Study of "Semantic Web" For Finding Relevant Information on Web

[1]Ms. Amrata Soni, [2]Ms. Jyoti Kukade
[1]Assistant Professor, [2]Assistant Professor
[1]Department of Computer Science,
[1]Mittal Institute of Technology, Bhopal, India

ABSTRACT— The main idea of this paper is to show current state of the Semantic Web concept. The large growth in the volume of data and with the extreme growth of web pages, today's search engines are not suitable anymore. Search engine is the most important tool to find any information in World Wide Web. Most existing Web search engines only generate a result list matching the user's query accurately. Based on extremely large and dynamic source of information, both normal and professional users obtain relevant information from information retrieval system with respect to the query that expresses the user's needs. Semantic Search Engine is born of traditional search engine to overcome the above problem. The Semantic Web is added advance features of the current web in which information is given well-defined meaning. Every page in semantic web contains semantic metadata that holds additional information's about the Web page. In this paper we made a brief survey on various assure features of some of the best semantic search engines and we have discussed the various approaches of semantic search. We have summarized the techniques, advantages of semantic web search engines that make it more efficient for making it more suitable search engine.

IndexTerms— RDF, OWL, Semantic web retrieval, Domain ontology.

I.   INTRODUCTION

The rapid growth of the Internet there is a requirement for a user to have an easy way of accessing information and services. In today's scenario on web huge semi structured database is present.

That provides data with redundancies & the unused amount of information. Due to this, information overload, we are facing new challenges for not only placing relevant information on precisely, but also accessing a variety of information from different resources directly. Efficient searching is required to get high quality results which are based on relevant matching between well-defined resources and user queries. Information retrieval technology has been the basis of the Web. When users use search engines to search for specific information, with the help of advance techniques the quality of such content will be improved. Most of the traditional search engines get the syntactically correct but have huge amount of information. Semantic Web technologies can significantly improve the information distribution process by overcoming the troubles of current web portals. In this sense, portals based on Semantic Web technologies represent the next generation Search engine.

The Semantic show this information to be accurately described in terms of well defined vocabularies. A semantic search engine gives selected results which the user is searching for. The main objective of Semantic Web is to make Web content understandable not only by humans, but also machine understandable. The Semantic Web is a novel approach for systematize information and it symbolize a large concern area for the global research community. We need to ensure that semantics are not lost during the whole life cycle of information retrieval. Today various semantic search engines present so far differ from each other through the outcome obtained & what the technologies concerned, which can be discussed in this paper. We investigate the evolution of web search engines & use of semantic web technology to generate relevant information.

II.   TRADITIONAL SEARCH ENGINE & ITS LIMITATIONS

Present World Wide Web is containing the global database that lacks the existence of a semantic methodology and hence it creates the difficulties for the machine to understand the information provided by the user in the form of search strings. As for results, the search engines return the ambiguous or partially ambiguous result data set; Semantic web is being to be developed to overcome the following problems for current web. The limitations are as follow:

1. The representation of information over the web content lacks of proper structure.
2. Ambiguity presented in outcome dataset resulting poor interconnection of information.
3. Unable to manage with huge number of users and content ensuring trust at all levels.
4. Incapability of machines to understand provided information due to lack of a universal format.
5. Automatic information flow is lacking.

Hakia [7] is a general purpose semantic search engine that search structured text like Wikipedia. Hakia calls itself a "meaning-based (semantic) search engine" [8]. They're trying to provide search results based on meaning match, rather than by the popularity of search terms. The presented news, Blogs, Credible, and galleries are processed by hakia's proprietary core semantic technology called QDEXing [7]. It can process any kind of digital artifact by its Semantic Rank technology using third party API feeds [9].

III.   SEMANTIC WEB DEFINITION

The most referred definition of the Semantic Web is given in the following:

"*The Semantic Web is an extension of the current webin which information is given well-defined meaning, better enabling computers and people to work in cooperation.*" (Tim Berners-Lee et al.) [20]

The main concept of the Semantic Web, proposed by Tim Berners-Lee, is to increase present data on the Web with machine interpretable metadata to allow better automation, integration, discovery and reuse across various application& web.

