National Conference on "Advanced Technologies in Computing and Networking"-ATCON-2015
Special Issue of International Journal of Electronics, Communication & Soft Computing Science and Engineering, ISSN: 2277-9477

A Search Engine Based on the Semantic Web

Ashlesha Khandalkar **P. D. Thakare** **Dr. M. V. Sarode**

***Abstract: W*eb search is a key technology of the Web, since it is the primary way to access content on the Web. Current standard Web search is essentially based on a combination of textual keyword search with an importance ranking of the documents depending on the link structure of the Web. For this reason, it has many limitations, and there are research activities towards more intelligent forms of search on the Web, called semantic search on the Web, or also Semantic Web search. we know that WWW allowing people to share the information globally from the big database repositories the amount of information grows billions of database hence to search particular information from these huge database it need to special mechanism which helps to retrieve information efficiently. now day various types of search Engines are available which makes information retrieving is difficult. but to proved the better solution in this problem ,semantic web engine are playing vital role basically main aim of [1]search engines is to providing the required information is small time with maximum accuracy .**

***Keyword-* Semantic search on the Web, Semantic Web, Ontology ,Information retrieval, Intelligent Search, Search Engine, XML, RDF**

I. INTRODUCTION

In this paper it presenting the new concept for semantic search engine which will answer the intelligent queries also efficiently and accurately. with keyword base search they usually provide result from blogs. the user cannot satisfaction with these result due to lack of trusts on blogs etc. to get trusted results search engine require searching [2][3]for pages maintain such information at some place. here the intelligent web based search .it use the power of XML metadata deployed on the page to search the queried information XML page will be consisted of built in and user defined tags. the metadata information of pages is extracted from xml into RDF. it get result very less time to the**.** answer the queries and provide more ccurate information.

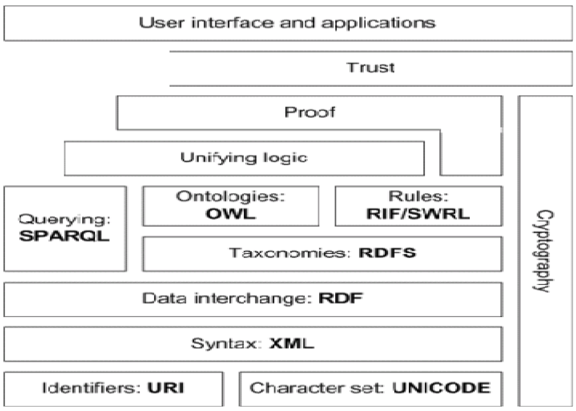
II. LITERATURE SURVEY

Information retrieval by searching information on the web is not a fresh idea but has different
challenges when it is compared to general information retrieval. different search engines return different search results due to the variation in indexing and search process. Google, Yahoo, it has been out there which handles the queries after processing the keywords. It[4] only search information given on the web page, recently, some research group's start delivering results from their semantics based search engines, and however most of them are in their initial stages. till none of the [4]search engines come to close indexing the entire web content, much less the entire internet.

Current web is the biggest global database that lacks the existence of a [5]semantic structure and hence it makes difficult for the machine to understand the information provided by the user. when the information was distributed in web, the two kinds of research problems in search engine i.e.

How can a search engine map a query to documents where information is available but does not retrieve in intelligent and meaning full information? hyperlink. how search engine can recognize efficiently such a distributed results?Semantic web [4] [5], can solve the first problem in web with semantic annotations to produce intelligent and meaningful information by using query interface mechanism and ontology's. Other one can be solved by the graph-based query models [6]. The Semantic web would require solving extraordinarily difficult problems in the areas of knowledge representation, natural language understanding. the following figure depicts the semantic web frame work it also referred as the semantic web layer cake by W3C.

Architecture:



The main purpose our architecture should serve is that provide a semantically unified interface for querying heterogeneous information sources. We do not aim at merging all possible sources together providing a cumulated view of all attributes. We argue that such an approach offers a very weak semantics, where the understanding of the semantic structure of all integrated sources is effectively left up to the user who is asking the query.[10][8][9] In System architecture, an underlying domain model consisting of hierarchies of concepts, relations, and possibly axioms is assumed to exist. This conceptual model (CM) is maintained centrally (at the schema level) but it is dynamically populated with instances during the query.resolution. the CM corresponds to an ontology and represents a semantic integration of the integrated data sources. it is described directly in RDF or RDF extended with some higherlevel ontology language. to create such a CM beforehand, ontology engineering tools are currently be available could be used. The main advantage of having an underlying [10]semantic model is that the way in which the data is structured (encoded) in the sources is transparent for the user, i.e. he can ask queries and interpret the results in understood.

