ISSN: 2321-7782 (Online)
Impact Factor: 6.047

Volume 4, Issue 7, July 2016

International Journal of Advance Research in Computer Science and Management Studies

Research Article / Survey Paper / Case Study

Available online at: www.ijarcsms.com

Semantic Web Search Engine

Nikhil Chitre
B.E, Department of Information Technology, MAEER's MIT
College of Engineering, Kothrud, Savitribai Phule
Pune University, Pune, India

Abstract: ***The World Wide Web permits people to share their data from huge database repositories globally. The quantity of data grows millions of databases. There are a lot of search engines that are available today, and retrieving meaningful data is complex. However to extract meaningful data intelligently, the semantic web technologies are playing a chief role to overcome this problem in search engine. This paper presents analysis of the search engine generations and the part of search engines in semantic search technologies and intelligent web. We have also summarized the various techniques and advantages of a few important semantic web search engines that have been developed so far. Most prominent part is how the semantic search engines are different from traditional searches.***

Keywords: Search Engine, Information retrieval, Semantic web, Intelligent Search.

I. INTRODUCTION

The Semantic Web is the actually the extension of World Wide Web that enables users to share the content beyond the limits of websites and applications. It has been defined in rather various ways: as a natural paradigm shift in our everyday use of the Web, or as a utopic vision, or merely as a web of data. Primarily, the Semantic Web has motivated and engaged a lot of users to create innovative semantic applications. Semantic search plays a major role as it promises to produce accurate answers to user's queries by taking benefit of the availability of explicit semantics of information. For example, when looking for news stories about the commerce students, having traditional searching technologies, often only those news entries in which the term "commerce students" appears. Hence, those entries which mention the commerce students by names but do not actually use the term "commerce students" directly will be actually missed out. However, such entries are often the ones that the user is in fact interested in. In the perspective of the semantic web, where the meaning of web content is made unambiguous, the semantic meaning of the term (which is the general concept in the example of commerce students) can be understood.

A lot of semantic searching tools have been recently created [8],[9],[10]. Review of the semantic search tools tells that while these tools do improve the performance of traditional searching technologies, they are though not suitable for naive users, i.e. end users who are not familiar with domain specific semantic information, SQL-like query languages or ontologies.

A large amount of the conventional search engines get the answers syntactically accurate but larger in amount. Semantic search engine allows the information to be accurately portrayed in terms of well defined expressions. A semantic web search engine gives selected answers which the user is looking for. The main goal of Semantic Web is to make the Web content easily understandable not by humans as well as machine understandable. It is needed to ensure that the semantics are not missing during the complete life cycle of information extraction. Several semantic web search engines developed so far vary from each other through the outcomes attained & technologies used which are discussed in detailed in the later sections.

II. Traditional Search Engine and its Limitations

The World Wide Web today is the global database that is short of the existence of a semantic structure and therefore it becomes complex for the machine to comprehend the information provided in the form of search strings by the user. As for outcomes, the search engines return the partially ambiguous or ambiguous outcome data set. To overcome the following major limitations of the current web, Semantic web is being developed.

- Automatic information transfer is missing.
- Unable to deal with huge amount of users and content ensuring trust at every level.
- Poor interconnection of information due to ambiguity of information.
- Inability of the machines to comprehend the provided information because of lack of a universal format.
- Regarding the representation of information, the web content lacks a proper structure.

Hakia [3] which is a general purpose semantic search engine that searches the structured text like the Wikipedia. They're trying to make available search results based on some meaning match, rather than by the popularity of search expressions. The presented news, galleries, Credible and Blogs are processed by QDEXing [3], which is hakia's proprietary core semantic technology.

