

# Semantic Web Technology in Digital Libraries: An Overview

Ashwini. P, Kavitha. S  
Dept. of library and information Science  
Annamalai University  
Annamalai Nagar  
Tamil Nadu, India

**ABSTRACT:** *Semantic web technology expands the usefulness of the digital libraries that will contain majority of data in future. This paper outlines the emerging concept of social semantic digital library and explores the potential of using semantic web technology for the digital libraries, we can also see that an overview of semantic tools and services available and different types of semantic tools and services.*

**Keywords:** Semantic Web, Semantic Digital Library, Digital Library

**DOI:** 10.6025/ijis/2019/11/4/113-117

**Received:** 6 April 2019, Revised 19 June 2019, Accepted 3 July 2019

© 2019 DLINE. All Rights Reserved

## 1. Introduction

The invention and advent of modern technology had lead a great change in the lives of mankind; it leaped to another level of lifestyle and culture, subsequently become a valuable invention and which has made internet accessible and more convenient to its user such as inventions formed the web. English scientist Tim Berners- Lee invented the World Wide Web in 1989. He wrote the first computer program in 1990 while employed at CERN in Switzerland. His vision was to design internet based hypermedia system which would facilitate the uploading of information from all corners of the globe so that everyone could benefit from that knowledge. The later developments in the web witnessed the evolution of semantic web, which was similarly derived from the version of Tim Berners-Lee in the sense of making information global and universal for its users. The current web contains billions of documents and has many administrative problems and limitation In addition to that the web content will only to accessible to humans. The solution to that problem is the Semantic Web. The semantic web is considered to be the extension for the current web.

Digital libraries provide high quality and well organized information but this data is often not connected with other pools on information such as personal homepage of authors.

The semantic web initiative of the World Wide Web (W3C) has been active for the last few years and has attracted user's interests. The main task of the semantic web is to add a meaning to the traditional web. In order to achieve that objective, layers of representational structure are needed. A digital library is a library in which collections are stored in digital formats and accessible by computers. In classical libraries interfaces the user mainly searches for title, author and keywords. Within an electronic interface this is done with one text entry field. Results are presented in a short form which could be extended to a longer description of the resource, which is available electronically or has at least the full description of the resource online. Semantic digital library is a digital library system that applies semantic technologies to provide machine support for the end user in their search for information. Users in semantic digital library would be able to log in and create

profiles that would translate to FOAF files and user ontologies to keep track of semantic preferences. The search and retrieval function would benefit from the ability to search semantically along with the traditional search methods the library currently employs. Semantic web technology essentially teaches the computer the relationship between the digital objects, their meaning and how users would like to interact with them.

## 2. Semantic Web Solutions

Resource description framework, Web Ontology Language and Extensible Markup Language (XML) will enable to providing machine readable descriptions which append meaning to the content on the web documents. In Semantic Web the machine performs automated information gathering and analysis similar to that of human beings which results in retrieval of more relevant and meaningful results. Tim Berners-Lee has called this linked data network as Giant Global Graph.

## 3. Components

The term semantic web is used to refer to the technologies and standards used for structuring and linking of data by providing a proper description of concepts, terms and their associations within a given knowledge domain. Such standards and technologies include under W3C are;

- a) Resource Description Framework (RDF)
- b) RDF Schema (RDFS)
- c) Simple Knowledge Organization System (SKOS)
- d) SPARQL, which is a RDF query language
- e) Notation3 (N3)
- f) N-Triples, is format for storing and transmitting data
- g) Web Ontology Language (OWL)

## 4. Semantic Web tools in Digital Libraries

Semantic tools are designed to support semantic functionalities, which are;

- a) **RDF conversion/ visualization tools:** Used for documenting and mapping.
- b) **Thesaurus/ KOS tools:** Used for representing and sharing knowledge organization system over the web.
- c) **Metadata schema standards:** Helps in transforming a flat metadata schema to a semantic web ontology.
- d) **Fully flashed project/ separate portal:** It incorporate features of semantic search and browse as well as follows semantic architectural models for storage of information.
- e) **Open Source Java Software/ Program/ Searching tool:** It provides the standard terminology to searching.
- f) **Knowledge Extraction tools:** Helps in creation of knowledge from structured and unstructured source.
- g) **Ontology Engineering Tool:** Any tool used for creating ontology or similar semantic documents.
- i) **Semantic Measures Tool:** Used for computation and analysis of semantic tools There are many semantic applications developed in recent past which can be classified under different semantic tools listed above.

