, 1‘305  
  - Wednesday, 11 – 13, RUD26, 1‘305

- **Contact**
  Ulf Leser
  Room: IV.401
  Tel: (030) 2093 – 3902
  Mail: leser (at) informatik.hu-berlin.de
Praktikum & Exams

• “Praktikum” by Andre Koschmieder
  - Presence & commitment are necessary
  - Implementation of file-/ buffer-/ index manager in C++
  - Two groups (?)
    • Tuesday, 13-15, RUD 26, Room 1'308
    • Wednesday, 13-15, RUD 26, Room 1'308
  - Starts next week

• Examination
  - 30 minutes, oral exam
  - Dates will be set mid-January
  - Admission: Passing the Praktikum
• Slides are available shortly after the lecture
• Please send me any errors
• Slides are
  - No script
  - No substitution for listening to the lecture
  - No substitution for reading a book
Contents

• Introduction
• Overview and architecture
• Storage and access methods
  – B*-Trees, Extensible hashing, index-sequential files …
  – Multidimensional indexing: Grid-files, kd-Trees, R-Trees …
• Query processing and optimization
  – Physical operators
  – Cost-based optimization
• Recovery
• Transactions and concurrency control
5 Schichten Architektur

- Mengenorientierter Zugriff
- Logischer Zugriff
- Speicherstrukturen
- Pufferverwaltung
- Betriebssystem
- Geräteschnittstelle
- Dateischnittstelle
- Systempufferschnittstelle
- Interne Satzschnittstelle
- Satzorientierter Zugriff

- Anfrageübersetzung, Zugriffspfadwahl, Zugriffskontrolle, Integritätskontrolle
- Sortierung, Transaktionsverwaltung, Cursorverwaltung
- Record Manager, Sperrverwaltung, Log / Recovery
- Speichermanagement, Puffermanagement, Caching-Strategien
- Externspeicherverwaltung
Guests

• January: Dr. Martin Grund, HPI Potsdam: Column-based main memory database systems
• January: Dr. Mike Sips, GFZ Potsdam: Spatial index structures
Literature

• Primary

• Other
    • Deutsche Übersetzung: „Grundlagen von Datenbanksystemen“, Pearson, 2002
Datenbanken@Informatik

- Datenbanken 1: Grundlagen (BA)
- Datenbanken 2: Implementierung (MS)
- Data Warehousing und Data Mining (MS)
- Informationsintegration (MS)
  - Inkl. Verteilter Anfrageoptimierung
- Informationssysteme
- Neue Konzepte und Techniken für Datenbanksysteme
- Techniken und Konzepte zum Schutz der Privatsphäre
- Text Analytics

- Schwerpunkt im Master
WBI

- Bachelor
  - Algorithmen und Datenstrukturen (9SP)
  - Grundlagen der Bioinformatik (5SP)
  - Proseminar „Maschinelles Lernen“ (3SP)

- Master/Diplom
  - Data Warehousing und Data Mining (8SP)
  - Algorithmische Bioinformatik (9SP)
  - Informationsintegration (8SP)
  - Text Analytics (8SP)
  - Implementierung von Datenbanken (8SP)
  - Grundlagen des Semantik Web (Wandelt, 5SP)
Laufende Studien-, Bachelor-, Magister-, Master- und Diplomarbeiten

Diplom-/Magister-/Masterarbeiten

Nils Alberti: Latent Semantic Indexing zur Aufschlüsselung philosophischer Texte **(Exposé)**
Diplomarbeiten Informatik
September 2012 - März 2013
Betreuung: Ulf Leser, Anke Lüdeling

Mario Lehmann: Identifikation von Erfassungsmeldungen über Focused Crawling und Textsegmentierung **(Exposé)**
Diplomarbeiten Informatik
September 2012 - Februar 2013
Betreuung: Ulf Leser, Felix Naumann (HPI Potsdam)

Benjamin Gehrels: Graphdatenbanksysteme - Überblick und Benchmark **(Exposé)**
Diplomarbeiten Informatik
September 2012 - März 2013
Betreuung: Ulf Leser, Andre Stefan Edlich (BHT Hochschule Berlin)

Lisa Thalheim: Point mutation analysis of four human colorectal cancer exomes **(Exposé)**
Diplomarbeiten Informatik
Mai 2012 - Januar 2013
Betreuung: Ulf Leser

Masterarbeit Bioinformatik
Juni 2011 - Dezember 2012
Betreuung: Ulf Leser, Knut Reinert

Studien-/Bachelorarbeiten

Diplomarbeit Informatik
Juli 2012 - November 2012
Betreuung: Astrid Rheinländer, Ulf Leser

David Asher: Classifying abstracts of biomedical literature for stem cell research **(Exposé)**
Diplomarbeit Informatik
Juni 2012 - Oktober 2012
Betreuung: Mariana Neves, Ulf Leser

Torsten Huber: Entity Linking - A Survey of Recent Approaches **(Exposé)**
Diplomarbeit Informatik
April 2012 - Juli 2012
Betreuung: Ulf Leser, Hans-Ulrik Uecker (DFKI), Peter Adolfs (DFKI)

Marten Richter: Visualisierung im Zeitverlauf von PubMed-Veröffentlichungen zu Proteinen, Genen und Protein-Protein-Interaktionen **(Exposé)**
Diplomarbeit Informatik
Januar 2012 - Mai 2012
Betreuung: Ulf Leser, Fouzia Moussawi

Mertz Bretschneider: Focused Crawling **(Exposé)**
Diplomarbeit Informatik
Juli 2011 - Dezember 2011
Betreuung: Astrid Rheinländer, Ulf Leser
Fragen ?