III. Semantic Web Search Engine

The semantic web search vastly improves the search accuracy of the data related to the query and the semantic search engine delivers the precise content that the users intent to know. There's obviously no denying about the power and the popularity of Google search engine. However, there are other methods to search the internet, and that is using semantic search engines. Using the semantic search engine we will make sure that it results in more appropriate and smart outcomes. The search engines are able to evaluate or retrieve the information and gives very accurate results for the search queries. The aim of semantic web is to provide one extra machine understandable layer, that will considerably make simpler the programming and the maintenance effort for the knowledge-based web services. Many of the semantic web languages have been created like RDF, DAML+OIL, OIL and OWL.

Approaches to semantic web:

There are mainly four approaches to a semantic search. Different semantic web search engines may use one or more of the given approaches. The goal of semantic search is to use the meaning (semantic) to improve the search experience of the users.

The **first approach** uses a contextual analysis in order to help to disambiguate the search queries. Let us see an example, the word "goal" refers to football or aim or something else entirely.

The **second approach** focuses on the reasoning. Given a collection of facts which are represented in the given system, additional facts can also be deduced from them. If the system is familiar with Paul's children, and it also knows who each of their children were, and hence a reasoning system can deduce who Paul's grandchildren were.

The **third approach** stresses on the natural language understanding. In this approach, the engines process the content that they index and also the search queries that the users submit so as to try to identify the objective of the information. It uses the syntax of the statement and the rules to recognize the users, places, the organizations, and so on. Powerset makes wide use of the natural language understanding.

The **fourth approach** makes use of ontology to represent the knowledge about a particular domain and also expand queries. When a user types a query for a term such as "car," the system then adds the terms from its ontology (e.g., "vehicle"

since a car is a type of vehicle) so as to make the search extra focused as well as more broad. This approach is used by a great number of semantic search engines.

Therefore, there is not only one approach to the semantic search. Many semantic search engines usually mix and match the approaches in various kind of ways to yield an exclusive search experience for the users. Every approach has a lot to contribute. Several types of searches are intended to fulfill various different kinds of functions.

