International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 2, February 2014)

Role of RDF in Semantic Web

Pallavi Grover[1], Dr.Sonal Chawla[2]
[1]*DAV College, Chandigarh.*
[2]*DCSA, Panjab University.*

***Abstract*— Information on web exists in decentralized way. Semantic Web is seen as the evolving extension of web which will hold information in machine understandable for. This paper has threefold objective. Firstly the paper talks about RDF model helps in representation of distributed information on web. It then focuses on RDF Model and Schema. Secondly the paper throws light on RDF structure, use of XML for representing information in distributed environment and how the two are complementary to each other. It also discusses the available formats of RDF. Lastly it highlights how RDF provides a mechanism to authenticate the information available on web.**

***Keywords*—RDF, RDF Schema, URI, RDF Reification, Semantic Web, XML.**

I. Introduction

The World Wide Web is a collection of distributed interlinked hypertext information across a network. This information is in the form of text, images, videos and other multimedia. It was originally built for human consumption, and although everything on it is machine-readable, this data is not machine-understandable. Since web contains voluminous information it is not possible to manage its automation manually. Use of meta-data is proposed to be solution for this. Metadata is data about data or in this context data describing web resources. Resource Description Framework, domain independent data model that enables interoperability between applications that exchange machine-compatible information on the internet. [1]

II. Resource Description Framework

RDF provides a foundation for processing metadata. It is a general method that decomposed knowledge into small pieces, using some rules about the semantics of those pieces. The idea is to eventually have a method so simple that it can express any fact in a structured manner that computer applications can do useful things with knowledge expressed in RDF. This eventually provides interoperability between such applications which exchange machine-understandable information on web.

III. Basic RDF Model And Schema

The RDF data model provides an abstract, conceptual framework for defining and using metadata. RDF is a data model based on a graph that uses URIs (Uniform Resource Identifier) for representing metadata. RDF model provides syntax used for encoding and transporting metadata such that it maximizes interoperability of independently developed web servers and clients.[2] Broadly we can say that RDF model defines mechanism for describing resources without making any assumptions about a particular domain or define semantics of any application beforehand. In other words the definition of mechanism should be neutral of domain and yet the mechanism should be suitable for describing information about any domain. [2] To facilitate the definition of metadata has class system, much like the object oriented programming, and collection of classes is referred to as Schema. Schemas are hieratically organized, generated with a specific purpose or domain and can be extended through subclass refinement. So when a schema has to be produced, which is similar to existing schema only incremental modifications are required to be done. It is because of this incremental extensibility that the agents processing metadata of unfamiliar origin can trace to schema of familiar origin to perform meaningful actions on metadata.

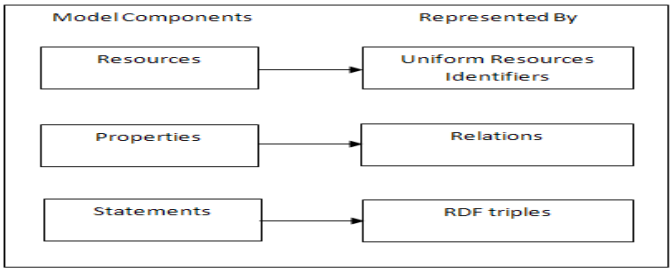

Figure I.Reperesentation of RDF Model

A. RDF Structure

As mentioned earlier RDF provides foundation for processing metadata [3]. This data about data is stored in RDF model in a particular format representing named properties and property values.

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 2, February 2014)

RDF properties can be thought of as attributes of resources or relationship between resources. For convenience it can be said that it bears resemblance to entity relationship diagram.

