

XML and Databases

(CS 345b)

Daniela Florescu (dflorescu@mac.com)

Donald Kossmann (donaldk@inf.ethz.ch)

Organization

- Instructors
 - Daniela Florescu: dflorescu@mac.com
 - Donald Kossmann: donaldk@inf.ethz.ch
- TA
 - Anish Das Sarma: anishds@stanford.edu
- Office Hours
 - Mondays, Wednesdays: 3 - 4 pm, Gates #414 (tentative)
- Structure + Grades
 - Classes
 - Homework: 2-3 assignments (20%)
 - Project (50%)
 - Exams (30%): midterm (Feb. 14), final (March 14)
- Course Web Page:
<http://www.stanford.edu/class/cs345b>

Goals

- **Overview of XML Technologies**
 - „Impress your boss“
 - Follow standards, products, recent developments
- **Use XML for projects (e.g., Mashups)**
 - Advantages, disadvantages of XML technologies
 - Typical application scenarios
 - Evaluate products (critically)
- **Own developments**
 - Packaging of Technologies, Abstractions
 - Compensate for missing functionality

References

- <http://w3c.org>
 - Most general XML standards
- <http://www.oasis-open.org>
 - More specific, industrial standards
- <http://www.informatik.uni-trier.de/~ley/db>
 - Database of research literature
- Product information from vendors

Tools

- **Oxygen, XML Spy, Stylus Studio**
 - XML Editor, Schema Editor, Query Editor
- **Apache Xerxes and family**
 - XML Parser, XML Schema validator, XSLT Engine
- **MXQuery, Saxon, ...**
 - XQuery Engines
 - (complete list at W3C Web site)

Schedule

1/10	intro & motivation
1/17 - 1/22	basics: namespaces, InfoSet, schema, ...
1/24	Web Services, REST, Mashups
1/29 - 2/12	XQuery, XSLT, Updates, XQueryP
2/14	midterm exam
2/19	RSS, Atom, pub/sub
2/21	IR, XQuery FullText
2/26	Data Integration (external presenter)
2/28 - 3/7	Implementation, XQuery Engines
3/12	Semantic Web
3/14	final exam

Overview (ctd.)

- **M4: XML and relational databases**
 - Shredding: Storing XML in tables
 - SQL / XML Standard
 - Products (IBM, Microsoft, Oracle)
- **M5: Implementation: Storage + Processing**
- **M6: Web Services**
 - SOAP, WSDL, UDDI
 - BPEL and XL
- **M7: Information Filtering and RSS**
- **M8: Semantic Web**
- **M9: Outlook**

Project

- Carry out in groups of 2-3 students
- Make a proposal after two weeks (Jan. 28)
- Mashups
 - Phase 1: build an interesting Web Service
 - Phase 2: build a Mashup with other Web Services
- Suggestions for other projects (if you are eager)
 - XQuery benchmark
 - Testing compliance of XQuery engines
 - Run XQuery on a PS3 (or so)
 - RSS Aggregator in XQuery
 - extend XQuery engine with cool new functionality
 - ... (ask us for more suggestions)