***Inference and Mediating*:[10][]12][13]** The RDF Mediator is the central component of the architecture. It maintains the

CM, provides query and inference services, and also the support for traversing the results. the CM consists of a class (concept) hierarchy together with class properties,

and a set of rules that correspond to axioms about classes or their properties. hence by applying these rules on the set of facts which are retrieved, [11][]12]it is possible to infer new facts. CM (ontology) for all possible applications is not feasible (for scalability reasons). however, the distributed approach where one instance of the architecture (and thus one CM or ontology) serves as an input source for another architecture instance, brings scalability also in an environment like WWW. The mediator contains an RDF parser, a query decomposition module, and a query engine, which uses for inference to support the traversal (and so the actual retrieval) of the results, the mediator also has to implement an analogy of the API, modified for the RDF data model after the mediator receives a query from the application layer it proceeds as follows.

First, it analyzes[13] whether the query's resolution demands inference rules to be applied and if so, which are the facts that are needed to evaluate the inference rules. engine assumes that the facts are known beforehand, which is however, in contradiction with the on-demand retrieval approach. That's why the initial query must retrieve also these facts that enable the inference engine to apply the rules.Second[14],[15] it decomposes the query into subqueries and distributes them among the brokers. The actual querying is triggered by a

***Application Layer*:** There are numerous of applications that can take advantage of the semantically unified interface provided by the architecture. [8]The types of applications can vary from search agents (machine processing) to hypermedia front-ends that guide a (human) user in query composition and CM exploration, and produce as a response to a query a full-featured hypermedia presentation supporting browsing and user/platform adaptation I

***XML Instance Layer*:[14][15]** The layer offers the serialized XML data that results from the previous layer. Sometimes (when no wrapping is required) the two layers can be considered as one. Note that assuming heterogeneity, we do not impose any particular structure the XML data sources should comply to. this allows us to leave the XML wrapping process for the source providers.

***XML2RDF Layer*:** This layer consists of XML2RDF brokers which provide the bridge between the XML instance layer and the mediator. the designer tailors each XML2RDF broker to its source by specifying a mapping from XML sources to the underlying CM. [16][17][18]this mapping is used by the XML2RDF broker while resolving a query coming from the mediator. to establish a mapping from XML instances to the CM requires

***Ontology*:** A Semantic Web search is done relative to a fixed underlying *ontology*, which defines an alphabet of elementary ontological ingredients, as well as terminological relationships between these ingredients. The ontology may either describe fully general

[19][20][21]for vertical ontology-based search on the Web. The former results into a general ontology-based interface to the Web similar to Google, while the latter produces different vertical ontology based interfaces to theWeb. There are many existing ontologies that can be used, which have especially been developed in the context of the Semantic Web, but also in biomedical and technical areas. Such ontologies are generally created and updated by human experts in a knowledge engineering process.[15] Recent research attempts are also directed towards an automatic generation of ontologies from text documents, eventually coming along with existing pieces of ontological knowledge .

characteristic features of Semantic Web

It present a novel approach to Semantic Web search, called *Serene*, which allows for a semantic processing of Web search queries relative to an underlying ontology, and for evaluating ontology based complex Web search queries that involve reasoning over the Web[11][13][14].

It develop the formal model behind this approach. In particular, and introduce Semantic Web knowledge bases and Semantic Web search queries to them. It also define the Object Rank ranking for Semantic Web search.

It provide a technique for processing Semantic Web search queries, which consists of an offline inference and an online reduction to a collection of standard [8]Web search queries. prove that this way of processing Semantic Web search queries is always ontologically correct. Furthermore, to identify a large class of Semantic Web knowledge bases where it is also complete.

