International Journal of Advance and Applied Research
www.ijaar.co.in
ISSN – 2347-7075 Impact Factor – 8.141
Peer Reviewed Bi-Monthly
Vol. 6 No. 18 March - April - 2025

Building A Semantic Web Technology For A Digital Library And It's Challenges

Miss. Pratiksha U. Deshmukh[1] & Prof. Dr Sanjay M. Salwe[2]
[1]*Research Scholar, Library and Information Science,*
Sant Gadge Baba Amravati University,
[2]*Ass. Prof.& Librarian, Head Ph.D. Research Center*
Shri Ramrao Sarnaik College of Social Work, Washim
Corresponding Author – Miss. Pratiksha U. Deshmukh
DOI - 10.5281/zenodo.14784834

Abstract:
This paper discusses about the concept of semantic web technologies and it's challenges. The Semantic Web has attracted a diverse, but significant, community of researchers, institutes and companies, all sharing the belief that one day the Semantic Web will have as big an impact on life as currently the WWW/Internet has. In this paper, we identify some of the major challenges the community faces in the coming years, and we outline solution directions. The major challenges concern: (i) the availability of content, (ii) ontology availability, development and evolution, (iii) scalability, (iv) multilinguality, (v) visualization to reduce information overload, and (vi) stability of Semantic Web languages. We will also say some words on the economic impact of the Semantic Web.

Origin of Semantic Web:

Many cognitive scientists have worked earlier on how to structure the knowledge semantically and enable the automated agents to access the web more intelligently and perform the work of the users on their behalf. Descriptive technologies such as XML, RDF, and OWL have been developed to address the limitations in using HTML. XML (Extensible Markup Language) provides a method for transmitting structured documents. It do not impose any semantic constraints or meaning on the data it carries. RDF (Resource Descriptive Framework) is a simple framework / data model to refer the content in the object. RDF is often represented in XML format. RSS is an RDF object. The concept of Semantic Web was first coined by Tim Berners-Lee,[1] who had also developed Hyper Text Markup Language (HTML), Hyper Text Transfer Protocol (HTTP), Uniform Resource Identifiers (URI) and World Wide Web (WWW). He visualized Semantic Web as a platform where the intelligent software agents will analyze a particular given situation and present with the best possible alternatives to the users. Tim Berners-Lee has shown how the Semantic Web would work technically and explained about ontologies and as well as their importance in constructing the Semantic Web Companion Web site.[2] He told that Semantic Web will act as an integrator across different applications and content in publishing, blogging and other areas, information applications and systems.

Some Definition of Semantic Web:

- Semantic web is an expansion of the current World Wide Web, the Semantic Web is a concept of a system that provides software

programmers with machine-interpretable metadata about the information and that has been published on web.

- 'The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users.

What Is The Semantic Web?

The word 'Semantics' has been derived from Greek word 'smantiká' (neuter plural of smantikós) which means the study of meaning. The study focuses on the relation between signifiers, such as words, phrases, signs, and symbols, and what they stand for, their denotation. Linguistic semantics deals with the study of meaning that is used to understand human expression through language. Other forms of semantics include the semantics of programming languages, formal logics, and semiotics. The terms semantics, metadata and ontologies are used synonymously to refer to Semantic Web. [7] The Semantic Web provides a common framework that allows data to be shared and reused across applications, enterprise, and community boundaries. It is a collaborative effort led by W3C [3] with participation from a large number of researchers and industrial partners. Its objective is to convert all the unstructured documents on the web into a web data. It is based on the Resource Description Framework (RDF). [47] "a web of data that can be processed directly and indirectly by machines." [7] W3C looks after the development of such Semantic Web standards1 . In their Semantic Web Activity Page W3C states: "the idea of having data on the Web defined and linked in such a way that it can be used by machines not just for display purposes, but for automation, integration and reuse of data across various applications."

Goals of Semantic Web:

The primary goal of the Semantic Web is to realize the full potential of the Web, making it cost effective for people, effectively record the knowledge by giving maximum impetus on machine consumption by designing and adopting the technologies which support such machine facilitated global knowledge exchange. "The Semantic Web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation."

Semantic Web Solutions:

Resource Description Framework, Web Ontology Language and Extensible Markup Language (XML) will enable in providing machine-readable descriptions which append meaning to the content on the web documents. In Semantic Web the machines performs automated information gathering and analysis similar to that of human beings which results in retrieval of more relevant & meaningful results. Tim Berners-Lee has called this linked data network as the Giant Global Graph [1]. 7. 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 included under W3C [7] 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 a format for storing and transmitting data g) Turtle (Terse RDF Triple Language) h) Web Ontology Language (OWL)

Semantic Web Resource Use:

The circulation policy always promotes healthy use of the collection and protects the library holdings. The arrival of digital libraries have eliminated many challenges of the circulation section such as lending the limited collection, defining loan periods and renewal policies, issue of lost and damaged items, fragile &rare materials etc. There should be Semantic Web resource use policy in the library which promote resources use, and protect the integrity of resources. The policy should clearly mention the access procedures for agents; provide them with a unique identification number, and borrowing privileges of resources. The information with regard to the availability of new resources both internally and externally shall be provided to the agents regularly. The Centralised information agents provide a cooperative approach to data sharing.

