

NEED OF FUTURE WEB TECHNOLOGY THE SEMANTIC WEB A BRIEF SURVEY

S.Raja Ranganathan¹Dr.M.Marrikkannan²Dr.S.Karthik³

¹Assistant Professor, Dept of CSE, SNS College of Technology-Coimbatore.

²Professor, Dept of CSE, Institute of Road and Transport Technology-Erode.

³Professor and Dean, Dept of CSE, SNS College of Technology-Coimbatore.

ABSTRACT-The World Wide Web (WWW) is a huge collection of information this allows the people to share information from the repositories of databases which are available globally. From day to day the information which is available inside database is growing rapidly also the number of user interaction with World Wide Web also increased. The search engines play a major role to extract these information to the user, there are many search engines available today but retrieving the most optimized result is a difficult task, however to overcome the problems in search engines for retrieving the optimized results for user queries, Semantic search technology plays a vital role. In this paper the detailed survey is made about the need of semantic technology and role of search engines in Expert based system and semantic search strategies.

KEYWORDS

Semantic technology, information retrieval, search engines, expert based system.

I. INTRODUCTION

The Semantic Web is able to describe things in a way which the computers can understand. This technology is the up gradation of the current searching methods [1] that provides the meaning of information in well-defined format that allows user to understand easily. In semantic web the solution for the problems is effectively overcome by its architecture itself. One of the main components in semantic web is (RDF) Resource Description Framework a new standard of W3C the search efficiency has been improved by multiple combinations made for user's keywords the RDF looks subject, Predicate and Object for each statement the user intend to search. The RDF is purely an XML language and RDF enables exchange and reuse of structured metadata. The second important component in semantic infrastructure is Ontology [2] this helps to make the relation among

the successful concepts. The ontologies use OWL web ontology language in different levels we can express they are OWL Lite, OWL DL and OWL Full

ordering by increasing level. The Semantic Web will support more efficient routing, expertise decision, integration and reuse of data and provide support for interoperability problem which cannot be resolved with current web technologies.

Many search engines search the keywords given by the user without any processing the keyword and publish the result without prioritization; in some cases these search results mislead the user to unwanted pages. This is because the topic wise search algorithms followed by the search engines these algorithms face problems for processing the user query's. for simple query this topic wise search is efficient but in day today life much more complicated intelligent based query arises in this situation these search engines produce vulnerable results. They show inaccurate result or some blind links in most of cases search engines shows "results not found". In future the topic based search engines we can't rely on.

In this paper we going to make a survey on existing literature about Semantic technology and semantic web search methodologies, by classifying the literature into few main categories, we review their characteristics respectively. In addition, the issues within the reviewed AI based semantic search methods and engines are analyzed and concluded based on perspectives.

II. BACK GROUND

The extraction of information from World Wide Web is not a new mechanism but we have to face challenges in information retrieval in many ways. There is different kind of search engines available in WWW each search engine follows a unique mechanism of indexing and processes of search of its own so the information extraction as well as the result produced by these search engines are not the same. Some of the popular search engines such as GOOGLE, YAHOO, BING and ALTA VISTA produce results based on queries after the keyword are processed. They only search

information available on the web page, recently updated, some research group's such as SWOOGLE start delivering results from their semantics based search engines, and however most of them are in their initial stages they face certain problems in matching ontology and combining keywords in RDF. The major problems facing by the search engines are they not able to gather content whole indexing in entire internet.

The World Wide Web is a huge collection of information which is available globally but huge lacks in existence of semantic mapping structure hence without any onto-semantic structure it is very difficult for the computer to understand the information which is provided by the user. In distributed environment web structure the following problems are widely faced they are

- How the queries are mapped to documents where the information available by the search engine but in most cases it not provide intelligent and meaning of the entire information what the user required?
- How the user queries are processed and stopping keyword makes search of user keywords?
- The web documents are connected one after another with the help of hyperlink so the results are distributed among different documents. The main problem here is how the search engine efficiently recognizes such distributed results?
- How the single crawler in a search engine able to search effectively a huge database which is available entire world with millions of servers got information of the user?