## 5. Semantic Webs Stack & Semantic Technology

The challenges in the current web are in the areas of integration, standardization, development of tools, and adoption by users. The following definitions are given by W3C consortium.

**a) HTTP:** It is an application protocol for distributed, collaborative, hypermedia information system to exchange or transfer hyper text.

**b) HTML:** It is the standard inextensible and rigid markup language derived from SGML for displaying web pages and other information in a web browser.

**c) eXtensible Markup Language:** XML is the extension of HTML, the popularity technology used to provide fundamental structure and syntax for the content in the web documents. XML allows for extensible data formats unlike HTML which is inextensible and rigid but does not provide semantic constraints.

**d) XML Schema:** It is a language describes the content structure in terms of constraints in an XML document.

**e) Metadata:** It is a general framework which allows providing machine understandable information.

**f) Logic:** It is the study of the principles of reasoning. Well understood formal semantics for expressing knowledge are provided by the Logic.

**g) Agents:** The software objects developed by using object oriented programming and component based software that work autonomously, proactively and intelligently are called agents.

**h) Uniform Resource Identifier:** URI is a unique name given to identify a resource over a network using specific protocols.

**i) Resource Description Framework:** RDF is a simple language used to create standard data models to refer resources, their relationship and data interchange on the web.

**j) Web Ontologies:** The term Ontology is taken from philosophy which means the study of the nature of existence. It is the branch of metaphysics which identifies and describes the things in the most general terms.

**k) SPARQL:** It is a protocol and query language for semantic web data sources. SPARQL can be used to express queries across diverse data sources, whether the data is stored natively as RDF or viewed as RDF via middle wear.

## 6. Development of a Semantic Web Library

The Semantic Web has emerged to address the shortcomings of HTML webpages by developing IT tools which are machine driven and required for integrated access across heterogeneous resources. The explicit meanings are given to the information which enables the machine to process without human intervention and put together it. The vision, goals, and mission of both the libraries and the semantic web are similar. Both of these have been developed for accessing information available in abundance and discovering the knowledge through cooperation and collaboration for the advancement of the society. The applicability of library functions for developing semantic library is discussed in the following sections;

### 6.1 Library Portals

Library Portals provides a gateway to information, services from multiple sources and access to the organizations resources. The use of Semantic Web Technologies in developing library portal facilitates users search access, and retrieval of learning resources. The implementation of library portals with semantic web services will fulfill the vision of libraries. It facilitated automated functions and task delegation to intelligent agents. The Semantic library portals should have automated interactions with a search engine at the resource combined with web ontology and the content is tagged with information

### 6.2 Semantic Web Selection

Effective sharing and reuse of selected high quality data is necessary for providing best services. The linked data will help in

select and sharing these complex metadata document resources. The semantic web selection faces challenges in identifying resources which are distributed across a number of heterogeneous information stores. The non- authoritative, inconsistent metadata in these heterogeneous resources, metadata with different vocabularies having similarities different fields of competency, semi structured repositories are the major challenges in selection. The library catalogue contains structured content which can be made available to the semantic web applications. The Semantic Web focuses on the mechanisms of describing the resources and making them available to the user.

### **6.3 Semantic Web Representation**

These are variety of Metadata schemes which guide in cataloging of digital resources i.e., metadata creation. The description given to the resource should enable a user to identify and select appropriate resources with respect to content, format, etc. The resources have to be represented by simple description, following formal ontology which are machine support and have deductive reasoning, by following functional requirements for bibliographic records and cataloging criteria.

### **6.4 The Semantic Web Resource use**

The arrival of digital libraries have eliminated many challenges of the circulation section such as lending the limited collection, defining the loan periods and renewal policies, issue of lost and damaged items etc.

