NB! Double-click on text with the quotation "_" and replace it with your own text. Don't remove or change any headers please. Leave the course code! It's not available.

FDA138 XML and Databases (CIS)

Lectures:

10 h (preferably, 5 sessions of 2 hours each).

Recommended for

Graduate students.

The course was last given:

New course.

Goals

This course aims at introducing techniques and methods related to XML and databases that have been developed during the last years, some of which are still the subject of active research.

Prerequisites

Common knowledge of basic notions in Computer Science. No knowledge of XML will be assumed. No specific knowledge in document management and/or in database systems is assumed.

Organization

5 sessions during week 34 and 35.

Contents

XML is emerging as the markup language of choice for data modeling and data interchange on the Web. Furthermore, many aspect of data modeling and data management on the Web are related to database systems. Indeed, the Web can be seen as a distributed information system and it provides databases with a convenient interface. This course aims at introducing to techniques and methods related to XML and databases that have been developed during the last years, some of which are still the subject of active research. The course will give an introduction to XML basics, to formalisms for specifying XML data schemas, to query and transformation languages for XML, and to indexing methods for XML data. The course will also present research results on some of these issues recently obtained at the University of Munich. No specific knowledge in document management and/or in database systems is assumed.

Course Summary

1. XML Basics

- 1.1 Markup Languages: Origins and Typology
- 1.2 Structure of an XML Document
 - 1.2.1 Document Prolog
 - 1.2.2 Elements and Attributes
 - 1.2.3 Entities and Notations
 - 1.2.4 Namespaces
 - 1.2.5 Character Sets
 - 1.2.6 Document Tree
- 1.3 XML vs. SGML
- 1.4 XML vs. HTML
- 1.5 References

- 2. XML, Databases, and Data Schemas for XML
 - 2.1 Features of Standard Data Models
 - 2.2 Advantages of Dispensable Data Schemas
 - 2.3 Semistructured Data
 - 2.4 DTD
 - 2.5 XML Schema
 - 2.7 References
- 3. Query and Transformation Languages for XML
 - 3.1 Need for Transformations
 - 3.2 Data Selection with XPath and XPointer
 - 3.3 Styling with CSS
 - 3.4 The Transformation Language XSLT
 - 3.5 The Query Language XQuery
 - 3.6 Xcerpt: Querying XML Data Reconsidered
 - 3.7 Streamed Evaluation of XPath
 - **3.8 References**

4. Indexing XML Data

4.1 Basics: Tag and/or Keyword Indexing

4.2 Path Indexing

4.3 Navigation Indices: Data Guides, Signature Files, and Improvements Thereof

4.4 References

Literature

Teaching material.

Teachers

François Bry, visiting professor.

Examiner

Jan Maluszynski.

Schedule

Fall 2002.

Examination

Written report.

Credit

2 credits.

Comments

Intensive course.