XML Data Management

Course Outline and Organisation

Werner Nutt

Objectives

- Overview of
  - XML data model,
  - usage of XML
  - technologies to handle XML (generate, parse, process, store)

- Present Languages to specify the structure of XML documents (DTDs and XML Schema)

- Present Languages to query and transform XML documents (XPath, XSLT, XQuery)

- Introduce tools that implement these languages (parsers, editors, validity checkers, query engines, etc.)

Students will …

- Try out the tools in the labs

- Solve exercises
  - write DTDs and XML schemas
  - query documents with XPath
  - transform documents with XSLT and XQuery

- Build small applications that process XML data, e.g.,
  - find differences between documents
  - write a Web application using XQuery

(Coursework!)

Course Content (1)

- Purpose and usage of XML

- XML Syntax:
  - XML markup rules
  - well-formed XML documents

- Specifying the structure of documents: DTDs

- More expressive schemas: XML schema

- Extracting data from documents: XPath

- Querying and constructing XML with XQ uery
.
- Transforming documents with XSLT

Course Content (2)

- Parsing and generating XML with a programming language:
    - DOM and SAX parsing in Java

Possibly:

- XML and relational databases:
    - PostgreSQL support for XML

- Native XML databases
    - BaseX, eXist

Student Input

I have given the course already twice, but

- I would like to try out new things
    - ➔ the course is not yet 100% prepared

If you have ideas, proposals for further

- topics
- activities,

let me know

I will see whether we can accommodate them

Course Format

- Lectures
  - introduce new concepts, give examples

- Labs
  - exercises (➔preparation for exam questions)
  - support for projects

- Coursework projects
  - Handling XML in Java:
    define and develop your version of xdiff
  - Web application in XQuery (?)

Books

Textbooks

- *Learning XML*, 2nd Edition By Erik T. Ray
- *XQuery,* By Priscilla Walmsley

Both books are available as
electronic editions at the library
website (need login and passwd)
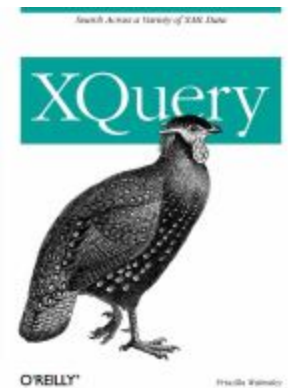
Additional reading:

- *Einstieg in XML: Grundlagen, Praxis, Referenz,*
  By Helmut Vonhoegen

Hardcover book, available in the library

Course Organization

- Lectures: Mon 8:30-10:30, SER-E 412

- Labs: Wed 14:00-16:00, SER-E 412 (alternating weeks)

- Office hours: Mo 14:00-15:00, POS 2.09

- Course web site will develop at

    `www.inf.unibz.it/~nutt/XMLDM1314`

    Will contain
    - slides
    - lab instructions
    - example code
    - pointers to literature and software

Assessment

Based on

- written exam

- coursework (cw)

Final mark = max { 30% x cw + 70% x exam,

exam}

This means:

- The final mark cannot get worse by the coursework

- You can improve the exam mark with the coursework

Example:

- exam = 20, cw = 29 $\Rightarrow$ final = 23

Final Exam

Will test technical skills, e.g.,

- write DTDs and/or XML schemas
- write XPath expressions
- write queries in XQuery
- write XSLT transformations

(see past exam on course page)

I may add questions that test general understanding, e.g.,

- role of XML in an application
- shortcomings of DTDs vs XML Schema
- how to build an application with XQuery