ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

XML and Databases

Peter Fischer

Web Science

http://websci.informatik.uni-freiburg.de

Some words on myself

- Since October 1st:
  Assistant/Junior Professor for Web Science
- Before Senior Researcher (“Oberassistent”) at Systems Group, ETH Zurich
- Research Interests:
  - XML/Web Technologies
  - Data Stream/Event Processing
- Systems-oriented approach

Goals of the lecture

- Overview of XML Technologies
  - „Man kann mitreden“
  - Follow standards, products, recent developments
- Understand what and why is different
  - Why isn‘t Java + SQL all you need?
  - How does this all relate to the NoSQL hype?

References

- http://w3c.org/XML/
  - 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 Editor
  - XML Editor, Schema Editor, Query Editor
- Apache Xerxes and family (open source)
  - XML Parser, XSLT Engine
- Saxon (open source + commercial)
  - XQuery Engine + XSLT Engine
- Micro XQuery/Zorba (research projects at ETH)
  - XQuery Engine + extensions
  - Eclipse plugin: editor, debugging
- Sausalito: Zorba in the Cloud
  - www.28msec.com

Overview

- M1: Motivation
- M2: XML Basics
  - XML, Infoset, Namespaces, XML Schema
- M3: XML Processing – Models and Languages
  - Data Models for XML
  - XPath / XQuery / XSLT
  - XQuery Update Facility
  - XQuery Scripting
  - Full Text

Overview (ctd.)

- M4: Implementation: Storage + Processing
  - Algebra/Execution Model
  - Query Rewrite
  - Data Representation+Storage
  - Indexing
  - Case Studies
- M5: SQL/XML
- M6: TBD – one out of
  - Web Services
  - XML Streaming/Event Processing
  - Cloud computing

Lecture+Exercises

- Lecture
  - Get concepts
- Exercise groups
  - Hands on experience with tools + languages
  - Deeper understanding of an XQuery engine
- Project
  - Implement a small project based on XQuery+extensions
  - 2 students per group
  - Mandatory for participating on exam

Schedule

- Exercise sheets will be handed out a week before they will discussed in the groups
- Exercise groups will start in week 2 (4.11.) (sheets this week, discussion next week)

- Project due early February
  - About 1 SWS dedicated to project