The infrastructure of the semantic web itself provide solution of the problems above stated hence it is basically AI based search technology in must produce meaningful information for the user queries with the help of ontology and RDF. The third problem is solved by graph based ontology models [6]. The need of the semantic arises in extra ordinary situations such as knowledge representation and natural language understanding. The below mentioned diagram explains the semantic framework

a. Limitations in Current Web

Even though the World Wide Web got huge collection of information in its distributed global database but still lacks in semantic structure. It is very difficult for the machines to understand the user

requirement in search space due to lack of semantic structure. Hence in most of the cases the search engines produces ambiguous results that mislead the user to wrong pages. The semantic web is developed to overcome some of the present problems.

- The contents available in web lack a proper structure regarding the representation of information.
- Ambiguity of information resulting from poor interconnection of information.
- There is lacking in automatic information tracking and transferring in web
- There is no universal format is followed so there is a lack in transparency and incapability of machines to understand the given information
- The crawler lose information in web without any universal structure

To overcome the above disadvantages the researchers work hard to implement semantic structure for example hokia a search engine works like Wiki it will match the user keyword and produce a meaning match result. The hokia following a semantic technology called Q-D-E-Xing [7] this is capable of processing Blogs, Newsletters and social networks also.]. It can process any kind of digital artifact by its Semantic web Rank technology using third party API [9].

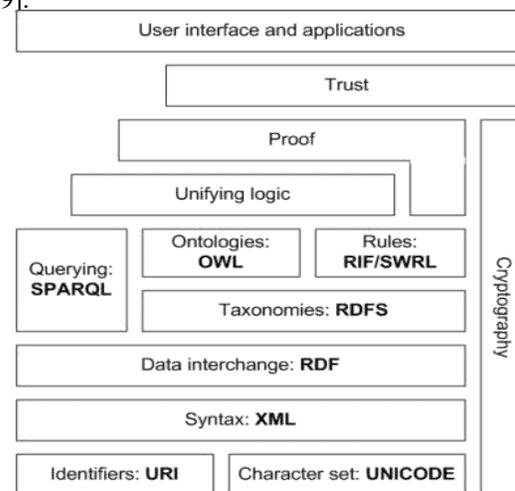


Fig.2.2. Isemantic web layer Infrastructure

III. INTELLIGENT SEMANTIC WEB

a. Intelligent Search Engines

The research conducted by the Patrick Lambrich and NahidShahmehri [13] about the precision and recall methodologies. The both scheme are widely used in semantic web to retrieve the information the user required but there is lot of problems enhancing the

performance. The new technique here adopted is subsumption information gathering technique this allows retrieving both desired topic as well as information about more specific topics. The results has been tested with new subsumption technique the performance of both precision and recall increased sufficiently testing carried out with small database the results are promising.

Anita Ferreira and John Atkinson discussed various issues on Artificial Intelligent search engines in his book the author described about spider web resource and intelligent information gathering model for intelligent searching and filtering using natural language feedback. The operation starts with natural language queries provided by a user (that is, general queries, general responses, feedback, and confirmation) and then passes them on to the discourse-processing phase the phase tells about RDF and Ontology, which generates the corresponding interaction turns (natural language output), arriving at a specific search request.

The Artificial intelligent based Semantic search engine with semantic interfacetechnology where included for several tests by FuMingHung[10]. The study is about the integration of ontology in Knowledge database concepts the ontology will map and form a graph based structure for every keyword the user intend to search. The efficiency has been dynamically improved by Description Logic Software the logic introduced for integration of ontology in Knowledge Database.

SathyasaiPrakash in his research [14] discussed about AI based search engine the new level of architecture is present here along with design specification for next generation semantic web search engine the author also described about push and pull models how the data being traversed in world Wide Web. The performance being measured by simulation studies using fuzzy satisfaction and heuristic search. The various analyses made after client and web interaction.