"*The Semantic Web is a web of data, in some ways like a global database.*" and "*Leaving aside the artificial intelligence problem of training machines to behave like people, the Semantic Web approach instead develops languages for expressing information in a machine process able form.*" (Tim Berners-Lee)

There is some confusion about the term "Semantic Web":

"*The fact that the programmer and the interpreter of the computer output use the symbols to stand for objects in the world is totally beyond the scope of the computer. The computer, to repeat, has syntax but no semantics.*"(John Searle)

"*Developing XML as a richer version of HTML was generally a good idea. But what botched the Semantic Web is that promoting a universal syntax does nothing to promote semantics. To avoid further confusion, it would be a good idea to rename it the syntactic web.*" (John Sowa)

We do not go into logical discussion about perfection of term "Semantic Web", but rather suggest that the use of the Semantic Web approach does not depend on the label we use for it. Tim Berbers-Lee has noted that the semantic in Semantic Web means machine process able.

IV.   COMPONENTS OF SEMANTIC WEB

The Semantic Web Layer, which describes its components and their relationships, is shown in Figure 1 (copied from [27]). Brief descriptions of each layer in the Semantic Web are as follows:

1. **URI and Unicode: In Semantic Web** Each object is identified by a unique URI (Uniform Resource Identifier) [19] assign to it. There are different types of URIs, such as Universal Resource Locators (URLs) and Uniform Resource Name (URNs).
2. **XML:** Extensible Markup Language (XML) [24] is a standard text format for serializing data using tags.
3. **RDF:** Resource Description Framework (RDF) [15] is a general-purpose language for representing information in the Web. RDF model is the RDF graph, whose nodes are represented by RDF URI references, blank nodes or plain literals, and arcs are labeled with RDF URI references.
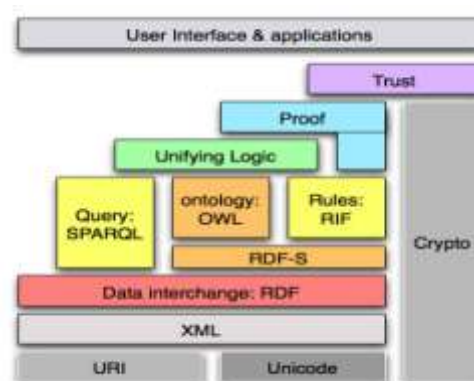

Figure 1: Semantic Web Components

V.   **APPLICATION OF SEMANTIC WEB**

1. In Software Engineering: Software engineering is the "Systematic Method for developing the software". Although the demand of software development being an engineering discipline is subject to ongoing discussions, there is no doubt that it has undergone basic changes during the last four decades. This assertion holds true both for emergence of new technology and sophistication of methodology. To decrease the complexity of software, the software implementations are abstracted in no of levels with the use of modeling and higher-level programming languages. Within the phase of Analysis and Design, the semantic application can be used in component reuse this is the potential application area during design phase.

2. In Requirements Engineering: Requirement engineering is the part of software engineering. This application can be used for Enhanced knowledge base using Semantic Retrieval technique, to describe requirements specification content and formally represent requirements base for simplified the complexity.

3. In Artificial Intelligence: Artificial Intelligence is the new era in science. The capturing knowledge is the main point to building huge and powerful Artificial Intelligence systems. In Artificial Intelligence it can use this application as computational models that enable certain kinds of automated reasoning.

4. In computer science and information science: Semantic Web is a formal description of knowledgebase as a set of object within a domain, and the relationships between those objects and classes. It shows the relation graph for describing the content of knowledgebase.

5. In web based system: This concept can be used to manage the classes and relationship between the objects which are used over the web for manage the knowledgebase.

VI.   **CONCLUSION**

In this paper, we reviewed research papers related to Semantic Web search and their characteristics. We focused upon the limitations of traditional search engine and how semantic will improve the searching capability for large amount of information present on web. In addition we give various definitions & issues of "Semantic Web". We conclude with various application of "Semantic Web".
In the future, our work will focus on the deeper and broader research in the field of intelligent semantic search, with the purpose of concluding the current situation of the field and promote the further development of intelligent semantic search engine technologies.

VII.   **REFERENCES**

[1]Berners-Lee, T., Hendler, J. and Lassila, O. "The Semantic Web", Scientific American, May 2001.

[2]Deborah L. McGuinness. "Ontologies Come of Age". In Dieter Fensel, J imHendler, Henry Lieberman, and Wolfgang Wahlster, editors. Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential. MIT Press, 2002.

[3]Ramprakash et al "Role of Search Engines in Intelligent Information Retrieval on Web", Proceedings of the 2nd National Conference; INDIACom-2008.

[4]T.Berner-Lee and M. Fishetti, Weaving the web "chapter Machines and the web,"Chapter Machines and the web, pp. 177-198, 1999.

[5]D.Fensal, W. Wahlster, H. Lieberman, "Spanning the semantic web: Bringing the worldwide web to its full potential, "MIT Press 2003.