## **7. RDF and Digital Libraries**

The digital libraries require active information access facilities and to acquire trusted and reliable information

### **7.1 Simple Knowledge Organization System(SKOS)**

It is an application of RDF which provides a data model for Knowledge Organization System (KOS) being developed by W3C. It is used to improve the recall, retrieval, precision and provides a number of searching methods for users.

### **7.2 Friend of a Friend (FOAF)**

FOAF is an application of RDF, a machine readable ontology which describes people. It provides a rich vocabulary to describe personal information. Each person will be identified by a unique identifier and this is used by the computer to locate people or groups with similar interests and allows for social relationship within group.

### **7.3 DBpedia**

DBpedia is a linked data project as part of the Wikipedia project by Tim Berners-Lee which is aimed at extracting the structured content from the information that structured information is then made available on the WWW?

## **8. Challenges and Threats**

The main challenges of semantic web are; privacy, censorship, double entry, network centric, and its cost in implementation, the vastness of the information on the web: overlapping and precise concepts which is leading vagueness : logical contradictions leading to ambiguity, variations and deception. The challenges to “unify logic” and “proof” layers of semantic web will require Web Ontology Language (OWL) and this is an area of active research. There was a hard criticism that developing of semantic web is time consuming has they have to create two formats for both human and machine. Micro formats have solved this problem partially.

## **9. Conclusion**

Traditional libraries have taken the shape of an interactive, accessible and efficient platform which is present for the user at any time of the day. The new forms of digital libraries that are semantic digital libraries have proved to produce more meaningful results for the user. The semantically intelligent integrated library system will provide effective functioning in the provisory on of library service. The development of semantic web and library function can be refined for making true impact on the society.

## **References**

[1] Berners-Lee, T. (1999). *Weaving the Web: The original Design and Ultimate Destiny of the World Wide Web by its Inventor*, San Francisco; Harper.

- [2] Berners-Lee, T., Hendler, J., Lassila, O. (2001). The Semantic Web. *Scientific American*, 284(5), Also available online <http://www.sciam.com/article.cfm?articleID=00048144-10D2-1C70-84A9809EC588EF21>.
- [3] WWW consortium W3C Website: <http://www.w3.org>.
- [4] McGuinness, D. L. (2003). Ontologies Come of Age. In Dieter, F., Hendler, J., Lieberman, H. and Wahlster, W. (Eds.) *Spinning the Semantic Web: Bringing the World Wide Web to its Full Potential*. Cambridge: MIT Press, 2003, p. 171-194.
- [5] P Majumdar – 2007, SEMANTIC WEB : THE FUTURE OF WWW, [http://ir.inflibnet.ac.in/dxml/bitstream/handle/1944/584/426-441\(cal%2007\).pdf?sequence=1](http://ir.inflibnet.ac.in/dxml/bitstream/handle/1944/584/426-441(cal%2007).pdf?sequence=1).
- [6] IFLA Digital Reference Guidelines. (2005). : <http://www.ifla.org/VII/s36/pubs/drg03.htm> International Federation of Library Associations and Institutions. (1963). Report: international conference on cataloguing principles, Paris, 9<sup>th</sup> -18<sup>th</sup> October, 1961. London: *Organizing Committee of the International Conference on Cataloguing Principles*, p. 91-96.
- [7] World Wide Web Consortium (2008). Representing Knowledge in the Semantic Web, <http://www.w3c.it/talks/2005/openCulture/slide7-0.html>.
- [8] Kruk, S., R., Cygan, M., Piotrowski, P., Samp, K., Westerski, A., Decker, S. (2006). *Building a Heterogeneous Network of Digital Libraries on the Semantic Web*, In: *Proceedings of the Semantics Conference*.
- [9] Metz, C. (2007). Web 3.0, PCMag.com, <http://www.pcmag.com/article2/0,2817,2102852,00.asp>.
- [10] Lynch, C. (2005). Where Do We Go From Here? The Next Decade for Digital Libraries, *D-Lib Magazine*, July/August 2005, 11 (7/8).