[6][7][8]It report on two prototype implementations of our approach in desktop search. Experiments with more than one million annotation facts show that the new methods are principally feasible and potentially scale to Web search actually much faster than desktop search, of larger search space).

The offline inference compiles terminological knowledge is called *completed* annotations. It prove that these have a polynomial size and can also be computed in polynomial time. experimental data show that they are also rather small in practice, especially since ontological hierarchies in practice are generally not that deep (a concept has at most a dozen super concept.[19]

III. Conclusion And Vision For The Future

[02][21]A semantic search on the Web (also called Semantic Web search), which is currently one of the hottest research topics in both the Semantic Web and the Web search community. In semantic search on the Web, the current strong research activities of the former to realize search on the Semantic Web are merged with the current strong research activities of the latter to add a semantics to Web queries and content when performing Web search. [16]It is through this integration that the reasoning capabilities envisioned in Semantic Web technologies are coming to Web search and the Web. it seen that , the formulation of queries and their result semantic search on the Web is ultimately directed by a third area, namely, the one of question answering systems, which is based on natural language processing. Although many approaches and systems to semantic search on the Web already

exist, the research in this area is still at the beginning, and many open research problems still persist. Some of the most pressing research issues are may be[17][18] (i) how to automatically translate natural language queries into formal ontological queries, and (ii) how to automatically add semantic annotations to Web content, or alternatively how to automatically extract knowledge from Web content. another central research issue in semantic search on the Web is (iii) how to create and maintain the underlying ontologies. This may be done either (a) manually by experts, e.g., in a Wikipedia like manner, where different communities may define their own ontologies ,or (b) automatically, e.g., by extraction from the Web, eventually coming along with existing pieces of ontological knowledge and annotations (e.g., from existing ontologies or [18][]19]ontology fragments, and/or from existing annotations of Web pages in micro formats or RDF), or (c) semi-automatically by a combination of (a) and (b). Clearly, the larger the degree of automation, the larger is also the potential size of ontologies that can be handled and the smaller are the costs and efforts for generating and maintaining them. So, for the very large scale of the Web, a very high degree of automation is desirable. A closely related important research challenge is (iv) the evolution and updating of and mapping between the ontologies that are underlying []17][18][19]20]semantic search on the Web, where it is similarly desirable to have a very high degree of automation. A further important issue is (v) how to consider implicit and explicit contextual information to adapt the search results to the needs of the users. For example, the needs and motivations of users may be defined in terms of ontology-based strict and soft (weighted) constraints and (conditional) preferences (e.g., similar to , which may then implicitly be expanded into the semantic search query and/or used in the computation of the ranking on objects and search results. Performing Web search in the form of returning simple answers to simple questions in natural language is still science fiction, let alone performing Web search in the form of query answering relative to some concrete domain or even general query answering. However, with the current activities towards semantic search on the Web, we are moving one step closer to making such science fiction become true, which ultimately aims at a human-like interface to the knowledge, information, services, and other resources available on the Web.

REFERENCES

1].Buitelaar, P. (2010). Ontology-based Semantic Lexicons: Mapping between terms and object descriptions. In: C.R. Huang, N. Calzolari, A. Gangemi, A. Lenci, A. Oltramari & L. Prevot(Eds.), Ontology and the Lexicon (pp. 212–223). Cambridge: Cambridge University Press.

2]A., Phipps, J., & Rubin, D. (2009). SKOS use cases and requirements. http://www.w3.org/TR/2009/NOTE-skos-ucr-20090818/, Accessed 19 October 2010.

3.]Shadbolt, N., Hall, W., & Berners-Lee, T. (2006). The semantic web revisited. IEEE Intelligent system21(3), 96–101.

4] A. Go´mez-Pe´rez and O. Corcho, "Ontology Languages for the Semantic Web," IEEE Intelligent Systems, vol. 17, no. 1, pp. 54-60,

5] J. Euzenat, "Research Challenges and Perspectives of the SemanticWeb Intelligent Systems," IEEE Intelligent Systems vol. 17, no. 5pp. 86-88, Sept.-Oct. 2002.

6] N.F. Noy, M. Sintek, S. Decker, M. Crubezy, R.W. Fergerson, andM.A. Musen, "Creating Semantic Web Contents with Protege-2000," IEEE Intelligent Systems, vol. 16, no. 2, pp. 60-71.