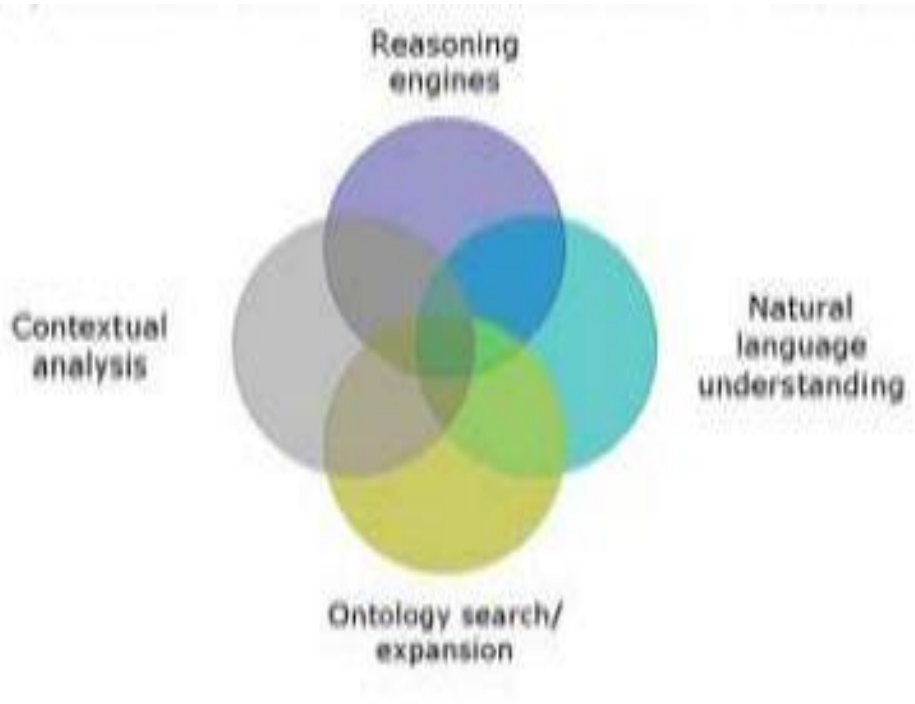

Fig1. Approaches to Semantic web

IV. TYPES OF SEMANTIC SEARCH ENGINE

The procedure of conveying enough meaning so to result in an action is semantic. A series of symbols can also be used to convey the meaning, and then this communication can affect the behavior. Semantics Web is the next generation of the Web that is driven by Semantic, where focus is on the part of semantics for automated approaches for exploiting of the resources in the Web. The word 'Semantic' also shows that the meaning of data can be discovered not only by users, but also by the computers. The Semantic Web was then made so as to extend the web which makes data simple to reuse everywhere.

A. Semantic search engines

Presently, a lot of semantic search engines are being developed and are implemented in various working environments, and these methods can be made use of in order to realize the present search engines. *Glauco Schmidt and Alcides Calsavara* proposes a novel type of service for semantic search engine. This search engine stores semantic data about the resources on the web and is able to solve difficult queries, also considering the context where Web resource is being targeted, and how this search engine may be utilized so as to permit the users to obtain information about certain commercial products and services and also about sellers and the service providers that can be hierarchically organized [2]. For the development of e-business applications, semantic search engines can be employed since it is based on widely accepted standards and strong theory.

Sara Cohen Jonathan Mamou et al presented XSEarch, a semantic search engine for XML [13]. It has an easy query language, that is suitable for a inexperienced users. It mainly returns the semantically related data fragments that satisfy the query given by the user. These query answers are the ranked using extended information-extraction methods. To facilitate the efficient implementation of XSEarch, advanced indexing techniques were developed.

Wang et al. project which is a semantic search methodology to extract information from the normal tables, and it has three main steps: recognizing the semantic relationships between the table cells; changing the tables into data in form of database; and retrieving the objective information by using query language. The research purpose defined by the authors is how we can use a given table and domain knowledge in order to convert that table into a database using semantics.

B. Ontology search engines

Georges Gardarin et al. discussed an ontology-based Web information system called SEWISE , to support the Web information description and extraction. SEWISE can map the text data from various Web sources into one XML structure and also make the concealed semantic in text available to the program. Textual data of interest is automatically retrieved by Web Wrappers from a variety of Web sources and then text mining methods such as classification and summarization are used to process the extracted text information.

Maedche et al. proposed an integrated method for ontology searching, reprocess and update [2]. In this architecture, ontology registry is designed that stores the metadata about the ontologies and the ontology server stores ontologies. The ontologies in these servers can be produced, duplicated and evolved. The ontology metadata in registry can be queried on and also recorded when a new ontology is produced. The ontology search in this registry is executed under two main situations - query-by-example is to limit search fields and the search terms, and query-by-term is to limit hyponyms of terms for the search.

V. SOME COMMON ISSUES

There are some common issues in the present semantic search engines and techniques are summarized as follows:

i. Identity intention of user

This plays a significant role in the intelligent semantic search engine. Consider the example, in *chiung-Hon leon lee* proposed a technique for analyzing request expressions to fit the user intention, as a result of that the service offered will be more suitable for user.

ii. High recall and Low precision

A few Intelligent semantic search engines are unable to show their significant performance in improving the precision and lowering the recall. In semantic flash search engine, proposed by Ding, the source of search engine is based on the first-50 returned results from Google which is not a semantic search engine, which could be high recall and low precision.

iii. Inaccurate queries.

The users use typically domain specific knowledge and they don't include all the potential synonyms and the variations in search query. Users have a problem but they aren't sure how to phrase their query.

iv. Individual user patterns can be extrapolated to the global users.

In previous search engines that presented disambiguation to search terms, a user could type in a search word that was ambiguous (e.g., Java) and search engine would return a list of options (programming language, coffee, island in the South Seas).

