AIT 321 WEB DATA MANAGEMENT 3(2+1)

Objective(s)

After learning the course, the student will be able to understand the overall vision of the Semantic Web, to analyze the current technology stack (URIs, XML, RDF/S, OWL), to understand how one could use these technologies for building something useful, to define and test an ontology, to define schema mappings and to install and use tools for semantic data management

UNIT I

Data Model

Introduction to Modeling Web Data, Semi structured data, XML, Web Data Management with XML, XML Standards, XML and syntax, XML Data Model, XLink, and XPointer

UNIT II

XPath and XQuery

Introduction, Basics of XPath and XQuery, XPath: Steps and path expressions, Evaluation of path expressions, Generalities on axes and node tests, Axes, Node tests and abbreviations, Predicates, XPath 2.0; FLWOR expressions in XQuery: Defining variables -the for and let clauses, Filtering - the where clause, The return clause, Advanced features of XQuery; XPath foundations.

UNIT III

Typing

Motivating Typing**,** Automata, Schema Languages for XML, Typing Graph Data: Graph Semistructured Data, Graph Bisimulation, Data guides.

XML Query Evaluation

XML fragmentation, XML identifiers: Region-based identifiers, Dewey-based identifiers, Structural identifiers and updates; XML evaluation techniques: Structural join, Optimizing structural join queries, Holistic twig joins.

UNIT IV

Ontologies, RDF, and OWL

Introduction, Ontologies by example, Web resources, URI, namespaces, RDF, RDFS: RDF Schema, OWL, Ontologies and (Description) Logics.

Querying Data through Ontologies

Introduction, Querying RDF data: notation and semantics, Querying through RDFS ontologies, Answering queries through DL-LITE ontologies.

UNIT V

Data Integration

Introduction, Containment of conjunctive queries, Global as view mediation, Local as view mediation, Ontology-based mediators, Peer-to-Peer Data Management Systems.

Building Web scale applications

Web search, web crawlers, web information retrieval, Web graph mining and hot topics in web search, Distributed systems, failure management, Required properties of a distributed

system, P2P networks, Hash-based structures, distributed indexing, Distributed computing with Map Reduce

Reference Book(s)

1. Serge Abiteboul, Ioana Manolescu, Philippe Rigaux, Marie-Christine Rousset and Pierre Senellart, "Web Data Management", Cambridge University Press, 2011
2. Bhavani Thuraisingham, "Web Data Management and Electronic Commerce", CRC Press, 2000
3. Bhavani Thuraisingham, "XML Databases and the Semantic Web", CRC Press, 2002
4. Athena Vakali and George Pallis, "Web Data Management Practices: Emerging Techniques and Technologies", IGI Publishing, 2007, ISBN-10: 1599042282; ISBN-13: 978-1599042282

Practical(s)

1. Create an XML file defining an article in newspaper.
2. Create an XML file containing list of students. Also create stylesheet file to display list in an HTML format.
3. Create an XML file containing list of students. Using XPath display following information
    a. Information of a student with ID No : 101
    b. All the student in the sorted order according to their CGPA
4. Create an XForm to collect information from staff member regarding their publications. Details like Year of Publication, National/International, Title, Conference/Journal etc.
5. From the above gathered information, using XQuery find out the number of publication in a specific year.
6. Demonstrate the use of AJAX.
7. Study of XMLSPY tool.
8. Create an RSS for the events occurring in your institute.
9. Write a program to read the articles in RSS created in above practical.
10. Study of RDF (Resource Description Framework)