GUJARAT TECHNOLOGICAL UNIVERSITY
MASTER OF COMPUTER APPLICATIONS (MCA)
Semester: V

Subject Name: **Web Data Management (WDM)**
             (Elective-II)

Subject Code**: 650009**   [w.e.f. June 2012]

Course Description

XML and, more generally, semistructured data, differs in shape from traditional relational data. More importantly, it is used in novel kinds of applications: exchange and retrieval on the Web, rather than local storage and processing. This course covers some major developments in semistructured data and XML that happened over the last few years: data models, syntax, query languages, schemas, query analysis, type-checking, publishing, indexes, and storage methods.

Prerequisite: Database Programming, Database Management Systems

Course Contents

Topics	Hours
Introduction: Modeling Web Data • Database Technology and Web Applications • Semistructured data • Web Data Management with XML • XML and syntax • XML Data Model, XLink, and XPointer	7
XPath and XQuery • Regular Path Expressions • XPath Basics • XPath steps and expressions • Path evaluations, axes, node tests, predicates • XQuery Syntax • FLWOR expression, advanced features • XUpdate	7
Typing • Automata on ranked trees, unranked trees • XML Schema, other schema languages • Graph semistructured data, graph bisimulation, data guides • XML query evaluation • XML identifiers, XML evaluation techniques	8
Ontologies, Querying and Data Integration • RDF, RDF Schema, OWL • Description Logic • Querying data through ontologies • Querying RDF data, querying through RDFS • Answering queries though DL	8

• Global-as-view (GAV) and Local-as-view (LAV) mediation • Ontology based mediation, 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 MapReduce	7

Text Book

S. Abiteboul, I. Manolescu, P. Rigaux, M. Rousset and P. Senellart, Web Data Management, Cambridge University Press, 2012

Reference Books

1. S. Abiteboul, P. Buneman and D. Suciu, Data on the Web: From Relational to Semistructured Data to XML, Morgan Kaufman Publisher
2. Research papers and W3C web site