VI. RESULT OF FEW SEMANTIC SEARCH ENGINE

A. DuckDuckGo

This is a feature-rich semantic web search engine that gives numerous reasons to put away Google. If a user searches for a keyword that has more than one connotation, it gives you the option to choose what you were originally searching for, along with its disambiguation outcomes. Consider the example, looking for the term Apple will display a big list of possible meanings including fruit, bank, company.

Fig2. DuckDuckGo

B. Kngine

The search results of kngine are divided into web results or image results. These are preceded by the information about the search word, which is known as 'Concepts'. Consider the example, searching for "iPhone 6S" will be then be preceded by the specifications of the device. Searching for a movie will be then preceded by information about the movie, its trailer links, reviews on it and quotes. Similarly searching for a city will also be preceded by the information about the city such as its local attractions, weather and restaurants and hotels. Currently kngine contains greater than eight million Concepts this is where the strength of the site resides.

Fig3. Kngine

C. Hakia

Hakia is one of the semantic web search engines that returns relevant results based on the concept match and not on the keyword match or its popularity ranking. But the engine also prompts to enter not just the keywords - but a question or a phrase or a sentence as well. They then provide the search outcomes based on the meaning match, rather than by the popularity of the search terms. A very important ability of hakia search engine is that it gets the outcomes using equivalent terms like, "kill=murder" or "treat=cure". These relations are deployed with suitable disambiguation techniques.

Fig4. Hakia

VII. CONCLUSION

This paper gives a brief summary of some semantic search engines that uses different approaches in various ways so as to yield exclusive search experience for its users. It is now clear that searching the web today is sort of a challenge and it is predictable that more than half of the difficult questions go unreturned. Semantic web search can improve the limitations of traditional web search. Future work includes developing a proficient semantic search engine technology that can deal with the limitations and the challenges efficiently and also is compatible with the global standards of the web technology.

References

1. G.Sudeepthi, G. Anuradha ,Prof. M.Surendra Prasad Babu "A Survey on Semantic Web Search Engine", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 2, No 1, March 2012.
2. G.Madhu, Dr.A.Govardhan, Dr.T.V.Rajinikanth "Intelligent Semantic Web Search Engines: A Brief Survey", International journal of Web & Semantic Technology (IJWesT) Vol.2, No.1, January 2011.
3. 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.
4. Jyoti chaurasia, Prof. Jagdish Raikwal "Survey on Semantic Web Search Engine: Using Domain Ontology", International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056 Volume: 02 Issue: 02 | May-2015.
5. S. Latha Shanmuga Vadivu, M. Rajaram, and S. N. Sivanandam "A SURVEY ON SEMANTIC WEB MINING BASED WEB SEARCH ENGINES", ARPN Journal of Engineering and Applied Sciences, VOL. 6, NO. 10, OCTOBER 2011.
6. Peng Wang "A Search Engine Based on the Semantic Web", May 2003.
7. D.Pratiba, Dr.G.Shobha "Semantic Web Search Engine", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 6, June 2013.
8. O. Corby, R. Dieng-Kuntz, and C. Faron-Zucker. Querying the Semantic web with Corese Search Engine. In Proceedings of 15th ECAI/PAIS, Valencia (ES), 2004
9. R. Guha, R. McCool, and E. Miller. Semantic Search. In Proceedings of the 12th international conference on World Wide Web, pages 700–709, 2003.
10. J. Heflin and J. Hendler. Searching the Web with SHOE. In Proceedings of the AAAI Workshop on AI for Web Search, pages 35 – 40. AAAI Press, 2000.
11. V. Lopez, M. Pasin, and E. Motta. AquaLog: An Ontology-portable Question Answering System for the Semantic Web. In Proceedings of European Semantic Web Conference (ESWC 2005), 2005.
12. C. Rocha, D. Schwabe, and M. de Aragao. A Hybrid Approach for Searching in the Semantic Web. In Proceedings of the 13th International World Wide Web Conference, 2004
13. 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.

