ISSN No. 0976-5697

Volume 2, No. 4, July-August 2011

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

Unified Structure for Exchange of University Information in Indian Context

Miss Nehal Daulatjada
Lecturer, SEMCOM College, Vallabh Vidyanagar
nehaldaulatjada@yahoo.com

Ms. Swaminarayan Priya R
Head & Asso. Prof., MCA, ISTAR
Vallabh Vidyanagar
Priya_swaminarayan@yahoo.co.in

Dr.P V Virparia*
Associate Professor, GDCST, Vallabh Vidyanagar
pvvirparia@yahoo.com

Dr.V R Rathod
Dean & Faculty, Dept. of Computer Science, Marwadi, Education Foundation's Group of Inst. Rajkot

Abstract: The Indian Scenario in information exchange is far outdated. The information transfer takes place in either physical form or in an electronic form via Email. The information about the universities is registered with the University Grants Commission (UGC). The collection of information of the approx. 500 universities by UGC periodically and updating in their database is a very tedious task. The authors have taken an in-depth view of the current scenario of exchanging information by the universities about its details and statistics. This research paper is an attempt to standardize the exchange of university relevant information through the use of Extensible Markup Language (XML) Schema.

Keywords: XML Schema, Indian Universities, UGC, Standardized structure, Instance document.

I. INTRODUCTION

Currently, Web content is formatted for human readers rather than programs. HTML is the predominant language in which Web pages are written for people, the information is presented in a satisfactory way, but machines will have their problems. Keyword-based searches and an intelligent agent might be able to identify the information but may not be able to distinguish the information. The Semantic Web approach to solve these problems is not the development of superintelligent agents, but if HTML is replaced by more appropriate languages, then the Web pages could carry their content on their sleeve. In addition to containing formatting information aimed at producing a document for human readers, they could contain information about their content [1].

XML is a W3C (World Wide Web Consortium) standard for document markup. Both Document Type Definitions (DTD) and XML schemas are mechanisms used to define the structure of XML documents. They determine what elements can be contained within the XML document, how they are to be used, what default values their attributes can have, and so on. Given a DTD or XML schema and its corresponding XML document, a parser can validate whether the document conforms to the desired structure and constraints. This is particularly useful in data exchange scenarios as DTDs and XML schemas provide and enforce a common vocabulary for the data to be exchanged. XML schemas differ from DTDs in that the XML schema definition language is based on XML itself. As a result, unlike DTDs, the set of constructs available for defining an XML document is extensible. In addition, stronger typing constraints on the data enclosed by a tag can be described because a range of primitive data types such as string, decimal, and integer are supported. This makes XML schemas highly suitable for defining data-centric documents [2].

An XML schema language is a formalization of the constraints, expressed as rules or a model of structure, that apply to a class of XML documents [3]. A schema expressed in W3C XML Schema syntax that describes the permitted content. XML Schema is the W3C-recommended schema definition language, expressed in XML 1.0 syntax, which is intended to describe the structure and constrain the content of documents written in XML. It is explicitly intended to improve on the schema functionality that was provided by the DTD, which was the original form of schema for XML documents that the W3C recommended in 1998 when XML was first released [4].

DTDs impose several restrictions with respect to expressivity. This means that only very simple languages can be defined by means of DTDs alone. For instance, one cannot define restrictions on how often a certain element may appear by only regular expressions; moreover, data types are hardly supported; etc [5].

These limitations are overcome by another mechanism for defining XML grammars, XML Schema. Beyond DTDs, XML schema allows advanced features such as the following:

A. Support for a basic set of data types (numbers, strings, and dates, etc.), which can be restricted further;
B. Definition of one's own element or attribute types, available for reuse via an inheritance mechanism which allows extension/restriction;
C. Namespace support;
D. XML Schema is an XML language itself, allowing developers to profit from tool support, the inherent extensibility of XML, the combination of several XML Schema files, etc.

II. PRESENT SCENARIO OF INFORMATION EXCHANGE

The information must be updated with the UGC database whenever demanded or whenever the data changes. This information is transferred in paper format or through electronic formats which may be created using proprietary software which leads to inter-operability issues or in an unorganized format which requires manual efforts. The information so collected at the other end needs to be integrated and sometimes manually entered in a uniform structure which is time-consuming and error-prone. The information gathered from different sources comes in different formats / fonts / structures. Sometimes insufficient details lead to incomplete record entries. Summarizing the collected information takes time and report generation is difficult.

Providing a central server would be one alternative to this problem. But if the Network connection is unavailable, this may not work. Another alternative is to create the XML Schema instance file on the client side in the specified format, by imposing restrictions on the input and using it for data transfer. Additionally this file may also be used on the University's website for a web page generation using XML Stylesheet Language Transformation (XSLT) which reduces the burden of creating a webpage to represent the same or some of the information from this instance document. At any time, the data file may be modified and the changes would be reflected on the webpage using that instance document.

