Datawarehousing and Analytics

Data-Warehouse-, Data-Mining- und OLAP-Technologien

Advanced Information Management

Bernhard Mitschang, Holger Schwarz
Universität Stuttgart

Winter Term 2014/2015
Departments of Institute of Parallel and Distributed Systems (IPVS)

Applications of Parallel and Distributed Systems
Prof. B. Mitschang, Prof. M. Herschel

Machine Learning and Robotics
Prof. M. Toussaint

Parallel Systems
Prof. S. Simon

Simulation of Large Systems
Prof. M. Mehl, Jun.-Prof. D. Pflüger

Distributed Systems
Prof. K. Rothermel

Infrastructure
Dipl.-Inf. M. Matthiesen

http://www.ipvs.uni-stuttgart.de
Applications of Parallel and Distributed Systems

Data and Metadata
- Repository Technologies
- Data Warehouse and Data Mining
- Domain-specific query optimization and processing
- Product Data Management
- Data Integration

Content and Semantics
- Focussed Semantic Search
- Scalable Content Management

Information Systems and Applications
- Database Middleware
- Information Services
- Generative Application Development
- Model-driven Engineering
- Technical Information Systems

- Data in the Cloud
- Federated Systems / Application Integration
- Metadata Management
- Content Management
- Business Intelligence / Business Processes
- Query Optimization
How to contact us

- **Lecture**
  Bernhard Mitschang
  Office: 2.359
  Tel.: 0711 685 88449
  bernhard.mitschang@ipvs.uni-stuttgart.de

- **Exercises and assignments**
  Holger Schwarz
  Office: 2.361
  Tel.: 0711 685 88244
  holger.schwarz@ipvs.uni-stuttgart.de
## Planned Schedule

<table>
<thead>
<tr>
<th>Time</th>
<th>Monday 9/29/14</th>
<th>Tuesday 9/30/14</th>
<th>Wednesday 10/1/14</th>
<th>Thursday 10/2/14</th>
<th>Tuesday 10/7/14</th>
<th>Wednesday 10/8/14</th>
</tr>
</thead>
<tbody>
<tr>
<td>09:00 – 11:15</td>
<td>Chapter 1 Introduction</td>
<td>Chapter 3 Design Process</td>
<td>Chapter 4 Monitoring</td>
<td>Chapter 6 Data Mining Introduction</td>
<td>Chapter 7 SQL &amp; OLAP</td>
<td>• Materialized Summary Data • Derivability</td>
</tr>
<tr>
<td></td>
<td>Chapter 2 Data Warehouse Architecture</td>
<td>• Conceptual Design</td>
<td>• Extraction</td>
<td>• Tools</td>
<td>SQL &amp; Mining</td>
<td>Database Support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Logical Design</td>
<td>• Transformation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Break</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12:00 – 14:15</td>
<td>Chapter 2 Data Marts</td>
<td>Chapter 3 Extended Dimension Table Design</td>
<td>Chapter 5 OLAP</td>
<td>Chapter 6 Assoc. Rules</td>
<td>Chapter 7 Examples and Miscellaneous</td>
<td>• Wrap up • Intro to Assignments</td>
</tr>
<tr>
<td></td>
<td>• Operational Data Store</td>
<td>• Extended Fact Table Design</td>
<td>• Architecture</td>
<td>• Classification</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Meta Data</td>
<td>• Physical Design</td>
<td>• Storage of Data Cubes</td>
<td>• Regression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Break</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15:00 – 16:30</td>
<td>• Intro SQL and ODPS</td>
<td>• Data Warehouse Architecture</td>
<td>• Conceptual and Logical Data Warehouse Design</td>
<td>• Storage of Data Cubes</td>
<td>15:00 – 17:15</td>
<td>Data Mining • Classification • Clustering • Association Rules</td>
</tr>
<tr>
<td></td>
<td>• Issues of data integration</td>
<td>• Monitoring</td>
<td>• Design</td>
<td>• ETL</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Transformation</td>
<td></td>
<td>• Cleansing</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Lectures
- **Monday**
  - Chapter 1 Introduction
  - Chapter 2 Data Warehouse Architecture

### Exercises
- **Tuesday**
  - Chapter 3 Design Process
  - Chapter 4 Monitoring
  - Chapter 6 Data Mining Introduction

- **Wednesday**
  - Chapter 5 OLAP
  - Chapter 6 Assoc. Rules
  - Chapter 7 SQL & OLAP

- **Thursday**
  - Chapter 6 Data Mining Introduction
  - Chapter 7 SQL & Mining

- **Tuesday**
  - Chapter 7 SQL & OLAP
  - Database Support

- **Wednesday**
  - Materialized Summary Data
  - Derivability

### Breaks
- **Monday**
- **Tuesday**
- **Wednesday**
# Exercises and Assignments

<table>
<thead>
<tr>
<th>Type</th>
<th>Description</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exercise</td>
<td>Detailed discussion of major topics, case studies, examples etc. (all)</td>
<td>September 29 – October 8</td>
</tr>
<tr>
<td>Assignment</td>
<td>Hands-on training for topics related to dbms (groups of 2-3 students)</td>
<td>Introduction on October 8 Due: TBA</td>
</tr>
<tr>
<td>Assignment</td>
<td>Hands-on training for ETL, OLAP and data mining (groups of 2-3 students)</td>
<td>Introduction on October 8 Due: TBA</td>
</tr>
</tbody>
</table>
Teaching Materials

• General information:
  http://www.ipvs.uni-stuttgart.de/abteilungen/as/lehre/lehrveranstaltungen/vorlesungen/WS1415/DWDM/en

• Slides, exercises, assignments, …
  ▪ Login to ILIAS: https://ilias3.uni-stuttgart.de/login.php
  ▪ Search and Join the course "Data Warehousing and Analytics"

Repository -> Engineering -> Computer Science -> Lehrveranstaltungen WS 14/15
Exams

- IMSE
  - Exam: Friday, December 12
- Informatik / Softwaretechnik / Infotech / Wirtschaftsinformatik / …
  - Register at your examination office
  - Make an appointment for the oral exam
- Appointments for oral exams
  - Annemarie Roesler
    Tel. 0711 685 88448
    Annemarie.Roesler@ipvs.uni-stuttgart.de
Books


Papers