[6]G. Bholotia et al.: "Keyword searching and browsing in database using BANKS," 18th Intl. conf. on Data Engineering (ICDE 2002), San Jose, USA, 2002.

[7]D. Tümer, M. A. Shah, and Y. Bitirim, An Empirical Evaluation on Semantic Search Performance of Keyword-Based and Semantic Search Engines: Google, Yahoo, Msn and Hakia, 2009 4th International Conference on Internet Monitoring and Protection (ICIMP ''09) 2009.

[8]"Top 5 Semantic Search Engines".http://www.pandia.com/.

[9]H. Dietze and M. Schroeder, GoWeb: a semantic search engine for the life science web. BMC bioinformatics,Vol. 10, No. Suppl 10, pp. S7, 2009.

[10]Fu-Ming Huang et al. "Intelligent Search Engine with Semantic Technologies"International journal of Web & Semantic Technology (IJWesT) Vol.2, No.1, January 2011.

[11]S. A. Inamdar1 and G. N. Shinde "An Agent Based Intelligent Search Engine System for Web mining" Research, Reflections and Innovations in Integrating ICT in education. 2008.

[12]Li Zhan, Liu Zhijing, Web Mining Based On Multi-Agents, COMPUTER SOCIETY,IEEE(2003).

[13]Patrick Lambrix et al, "Dwebic: An Intelligent Search Engine based on Default Description Logics"-1997.

[14]K. SatyaSai Prakash and S. V. Raghavan "Intelligent Search Engine: Simulation to Implementation", In the proceedings of 6th International conference on Information Integration and Web-based Applications and Services(iiWAS2004), pp. 203-212, September 27 - 29, 2004, Jakarta, Indonesia, ISBN 3-85403-183-01.

[15]D. Beckett. *RDF/XML Syntax Specification (Revised)*, February 2004. http://www.w3.org/TR/2004/REC-rdf-syntaxgrammar-20040210/.

[16] XiajiongShen Yan XuJunyang Yu Ke Zhang"Intelligent Search Engine Based on Formal Concept Analysis" IEEE International Conference on Granular Computing, pp 669, 2-4 Nov, 2007.

[17]Sanjibkumar, Sanjay kumarmalik "TOWARDS SEMANTIC WEB BASED SEARCH ENGINES" National Conference on "Advances in Computer Networks & Information Technology (NCACNIT-09) March 24-25,

[18]F. F. Ramos, H. Unger, V. Larios (Eds.): LNCS 3061, pp. 145–157, Springer-Verlag Berlin Heidelberg 2004.

[19]Cohen, S. Mamou, J. Kanza, Y. Sagiv, Y "XSEarch: A Semantic Search Engine for XML" proceedings of the international conference onvery large databases, pages 45-56, 2003.

[20]D. Bhagwat and N. Polyzotis, "Searching a filesystem using inferred semantic links," in Proceedings of HYPERTEXT '05 Salzburg, 2005, pp. 85-87.

[21]H. L. Wang, S. H. Wu, I. C. Wang, C. L. Sung, W. L. Hsu, and W. K. Shih, "Semantic search on Internet tabular information extraction for answering queries," in Proceedings of CIKM '00 McLean, 2000, pp.243-249.

[22]E. Kandogan, R. Krishnamurthy, S. Raghavan, S.Vaithyanathan, and H. Zhu, "Avatar semantic search: a database approach to informationretrieval," in Proceedings of SIGMOD '06 Chicago, 2006, pp. 790-792.

[23]A. Maedche, B. Motik, L. Stojanovic, R. Studer, and R. Volz, "An infrastructure for searching, reusing and evolving distributed ontologies," in Proceedings of WWW '03 Budapest, 2003, pp. 439-448.

[24]T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau. *Extensible Markup Language (XML) 1.0 (Third Edition)*, February 2004. http://www.w3.org/TR/2004/RECxml-20040204/.

[25]D. Ding, J. Yang, Q. Li, L. Wang, and W. Liu, "Towards a flash search engine based on expressive semantics," in Proceedings of WWW Alt.'04 New York, 2004, pp. 472-473.

[26]Chiung-Hon Leon Lee, Alan Liu, "Toward Intention Aware Semantic Web Service Systems," scc, vol. 1, pp.69-76, 2005 IEEE International Conference on Services Computing (SCC'05) Vol-1, 2005.

[27]M. H. Butler. *Is the Semantic Web hype?* http://www.hpl.hp.com/personal/marbut/ isTheSemanticWebHype.pdf.