III. USING XML SCHEMA FOR UNIVERSITY INFORMATION EXCHANGE

Looking at this problem that is faced by many universities in managing their information needs, it is quite obvious that XML schema can, to a great extent, help in sharing the information. XML forms the base of many information representation and exchange technologies. As developers create new XML vocabularies, they often need to describe those vocabularies to share, define, and apply them, which we will, using XML Schema. Hence our attempt is to provide a very simple mechanism by which the universities are able to store their information as well as exchange it.

IV. STRUCTURE DESCRIPTION

For our Standardization, we have used the most common and obvious of information that a university may need to communicate [6,7,8]. Following is a list of the details and its element name used in our structure:

A. University Registration Number :RegistrationNo
B. University Name: UniversityName
C. University Address: UniversityAddress
D. University Establishment Year: UniversityEstablishmentYear
E. 5. University Type: UniversityType (State, Deemed, Government, Technical, Central, Affiliated, State Private)
F. Vice-Chancellor Name: VCName
G. University Motto: UniversityMotto
H. NAAC Accreditation Grade:NAACGrade (A,B,C,D, {empty})
I. Number of Masters Courses: MastersCourses
J. Number of Under-Graduate Courses: UGCouses
K. 11. Number of Departments: NumberOfDepartments
L. 12. Total University Staff: UniversityStaff
M. 13. Total Department Staff: DepartmentStaff
N. 14. Number of Research Projects: ResearchProjects
O. 15. Number of Research Students(PhD): PhdStudents
P. 16. Number of Research Students(MPhil): MphilStudents
Q. 17. Total revenue through Research Projects: ResearchRevenue
R. 18. Total International Students: InternationalStudents
S. 19. Total National Students: NationalStudents
T. 20. Number of students in Masters Courses: MastersStudents
U. 21. Number of students in Under-Graduate Courses: UGStudents
V. 22. Website Address: WebAddress
W. 23. Chancellor Name: CName
X. 24. Registrar Name: RName
Y. 25. Jurisdiction: Jurisdiction
Z. 26. Girls Colleges: GirlsColleges
AA. 27. Recognised Research Institutes: ResearchInstitutes
BB. 28. Total Department Teacher Staff: DepartmentTeachers
CC. 29. Total College Teacher Staff: CollegeTeachers
DD. 30. Library University exists: LibraryExist
EE. 31. UGC Recognition: UGCRecognition
FF. 32. UGC Funded: UGCFunded
GG. 33. Membership : Membership

V. XML SCHEMA STRUCTURE [LISTING 1]

Based on the above details collected from various sources, the following structure was derived. It has been tested using XSV on the W3C XML Schema Validator site[9]. This XML Schema can be used in Local-site specific application as the data structure and an application can be developed which will ensure the storage of inputs using this structure and result into an XML output.

```
<?xml version='1.0' encoding='UTF-8'?>
<schema xmlns="http://www.w3.org/2001/XMLSchema">
<element name="University">
<complexType>
<sequence>
<element name="RegistrationNo" type="string" />
<element name="UniversityName" type="string" />
<element name="UniversityAddress" type="string" />
<element name="UnivEstablishmentYear"
type="UnivEstablishmentYear" />
<element name="UniversityType" type="UniversityType"
/>
<element name="VCName" type="string" />
<element name="UniversityMotto" type="string" />
<element name="NAACGrade" type="NAACGrade" />
<element name="MastersCourses" type="NonZero" />
<element name="UGCourses" type="NonZero" />
<element name="NumberOfDepartments" type="NonZero"
```

```
/>
<element name="UniversityStaff" type="NonZero" />
<element name="DepartmentStaff" type="NonZero" />
<element name="ResearchProjects" type="ZeroMore" />
<element name="PhDStudents" type="ZeroMore" />
<element name="MphilStudents" type="ZeroMore" />
<element name="ResearchRevenue" type="ZeroMore" />
<element name="InternationalStudents" type="ZeroMore"
/>
<element name="NationalStudents" type="NonZero" />
<element name="MastersStudents" type="NonZero" />
<element name="UGStudents" type="NonZero" />
<element name="WebAddress" type="anyURI" />
<element name="CName" type="string" />
<element name="RName" type="string" />
<element name="Jurisdiction" type="string" />
<element name="GirlsColleges" type="ZeroMore" />
<element name="ResearchInstitutes" type="ZeroMore" />
<element name="DepartmentTeachers" type="ZeroMore" />
<element name="CollegeTeachers" type="ZeroMore" />
<element name="LibraryExist" type="YesNo" />
<element name="UGCRecognition" type="YesNo" />
<element name="UGCFunded" type="YesNo" />
<element name="Membership" type="string" />
</sequence>
</complexType>
</element>
<simpleType name="UniversityEstablishmentYear">
<restriction base="gYear">
<minInclusive value="1900" />
</restriction>
</simpleType>
<simpleType name="UniversityType">
<restriction base="string">
<enumeration value="State" />
<enumeration value="Deemed" />
<enumeration value="Technical" />
<enumeration value="Central" />
<enumeration value="State Private" />
<enumeration value=""/>
</restriction>
</simpleType>
<simpleType name="NAACGrade">
<restriction base="string">
<enumeration value="A" />
<enumeration value="A+" />
<enumeration value="B" />
<enumeration value="B+" />
<enumeration value="C" />
<enumeration value="C+" />
<enumeration value="D" />
<enumeration value="" />
</restriction>
</simpleType>
<simpleType name="NonZero">
<restriction base="decimal">
<minExclusive value="0" />
</restriction>
</simpleType>
<simpleType name="ZeroMore">
<restriction base="decimal">
<minInclusive value="0" />
</restriction>
</simpleType>
<simpleType name="YesNo">
<restriction base="string">
<enumeration value="Yes" />
<enumeration value="No" />
</restriction>
</simpleType>
</schema>
```

