Semistructured Data
Lecture with exercises (VU 2.0)

http://www.dbai.tuwien.ac.at/education/ssd/

Andreas Pieris   Wolfgang Fischl

Database and Artificial Intelligence Group
Institute for Information Systems
Vienna University of Technology

Summer term 2015
Course procedure

- **Register** to the course in TISS until 5\(^{th}\) March 2015. Adding the course to your favorites is not a registration!

- Every lecture and general instructions to the exercises are given on Thursdays starting at 9:15 a.m. (sharp) in HS8 (Heinz Parkus), Karlsplatz 13.

- The assignment discussion for the exercise is going to be in the Informatiklabor (Favoritenstraße 9–11).
Course contents

- Basics
  - XML Fundamentals
  - XML Namespaces

- Schema languages
  - DTD
  - XML Schema

- APIs
  - SAX
  - DOM

- Query languages
  - XPath
  - XQuery
  - XSLT
Textbook

XML in a Nutshell
Eliotte Rusty Harold, W. Scott Means
O’Reilly, 3rd Edition, 2004

German Edition is currently out of print, but can be bought as an eBook:
http://www.oreilly.de/catalog/pdf_xmlnut3ger/
Further Resources

- **W3C Recommendations**
  - [http://www.w3.org/TR/xml/](http://www.w3.org/TR/xml/)
  - [http://www.w3.org/TR/xml-names/](http://www.w3.org/TR/xml-names/)
  - [http://www.w3.org/TR/xmlschema-0/](http://www.w3.org/TR/xmlschema-0/)
  - [http://www.w3.org/TR/DOM/](http://www.w3.org/TR/DOM/)
  - [http://www.w3.org/TR/xpath](http://www.w3.org/TR/xpath)
  - [http://www.w3.org/TR/xpath20/](http://www.w3.org/TR/xpath20/)
  - [http://www.w3.org/TR/xslt](http://www.w3.org/TR/xslt)
  - [http://www.w3.org/TR/xquery/](http://www.w3.org/TR/xquery/)
Further Resources

**Java und XML** (German)
M. Scholz und S. Niedermeier

**XQuery** (English)
P. Walmsley

Additionally:

- [http://www.json.org/](http://www.json.org/)
- [http://www.saxproject.org/](http://www.saxproject.org/)
- [http://w3schools.com/](http://w3schools.com/)
- [http://www.cafeconleche.org/books/xmljava](http://www.cafeconleche.org/books/xmljava)
  (Online version of the book “Processing XML with Java”)
Exercises

General procedure

Three exercises are to be solved individually.

- Exercise 1: online multiple choice test
  - Exercise 1 open from 7\textsuperscript{th} March until 15\textsuperscript{th} March
  - Exercise 1 is compulsory for further participation of the course and a positive certificate.

- Exercise 2 and 3 are going to be discussed with our teaching assistants.
  - The details of the assignment will be provided on our course webpage.
  - Registration for the assignment discussion will be through our course manager.
  - A change of the date and time for the assignment discussion is possible until the day before.
  - Missing the assignment discussion means 0 points on the exercise.
Exercises
How to obtain good solutions

- Recommended strategy:
  - Discuss with your colleagues
  - But, write your own solution

- Unrecommended strategy:
  - Copy solutions from others

- For ambiguities and questions:
  - Ask in the next lecture
  - use the TISS forum
  - write an e-mail: ssd@dbai.tuwien.ac.at
Exercises

Important dates

- **Exercise 2:**
  - Due by 3/5/2015
  - Assignment discussions start on 11/5/2015 until 22/5/2015

- **Exercise 3:**
  - Due by 3/6/2015
  - Assignment discussions start on 10/6/2015 until 19/6/2015

- **Assignment discussions for Exercises 2 and 3:**
  - In the Informatiklabor at the teaching assistants
  - Bring your student identity card
  - Teaching assistants check your solutions on completeness and correctness
  - Teaching assistants also ask questions about the underlying principles
Exercises
Grading

Points:
- Exercise 1: no points, but necessary for course admission
- Exercise 2: 0–10 points
- Exercise 3: 0–15 points

Points are awarded for the written solutions and for the answers to the questions during the assignment discussion. In extreme cases one can receive 0 points for a syntactic correct solution.

Teaching assistants write a report of the discussion and make a suggestion for the number of points.
Exercises

CourseManager

- Same system as in DM and DBS
- Link is available on the course webpage or http://minteka.dbai.tuwien.ac.at/cman/
- Accounts will be created after the registration deadline in TISS for all in TISS registered students. Afterwards you will be able to receive a password through the CourseManager (or use the password from previous years).
- Username is uMATRNR, e.g.: u1234567
- Please make sure that you have added in TISS a correct eMail adress.
- The CourseManager is used for the multiple-choice test, for the upload of Exercises 2 and 3 and for the subscription to a time slot for the assignment discussion.
Exam

General remarks

- First exam date: Wednesday, 24th June 2015, 18:00–19:40
- 3 further exam dates in the winter term 2015 (announced in TISS)
- Registration for the exam in TISS.
- Admission requirements:
  - Attention: STEOP!
  - Correct registration to the course
  - this means in particular solving Exercise 1
# Exam

## Final grade

- **Necessary for a positive certificate:**
  - At least 50 (out of 100) points have to be scored in total (exercises + written exam).
  - At least 37.5 (out of 75) points have to be scored in the written exam.

- **Grading scheme (100 points reachable):**

<table>
<thead>
<tr>
<th>Points scored ((p))</th>
<th>Grade</th>
</tr>
</thead>
<tbody>
<tr>
<td>(p &lt; 50) or (&lt; 37.5) scored in the written exam</td>
<td>5</td>
</tr>
<tr>
<td>(50 \leq p \leq 61)</td>
<td>4</td>
</tr>
<tr>
<td>(61 &lt; p \leq 74)</td>
<td>3</td>
</tr>
<tr>
<td>(74 &lt; p \leq 87)</td>
<td>2</td>
</tr>
<tr>
<td>(p &gt; 87)</td>
<td>1</td>
</tr>
</tbody>
</table>
Further information

- **Homepage**
  http://www.dbai.tuwien.ac.at/education/ssd/
  please read carefully and visit regularly!
  - General informations
  - Slides and other material

- **TISS:**
  - Messages in the TISS forum (e.g.: for changes on short notice)
  - Add the course Semistructured Data to your favorites!

- **E-mail:**
  - E.g.: account information for the course manager
  - A working e-mail address has to be added to your TISS profile

- **Questions:**
  - To the contents: during/after the lecture
  - Organizational matters: send an e-mail to
    ssd@dbai.tuwien.ac.at