Artificial Intelligence and Expert Systems:

Semantic Web can be fully realized with the use of artificial intelligence. It is the study and design of intelligent agents which perceives its environment and takes actions that maximize its chances of success. John McCarthy, has coined this term in 1955, and defined it as "the science and engineering of making intelligent machines". An expert system is a computer system that has the ability to make decisions like a human expert. These are particularly designed and developed to solve complex problems by reasoning about knowledge, like human expert, and not by computer algorithms and procedures.

Development of a Semantic Library:

The Semantic Web comes in handy for the Librarians in providing effective library services. Using the experiences and knowledge of the Librarians the appropriate metadata can be embedded into the existing collections. As the Libraries are information gatekeepers they should bring information and people together. Semantic Web is a remarkable tool for Libraries where it protects proprietary information, and helps in sharing the wealth of knowledge. The Semantic Web has emerged to address the shortcomings of HTML web pages 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 machines to process without human intervention and put together it. "The Semantic Web is not a separate web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation" 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 society.

Library Portals:

The Library portals provide a gateway to information, services from multiple sources and access to the organization's resources. The use of Semantic Web technologies in developing Library portals facilitates users' search, access, and retrieval of learning resources. The portal should aim to provide access to a coalition of learning repositories with learning resources available in different formats. The implementation of Library portals with Semantic Web services will fulfill the vision of Libraries. International Journal of Library Science. The large collections of learning resources are semantically annotated adopting various technologies that facilitate user's access to the content in one or more learning repositories. Ontologies are used for annotating information to the web content

and expressing its semantics in a machine-readable manner. The Ontology schema will be able to give more flexibility in providing semantic description to the content in learning object repositories and, at the same time, it facilitates automated functions and task delegation to intelligent agents. The library portals search interface should have the capabilities for searching across the heterogeneous resources. The Semantic Library portal should have automated interaction with a search engine at the resource, combined with web ontologies, and the content is tagged with information. The adoption and implementation of technologies will enable ontology-facilitated sharing and reuse of learning resources. Such a portal will allow the library to provide best services.

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 approaches for selecting the documents are a) Taking semantic similarity as metadata analysis to support the user comparison for resources b) Development of huge amount of ontology driven metadata describing complex features as linked data 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 in different fields of competency, semi-structured repositories are the major challenges in selection. The distinct decentralized repositories provide more structured semantic metadata about the learning resources. The search for the resources in WWW or such repositories can be done using ontology driven which extends the combinations of the search terms which are similar semantically. The quality data which is of high quality can be obtained by identifying the resources using explicit metadata statements, compared with annotations of co-occurrences, then ranked and selected. The structure of ontology heavily relies on the measure of relatedness and match. 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. The collection development policies have to be reviewed and revised periodically as per the present requirements and demands. The collection developed should have the characteristics of semantic web. The web resources have to be identified, selected and information has to be semantically encoded with the defined type of format(s) that will be tagged. The people who are involved in the collection development and metadata creation should be guided in their selection and tagging of resources. A Semantic Web Selection Policy has to be drafted carefully. [40]

There are still several problems to solve before making this happen, including, but not limited to:

- **The availability of content**. Currently, there is little Semantic Web content available. Existing web content should be upgraded to Semantic Web content including static HTML pages, existing XML content, and dynamic content, multimedia and web services.
- **Ontology availability, development and evolution.** Ontologies will become a key piece, as they allow expliciting the semantics of Semantic Web content. A big effort must be made in the creation of common widely used ontologies for the Semantic Web, on the provision of adequate infrastructure for ontology development, change management and mapping, and, in this distributed web

environment, on the adequate control of the evolution of ontologies and the annotations referring to them.

- **Scalability**. A significant effort must be made to organize Semantic Web content, store it and provide the necessary mechanisms to find it. All these tasks must be performed and coordinated in a scalable manner, as these solutions should be prepared for the huge growth of the Semantic Web
- **Multilinguality**. This problem already exists in the current Web, and should also be tackled in the Semantic Web. Any Semantic Web approach should provide facilities to access information in several languages, allowing the creation and access to SW content independently of the native language of content providers and users.
- **Visualization**. Intuitive visualization of Semantic Web content will become more and more important to solve the increasing amount of information overload, as users will demand the easy recognition of relevant content for their purposes. New techniques must be explored that differ from the usual hypertext structure visualization of the current web.
- **Stability of Semantic Web languages**. Finally, standardization efforts must be performed urgently in this emerging field, in order to allow the creation of the necessary technology that supports the Semantic Web

In this paper, we will explain those challenges1 and provide some solution directions. But before that, we will sketch some relevant notions that have emerged in the short history of the Semantic Web. At the end of the paper we will also say something about possible commercial applications exploiting the Semantic Web.