Following are the major parts RDF model

- *Resources:* Everything being described by RDF expressions is called resource. [5] A resource may vary from being an entire web page or certain part of web page or XML element within the document source or might be a whole collection of pages; e.g. an entire web site. A resource might also be an object that is not directly accessible via the Web. For example A table. Anything can have a URI. Resources are always named by URIs plus optional anchor IDs [5].
- *Properties:* A property is a characteristic, attribute, a relation or some other specific aspect which is used to describe the resource. Each property has a specific meaning. Basically it provides some sort of description to resources which can also be in terms of relationship between resources or other properties.
- *Statements:* A combination of specific resource together with a named property plus the value of that property for that resource is called an RDF statement. A statement has three different parts which are described as under:

➢ *A subject:* Start node of the edge.
➢ *A predicate:* Describes a relation between the subject and object as a verb or type of edge or property.
➢ *An object:* End node of the edge.

RDF Model using RDF triples		
Subject <Start node >	Predicate < Relation between Subject and Object >	Object <End Node>

Figure II. Representation of RDF triples

In terms of object oriented approach, we can conveniently say that properties are analogous to instance variable and resources to objects of a class. The RDF data model is a syntax-neutral way of representing RDF expressions [5]. We can use the data model representation to evaluate equivalence in RDF expressions. Two RDF expressions are said to be equivalent if and only if their data model representations are the same and do not alter the meaning of the expressions in context. Here syntactic variations are permitted in an expression.

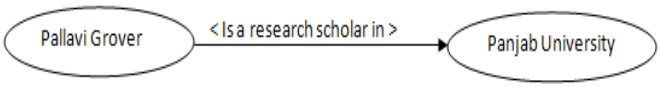

Figure IV. Graphical Representation of an example using RDF triple

Pallavi Grover	< is a research scholar in >	Panjab University
Subject	Predicate	Object

Figure V. Tabular Representation of an example using RDF triple

B. RDF Format

This specification of RDF uses the Extensible Mark-up Language [XML] encoding as its interchange syntax. RDF/XML format is a w3c standard and is based on XML.RDF also requires the XML namespace facility to precisely associate each property with the schema that defines the property. Before further elaborating RDF format let us understand XML with respect to both HTML and RDF.

XML

XML stands for Extensible Markup Language. It defines a set of rules for encoding documents on web, in a format that is both human-readable and machine-readable. It is defined in the XML 1.0 Specification produced by theW3C [2].It is also a marked up language like HTML but also varies from it in certain ways. The basic purpose of XML to carry data and not to display data as incase of HTML. XML and HTML were designed with different goals. With focus on what data is, XML was designed to transport and store data whereas how the data looks is on web was the focal point of HTML's design. XML is a complement to HTML and not its replacement in any way. Lastly HTML uses some predefined tags whereas XML uses user-defined tags. As written earlier XML defines rules for encoding documents in a format that is both human and machine readable. Also since XML allows users to define their own tags, it is this feature which largely contributes to its letdown when it needs to be used Semantic web because it lacks consistency.

Example: Let this be the statement which has to be parsed into a machine understandable language, "www.google.com was created by Larry Page and Sergey Brin"

XML code for this statement could be:

International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347 - 6435 (Online) Volume 2, Issue 2, February 2014)

```
<author>
<uri>www.google.com</uri>
<name>Larry Pane and Sergey Brin</name>
</author>
```

Figure VI. XML code for representing information

Alternatively it could be,

```
<document href=" www.google.com ">
<author> Larry Pane and Sergey Brin </author>
</document>
```

Figure VII. Alternate XML code for representing information

Clearly these XML codes lack consistency and the user defined tags become a hindrance here. When a machine will parse the above two XML codes, different XML trees will be produced. There can be more ways in which the above code could have been coded. This means there exist a large number of ways in which XML maps onto logical tree. But eventually our requirement is to convert a query into a single statement no matter how it occurs in the code. This is exactly what RDF does. It provides a standard way of writing statement such that whatever occurs in the document, the effect produced or the meaning derived is the same.