The agents [11] a software module works as assistant helps the user to make efficient search the research presented by Inamdar and Shindi. Many search engines we are using today uses a text based search mechanism. This searching methodology only sees text in web pages and extracts the text matched web page. To overcome this author discuss about agents each user is assisted by his/her own personal agent to search the web. The major goal of each personal agent is to propose to its user and to other agent's

links to web pages that are considered relevant for their search. The information both internally and externally used by local personal Agents. [12] These agents are a mainly software agent which mainly runs inside each server.

IV. SEMANTIC AND ONTOLOGY

Li Ding, Tim Finin, Anupam Joshi, Rong Pan, R. Scott Cost, Yun Peng, PavanReddivari, Vishal C Doshi, [25] and Joel Sachs presentsSWOOGLE a search engine uses a crawler indexing and ontology based retrieval mechanism for semantic web. The World Wide Web got huge collection of Concepts (Web pages) the main aim of this search engine is to extract metadata for each successfully discovered concept. The mapping carried out by forming the relationship between concepts this is effectively done with the help of ontology. The characters use N-Gram and URI this keyword that effectively matches the web concepts. Finally the information collected are stored and ranked with the help of rank algorithm.

Stefano Mazzocchi[26] discussed about the information resource sharing and how the information networks are created from seamlessly connected sources of information. This information is distributed throughout the world in decentralized manner so no control over the nodes. There is a need of data to be centralized but the problem is cost for implementing it amount of information stored with amount of people involved. The knowledge representation for information also prescribed how much knowledge the system able to provide users from its effective search results

V. COMMON ISSUES IDENTIFIED IN SEMANTIC WEB

The above discussion we have seen various aspects of artificial intelligent based semantic web search Technology. But lot of issues there once we implement this technology in real world. Some of the common issues arise as follows

a) High recall but low precision

We not able to say in all cases the semantic technology show its performance due to low precision in mapping concepts (pages) and unwanted recall is high. For example the semantic search engine Ding [25] is using the Google results and form a metadata for the top results obtained from Google in several test cases the semantic produce low precision but high in recall. Hence more optimized

mapping required to improve precision.

b) Identify User intention

For any semantic search engine it must act upon what the user is intend to search, using its intelligence it produce the result more relevant to user what he thought to search. For example the analysis has been made in research [26] how to match user intention and Semantic search hence it produce most suitable results for user.

c) Inaccurate queries phrase.

The user who search in search engine only have domain specific knowledge so he give the keyword or sentence. In search space the user will not include any synonyms or potential variation in query, that exactly matches results here the user have a problem but aren't sure how to phrase.

d) Crawler Efficiency

The World Wide Web got trillions of distributed information for the topic based search engines it only extract the topic in pages and produce a result to user. But once we go for semantic based search engines it is capable of making multiple choices for single user keyword. Second thing is it forms metadata key based search for processing each web documents or pages so single crawler is not sufficient to do the task.

VI. CONCLUSION

In this paper, we saw about a brief survey of the existing literature regarding Artificial intelligent based semantic search technologies. The review also made for various characteristics of semantic web respectively. In addition, the issues within the reviewed Artificial intelligent based semantic search methods and engines are concluded based on four perspectives High recall but low precision, Identify User intention, Inaccurate queries phrase and Crawler Efficiency. In the future, our work will focus on the deeper and broader research in the field of Artificial intelligent based semantic search techniques, with the purpose of concluding the current situation of the field and promote the further development of Artificial intelligent based semantic search technologies.

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. Fensel, 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 Hakkia, 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"
- [11] S. A. Inamdar 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. SatyaSaiPrakash 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] Dan Meng, Xu Huang "An Interactive Intelligent Search Engine Model Research Based on User Information Preference", 9th International Conference on Computer Science and Informatics, 2006 Proceedings, ISBN 978-90-78677-01-7.
- [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 on very large databases, pages 45-56, 2003.
- [20] D. Bhagwat and N. Polyzotis, "Searching a file system 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 information retrieval," 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] www.georges.gardarin.free.fr/Articles/Sewise_NLDB2003.pdf.
- [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.