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 2016
Course procedure

- **Register** to the course in TISS until 17\textsuperscript{th} March 2016. 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) until 10:45 a.m. 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, JSON
  - XML Namespaces
- Schema languages
  - DTD
  - XML Schema
- APIs
  - SAX
  - DOM
- Query languages
  - XPath
  - XQuery
  - XSLT
Textbooks

- **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:

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

- **XQuery** (English)
  P. Walmsley
Further Resources

- W3C Recommendations
  - http://www.w3.org/TR/xml/
  - http://www.w3.org/TR/xml-names/
  - http://www.w3.org/TR/xmlschema-0/
  - http://www.w3.org/DOM/
  - http://www.w3.org/TR/xpath
  - http://www.w3.org/TR/xpath20/
  - http://www.w3.org/TR/xslt
  - http://www.w3.org/TR/xquery/

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

General procedure

Two exercises are to be solved individually.

- Exercises 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 TUWEL.
- 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 TUWEL forum
  - write an e-mail: ssd@dbai.tuwien.ac.at
Exercises

Important dates

- Exercise 1:
  - Due by 5/5/2016
  - Assignment discussions start on 9/5/2016 until 20/5/2016

- Exercise 2:
  - Due by 8/6/2016
  - Assignment discussions start on 10/6/2016 until 21/6/2016

- Assignment discussions:
  - In the Informatiklabor (exact location announced in TUWEL) with 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: 0–10 points
  - Exercise 2: 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

TUWEL

- For the organization of the exercises we will use TUWEL.
- Link is available on the course webpage or http://tuwel.tuwien.ac.at
- All registered students will be invited to the course after the deadline in TISS.
- Please make sure that you have added in TISS a correct eMail adress.
- TUWEL is used to Exercises and to subscribe for a time slot for the assignment discussion.
Exam

General remarks

- First exam date: Wednesday, 22\textsuperscript{nd} June 2015, 18:30–20:30
- 3 further exam dates in the winter term 2016 (announced in TISS)
- Registration for the exam in TISS.
- Admission requirements:
  - Attention: STEOP!
  - Correct registration to the course in TISS
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>

- A certificate is issued once you have uploaded a solution to one of the exercises and for every attempt at the final exam.
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 TISS (e.g.: for changes on short notice)
  - Add the course Semistructured Data to your favorites!

- Questions:
  
  - To the contents: during/after the lecture
  - TUWEL forum
  - ssd@dbai.tuwien.ac.at