7] E. Selberg and O. Etzioni, "The MetaCrawler Architecture for Resource Aggregation on the Web," IEEE Expert, vol. 12, no. 1,pp. 11-14, 1997.

8] B. Liu, W.S. Lee, P.S. Yu, and X. Li, "Partially Supervised Classification of Text Documents," Machine Learning: Proc. 19th Int'l Conf. (ICML 2002), pp. 387-394, 2002.

9] S.M. Pahlevi and H. Kitagawa, "Taxonomy-Based Adaptive Web Search Method," Proc. Third IEEE Int'l Conf. Information Technology:Coding and Computing (ITCC 2002) pp. 320-325, 2002.

10] G. Cheng, W. Ge, and Y. Qu. Falcons: Searching and browsing entities on the Semantic Web. In *Proceedings WWW-2008*, pp. 11011102. ACM Press, 2008.

11] L. Ding and T. Finin. Characterizing the SemanticWeb on theWeb. In *Proceedings ISWC-2006*, *LNCS* 4273, pp. 242–257. Springer, 2006

12] L. Ding, T. W. Finin, A. Joshi, Y. Peng, R. Pan, and P. Reddivari. Search on the Semantic Web. *IEEE Computer*, 38(10):62–69, 2005.

[20] C. d'Amato, F. Esposito, N. Fanizzi

13]. C. d'Amato, F. Esposito, N. Fanizzi, B. Fazzinga, G. Gottlob, and T. Lukasiewicz. Inductive reasoning and Semantic Web search. In *Proceedings SAC-2010*, pp. 1446–1447. ACM Press, 2010.

14] C. d'Amato, N. Fanizzi, B. Fazzinga, G. Gottlob, and T. Lukasiewicz. Combining SemanticWeb search with the power of inductive reasoning. In *Proceedings URSW-2009*, *CEURWorkshop Proceedings* 527. CEUR-WS.org, 2009.

15] D. Damljanovic, M. Agatonovic, and H. Cunningham. Natural language interface to ontologies: Combining syntactic analysis and ontology-based lookup through the user interaction. In *Proceedings ESWC-2010*, Part I, *LNCS* 6088, pp. 106–120. Springer, 2010.

16]C. d'Amato, N. Fanizzi, B. Fazzinga, G. Gottlob, and T. Lukasiewicz. Combining Semantic Web search with the power of inductive reasoning. In *Proceedings SUM-2010*, *LNCS* 6379, pp. 137–150.Springer, 2010.

17]J. de Bruijn and S. Heymans. Logical foundations of (e)RDF(S): Complexity and reasoning. In *Proceedings ISWC-2007*, *LNCS* 4825, pp. 86–99. Springer, 2007.

18] M. Fern´andez, V. Lopez, M. Sabou, V. S. Uren, D. Vallet, E. Motta, and P. Castells. Semantic search meets the Web. In *Proceedings ICSC-2008*, pp. 253–260. IEEE Computer Society, 2008.

19] G. Kasneci, F. M. Suchanek, G. Ifrim, M. Ramanath, and G. Weikum. NAGA: Searching and ranking knowledge. In *Proceedings ICDE-2008*, pp. 953–962. IEEE Computer Society, 2008.

20] H. P´erez-Urbina, I. Horrocks, and B. Mo Springer *ISWC-2009*, *LNCS* 5823, pp. 489–504., 2009.

21] S. Zhang, C. Zhu, J. K. O. Sin, and P. K. T. Mok, "A novel ultrathin elevated channel low-temperature poly-Si TFT," IEEE Electron DeviceLett., vol. 20, pp. 569–571, Nov. 2010.

AUTHOR'S PROFILE

Ashlesha V. Khandalkar
A. V. Khandalkar has completed B.E. Degree in Computer Engineering from Sant Gadge Baba Amravati University, Amravati, Maharashtra. She is persuing Masters Degree in Computer Science and Engineering from Jagadambha college of engineering ,Yavatmal of Amravati university.

Prof P.D.Thakare
P.D Thakare has completed B.E Degree in Computer science & engg.from sant Gadge Baba Amravati University, Amravati, Maharashtra. he has completing PG in Amravati university (S.G.B.A.U) computer science and engineering