Advanced Databases

Winter Term 2013/14

Prof. Dr. Dietmar Seipel
University Würzburg
**Databases and Knowledge Bases**
(Research Group at Chair for Computer Science I)

- **Data Modelling:**
  - tables and rules: relational, deductive
  - semi-structured: HTML, XML
  - ontological: Web 2.0, Semantic Web

- **Declarative Data Access:**
  - query languages
  - data mining

- **Projects:** linguistics, bio-informatics, software engineering, . . .
### Relational Database

#### COURSES

<table>
<thead>
<tr>
<th>CourseId</th>
<th>CourseName</th>
<th>Language</th>
<th>ECTS</th>
<th>Prereq</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB1</td>
<td>Databases</td>
<td>German</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DB2</td>
<td>Advanced Databases</td>
<td>English</td>
<td>3.5</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

#### SCHEDULES

<table>
<thead>
<tr>
<th>CourseId</th>
<th>Part</th>
<th>Start</th>
<th>End</th>
</tr>
</thead>
<tbody>
<tr>
<td>DB1</td>
<td>Lectures</td>
<td>14.10.2013</td>
<td>02.12.2013</td>
</tr>
<tr>
<td>DB2</td>
<td>Introduction</td>
<td>03.12.2013</td>
<td>09.12.2013</td>
</tr>
<tr>
<td>DB2</td>
<td>Lectures</td>
<td>09.12.2013</td>
<td>04.02.2014</td>
</tr>
</tbody>
</table>
XML Data (Semi–Structured)

```xml
<?xml version="1.0" encoding='ISO-8859-1'?>
<students courseid="DB2">
   <group name="Space Master"/>
   <group name="Computer Science">
      <type>Master</type>
      <type>Diploma</type>
      <type>Teaching</type>
   </group>
   <group name="Mathematics"/>
</students>
```
Project in Linguistics:
Declarative Parsing and Annotation of Electronic Dictionaries

information extraction produces XML data, which can be analyzed further; e.g., by network analysis (with comparison to genomes) or data mining
Advanced Databases

... collect / connect data from various domains / sources, such that they can be efficiently accessed and modified by many people.

Knowledge Types of Data Sources:

- XML
- Semantic Web Data (OWL, SWRL)
- Codebases

Data Mining discovers implicit / hidden knowledge.
Data Mining Workflow
Possible Applications in Space Technology

- Sensor Data for Satellites
  - Storage
  - Data Mining
- E.g., Evaluation of Weather Information
- Data Exchange using XML
Topics (planned):

DB1 – Databases
1. Introduction to Databases, ER Diagrams
2. Relational Data Modelling, Query Language SQL
3. Database Schema Design (FDs, Normal Forms)
4. Transactions

DB2 – Advanced Databases
5. Data Warehousing and Data Mining
6. Web Databases
7. XML Data Modelling, Query Language XQuery
8. Deductive and Object–Oriented Databases, Semantic Web
ER Diagram
## Relational Database

<table>
<thead>
<tr>
<th>FNAME</th>
<th>MINIT</th>
<th>LNAME</th>
<th>SSN</th>
<th>BDATE</th>
<th>ADDRESS</th>
<th>SEX</th>
<th>SALARY</th>
<th>SUPERSSN</th>
<th>DNO</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>B</td>
<td>Smith</td>
<td>4444444444</td>
<td>1955-01-09</td>
<td>731 Fondren, Houston, TX</td>
<td>M</td>
<td>30000</td>
<td>222222222</td>
<td>5</td>
</tr>
<tr>
<td>Franklin</td>
<td>T</td>
<td>Wong</td>
<td>2222222222</td>
<td>1945-12-08</td>
<td>638 Voss, Houston, TX</td>
<td>M</td>
<td>40000</td>
<td>111111111</td>
<td>5</td>
</tr>
<tr>
<td>Alicia</td>
<td>J</td>
<td>Zelaya</td>
<td>7777777777</td>
<td>1958-07-19</td>
<td>3321 Castle, Spring, TX</td>
<td>F</td>
<td>25000</td>
<td>333333333</td>
<td>4</td>
</tr>
<tr>
<td>Jennifer</td>
<td>S</td>
<td>Wallace</td>
<td>3333333333</td>
<td>1931-06-20</td>
<td>291 Berry, Bellaire, TX</td>
<td>F</td>
<td>43000</td>
<td>111111111</td>
<td>4</td>
</tr>
<tr>
<td>Ramesh</td>
<td>K</td>
<td>Narayan</td>
<td>5555555555</td>
<td>1952-09-15</td>
<td>975 Fire Oak, Humble, TX</td>
<td>M</td>
<td>38000</td>
<td>222222222</td>
<td>5</td>
</tr>
<tr>
<td>Joyce</td>
<td>A</td>
<td>English</td>
<td>6666666666</td>
<td>1962-07-31</td>
<td>5631 Rice, Houston, TX</td>
<td>F</td>
<td>25000</td>
<td>222222222</td>
<td>5</td>
</tr>
<tr>
<td>Ahmad</td>
<td>V</td>
<td>Jabbar</td>
<td>8888888888</td>
<td>1959-03-29</td>
<td>980 Dallas, Houston, TX</td>
<td>M</td>
<td>25000</td>
<td>333333333</td>
<td>4</td>
</tr>
<tr>
<td>James</td>
<td>E</td>
<td>Borg</td>
<td>1111111111</td>
<td>1927-11-10</td>
<td>450 Stone, Houston, TX</td>
<td>M</td>
<td>55000</td>
<td>NULL</td>
<td>1</td>
</tr>
</tbody>
</table>

1: Headquarters
2: Research
3: Administration
4: Research
5: Headquarters
Schedule for DB2:

- Introduction (selected topics from DB1):
  - Tuesday, 03.12., 08:30 – 10:00, Turing–Hörsaal
  - Monday, 09.12., 10:00 – 11:30, ÜR1

- Lectures: Turing–Hörsaal
  - Monday, 12:15 – 13:45, and Tuesday, 08:30 – 10:00
  - Start: Monday, 09.12., 12:15

- Labs: Theory (Monday, 10 – 12)
  + Computer Labs (1 hour per week)

- Written Exam at the end of the semester: Tuesday, 28.01.2014