Challenges:

The primary challenges of the Semantic Web are Privacy, Censorship, Double Entry/ Increased Workloads, Network-centric, and it's Cost in implementation. The other challenges for the Semantic Web include the vastness of the information on the web; overlapping and imprecise concepts which is leading to vagueness; logical contradictions leading to ambiguity, variation, and deception. Automated reasoning systems will have to deal with all of these issues in order to achieve the goals of the Semantic Web. The challenges to the "unifying logic" and "proof" layers of the Semantic Web will require extensions to the Web ontology language (OWL) and this is an area of active research. The major possible threat is from the Vendors as they are incapable and reluctant to make necessary changes due to their failure in coordinating, apathy and indifference attitude towards the new technology. The other possible threats are particular to libraries concerned. The libraries have to identify the strategies to deal with the vendors and they need to begin adopting standards-based APIs, such as XML-based web service layers.

Conclusion:

This article explored the possibilities of performing library functions and providing Library Services using Semantic Web Technologies. An inquiry into the system indicated that many of the Library functions and services can be applied to the Semantic Web. And In this paper we identified what in our opinion are some of the main challenges the Semantic Web faces in the coming years. The major challenges include: (i) the availability of content, (ii) evolving ontologies, (iii) scalability, (iv) multilinguality, (v) visualization to reduce information overload, (vi) stability of Semantic Web languages.

The librarians should acquire all the latest IT skills useful in developing and maintaining digital libraries, and develop communication skills in different languages and ontologies for better dissemination of

information and provision of services and to reach the largest clientele. Even though the libraries have been adopting the latest IT technologies, they still have to embrace the Semantic Web. The Librarians who have skills, talent and knowledge are the advocates of Semantic Web and they can fulfill the vision and mission of the Semantic Web. The primary challenges of the Semantic Web are Privacy, Censorship, Double Entry/ Increased Workloads, Network-centric, and it's Cost in implementation. The other challenges for the Semantic Web include the vastness of the information on the web; overlapping and imprecise concepts which is leading to vagueness; logical contradictions leading to ambiguity, variation, and deception. The research should be done to tackle these challenges and threats.

References:

1. Semantic web, http://en.wikipedia.org/wiki/Semantic_Web.
2. Brooks, T.A. (2202), The Semantic Web, Universalist Ambition and Some Lessons from Librarianship. Information Research, 7(4): http://informationr.net/ir/7-4/paper136.html
3. Gerber, AJ, Barnard, A & Van der Merwe, Alta (2006), A Semantic Web Status Model, Integrated Design & Process Technology, Special Issue: IDPT 2006
4. 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, 9th - 18th October, 1961. London: Organizing Committee of the International Conference on Cataloguing Principles, p. 91-96.
5. Shadbolt, Nigel; Berners-Lee, Tim; Hall, Wendy (May/June 2006). "The Semantic Web Revisited". IEEE Intelligent Systems 21 (3): pp. 96–101.
6. Berners-Lee, T. (1999), Weaving the Web: The original Design and Ultimate Destiny of the World Wide Web by its Inventor, San Francisco; Harper.
7. Greenberg, J. and Robertson. D. W. (2002). Semantic Web Construction: An Inquiry of Authors' Views on Collaborative Metadata Generation. DC-2002: Metadata for e-Communities: Supporting Diversity and Convergence. Proceedings of the International Conference on Dublin Core and Metadata for e-Communities, 2002, Florence, Italy. October 13-17. Firenze University Press (ISBN: 88-843-043-1), pp. 45-52. Also available at: http://www.bncf.net/dc2002/ program/ft/paper5.pdf.
8. 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
9. Grigoris Antoniou, and Frank Van Harmelen - 2004 - A Semantic Web Primer - page 7-8
10. Greenberg, J. Advancing the Semantic Web via Library Functions, http://ils.unc.edu/mrc/pdf/greenberg07advancing.pdf
11. Kruk, S.R. Building Semantic Web Services Infrastructure for Digital Libraries,http://marcont.corrib.org/marcont/pdf/eswc2005_swsl2l.pdf
12. Brooks, T. A. (2002). The Semantic Web, Universalist Ambition and Some Lessons from Librarianship.

Information Research, 7(4): http://informationr.net/ir/7-4/paper136.html

13. I O'Murchu, J. Breslin, and S. Decker. Community portal survey. DERI Research Report - Semantic Web Portal Project D9, DERI.Galway, April 2004.
14. C. Bizer and T. Gauss. Rdf book mashup - serving rdf descriptions of your books,
    a. Nov. 2006.
15. Benjamins R., Decker S., Fensel D., Gómez-Pérez A. "(KA)2: Building Ontologies for the Internet": A Mid Term Report. International Journal of Human Computer Studies (IJHCS). 51(3). September 1999. 687-712.
16. Brickley, D., Guha, R.V. Resource Description Framework (RDF) Schema Specification. W3C Proposed Recommendation. March, 2000. http://www.w3.org/TR/rdf-schema/