VI. ABOUT THE SCHEMA

In our Schema, we are using predefined data types like string, any URI, gYear and decimal as well as user-defined datatypes like University Establishment Year, University Type, NAAC Grade, Non Zero, Zero More and Yes No. The Schema assumes that, there are five types of universities i.e. State, Deemed, Central, Technical, and State Private. The NAAC Grades are taken as A+, A,B+,B,C+,C and D.

The XML Schema structure so depicted can be used effectively for exchanging information. Additional information, specific to any university can also be incorporated. This addition will not affect the information exchange as long as the specified basic element names are not modified or duplicated, otherwise it will result in an error. The individual universities will have the freedom to use software as per their requirement or availability but will ensure the data will be stored in an XML file using the given element names only, thus standardizing the formats across all universities.

VII. CONCLUSION

XML has been used to ensure Platform-neutral data transfer across the World Wide Web. Our structure follows the XML Schema recommendations and will ensure standard structure of information representation by all the universities in the Indian context.

VIII. REFERENCES

[1] Grigorsi Antoniou and Frank van Harmelen, The MIT Press, 2004, “A Semantic Web Primer”

[2] Chaudhri.A.B, Rashid.A, Zicari.R, XML Data Management: Native XML and XML-Enabled Database Systems, Addison Wesley, 2003

[3] Van Der Vlist.E, XML Schema, O'Reilly, 2002

[4] Wyke.A, Watt.A, XML Schema Essentials, Wiley Computer Publishing, 2002

[5] Dieter Fensel · Holger Lausen · Axel Polleres Jos de Bruijn · Michael Stollberg · Dumitru Roman John Domingue, Enabling Semantic Web Services - The Web Service Modeliing Ontology Springer, 2007

[6] Inflibnet Web site: http://www.inflibnet.ac.in/universitydirectory/ accessed on 7-Nov-2010

[7] Indian education portal: http://www.indiastudycenter.com/Univ/Gujarat-Universities.asp accessed on 7-Nov-2010.

[8] Sardar Patel University Website : http://www.spuvvn.edu/glance.html accessed on 8-Nov-2010

[9] World Wide Web Consortium official website : http://www.w3.org/2001/03/webdata/xsv accessed on 9-12 Nov,2010

AUTHOR'S PROFILE

Miss Nehal Daulatjada, Lecturer, SEMCOM College, Vallabh Vidyanagar. She is born in Surat, Gujarat-India, 15th November'74. She has done her graduation in Statistics (Gold med.) in 1995 and later pursued her Masters in Computer Application in 1998, both from Sardar Patel University, Vallabh Vidyanagar – India. Currently she is pursuing her Doctoral studies in the area of Web Ontology. Her areas of interest include System Analysis, Database Design, Web technologies and Ontology Development.

Ms. Swaminarayan Priya R. is working as a Head & Associate Professor in MCA department of Institute of Science & Technology for Advanced Studies and Research, Vallabh Vidyanagar. She has done her graduation in Mathematics in 1994 and later pursued her Masters in Computer Application in 1997, both from Sardar Patel University, Vallabh Vidyanagar – India. Currently she is pursuing her Doctoral studies in the area of web ontology. Her publications include 5 papers in International Journal, 5 papers in National Journals and 20 papers in International/National conferences/seminars. Her research interests include the areas of Semantic Web and Ontology Development.

Dr. Paresh Virparia is working as a Senior Associate Professor in the Department of Computer Science of Sardar Patel University, Vallabh Vidyanagar. He completed his MCA in 1989 from Sardar Patel University and Ph. D. in 2002 from Sardar Patel University. He is recognized Ph.D. guide in Computer Science at Sardar Patel University, V V Nagar and Kadi Vishva Vidyalaya, Gandhinagar. Currently, NINE students are pursuing their Ph. D. under the guidance of him. His publications include 8 papers in International Journal, 6 papers in National Journals and 30 papers in national conferences/seminars. His research interests include the areas of Computer Simulation & Modeling, Networking and IT enabled services.

Dr. V.R.Rathod, Dean and Faculty of Computer Applications, Marwadi Education Foundations group of Institutions, Rajkot. He has more than 42 years of professional and research experience in the field of Computer Science. He has presented more than 60 research papers in National and International conferences. He has published more than 20 research papers in National and International journals.