www.google.com	Created by	Larry Page and Sergey Brin
Resource	Property	Statement

Figure VIII. Representation of information through RDF

IV. Use Of RDF In Binary Relationships

Until now we have discussed at length RDF structure, format, role and purpose of existence in Semantic Web Architecture. It is evident from the discussion that RDF, by default, stores information in the form of triplets. Now suppose there exists a scenario where incomplete information is available i.e., one out of those three nodes is missing. Or precisely only binary relations exist. In this situation both the storage and query could have been affected. But RDF provides a concept of using blank nodes in these situations. These blank nodes or the anonymous nodes can be used to represent more complex relations. These nameless nodes can exist because of varied reasons ranging from author lacking knowledge about the document to not willing to disclose or provide a name to a node.

Whatever might be the case, RDF gives an option that these anonymous nodes be given a local identifier so that they may be referred multiple times within the same document. Since these are local identifiers, they don't have any meaning beyond the scope of the document.

Local_Identifier	Predicate	Object
Subject	Local_Identifier	Object
Subject	Predicate	Local_Identifier

Figure IX.RDF storage and query format for binary relationships

V. RDF Reification

Reification generally refers to making something real, bringing something into being, or making something concrete. [5] In terms of computer science we can say reification is constructing a data model for a previously abstract concept. Plethora of information is accessible due to Internet explosion which has bought with it certain issues, trustworthiness standing tall among them. The importance of provenance information as a means to trust and validate the authenticity of available data cannot be stressed enough in today's web-enabled world. [6] By Provenance information it s meant that who should be held responsible for the data or how the data came into being. This assists in the process of verifying the authenticity of the data. Semantic web technologies such as Resource Description Framework (RDF) include the ability to record such provenance information through the process of reification. Reification is an important RDF concept that provides the ability to make assertions about statements represented by RDF Triples [7]. It not only helps associate trust with RDF triplets but also helps in ranking of authenticity of the triples

VI. Conclusion

Semantic Web leaves its impact on information technology and different research areas such as Knowledge Engineering / Management, Software Agents and Web Services. An important objective of the Semantic Web is to hand over most of the information to software agents that we are doing by ourselves now a day on the web. The paper discussed a few essential components of semantic web; Extensive Mark-up Language (XML) provides a surface syntax for structured documents, XML schema is a language for restricting the structure and content element of the XML document. Resource Distribution Framework (RDF) is a simple data model for referring objects and how they are related; RDF schema is a vocabulary for describing properties and classes of RDF resources.

VII. FUTURE SCOPE

There is variety of challenges in representing information on the web and there are increased demands and expectations for making information available both for humans and software agents. Humans being want to express information in natural language whereas formal representation is required by software agents. A satisfactory solution to web information representation requires minimizing human effort while still satisfying the software agents' demands.RDF has been recognised as one of those but has certain limitations associated with it. Support for many way relationships

REFERENCES

[1] W3C Recommendation, RDF Primer (2004), http://www.w3.org/TR/rdf-primer/

[2] http://www.w3schools.com/xml/xml_whatis.asp

[3] Brickley and R. Guha, "Resource Description Framework (RDF) Schema Specification," W3C Candidate Recom-mendation, Mar. 2000; available at http://www.w3.org/TR/2000/CR-RDF-schema-20000327.

[4] http://en.wikipedia.org/wiki/XML

[5] http://www.w3.org/TR/PR-rdf-syntax/

[6] http://www.w3.org/TR/1998/WD-rdf-syntax-19981008/

[7] http://en.wikipedia.org/wiki/Reification S.Ramanujam. A.Gupta, L.Khan, S.Seida, B.Thuraisingham , " A relational Wrapper for RDF Reification" http://link.springer.com/chapter/10.1007%2F978-3-642-02056-8_13

[8] O.Harting, "Provence information in the wed of data"

[9] http://www.w3.org/TR/rdf-primer/

[10] Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems

[11] Fröhlich, B. and Plate, J. 2000. The cubic mouse: a new device for three-dimensional input. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems