

Survey on Semantic Web Search Engine: Using Domain Ontology

Jyoti chaurasia¹, Prof. Jagdish Raikwal²

¹ Student, Computer Science, DAVV University, MP, India

² Professor, Computer Science, DAVV University, MP, India

Abstract - *The tremendous growth in the volume of data and with the terrific growth of number of web pages, traditional search engines based on keyword based searching, now a days are not appropriate and not suitable anymore. Search engine is the most important tool to discover any information in World Wide Web. Semantic Search Engine is born of traditional search engine to overcome the above problem. The Semantic Web is an extension of the current web in which information is given well-defined meaning. Semantic web technologies are playing a crucial role in enhancing traditional web search, as it is working to create machine readable data. In this paper we propose a semantic search engine based on domain ontology. The most prominent part is that how the semantic search engines differ from the traditional searches and their results are shown by giving a sample query as input.*

Key Words: RDF, OWL, DAML+OIL, OIL, QDEX.

1. INTRODUCTION

The rapid growth of Internet has given user an easy way of accessing information and services. Web is a huge semi structured database that provides with vast amount of information. With ever-increasing information overload, we are facing new challenges for not only locating relevant information precisely but also accessing variety of information from different resources automatically. Efficient searching is required to get high quality results which are based on appropriate matching between well-defined resources and user queries. When users use search engines to search for specific information, the quality of the search results will be improved significantly if they make use of semantic search. Most of the traditional search engines get the answers syntactically correct but larger in amount.

The Semantic allows the information to be precisely described in terms of well-defined vocabularies. A semantic search engine gives selected results which the user is searching for. The main objective of this proposed method is to provide the hyperlinks with the keywords in the Semantic Web search to make easy and fast access of hyperlinks from the large amount of data scattered over the Web. It is very easy to extract useful knowledge from structured data using proven algorithms and patterns. But the problem comes when we have unstructured data to work with. It becomes very difficult to extract knowledge from the un-structured data because of non-availability of proven algorithms, schemas, patterns and information systems. Through this paper, we propose an algorithm to overcome the various challenges in the field of unstructured data mining using semantic web techniques

2. CONVENTIONAL SEARCH ENGINE & ITS LIMITATIONS

Today's World Wide Web is the global database that lacks the existence of a semantic structure and hence it becomes difficult 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 developed to overcome the following main limitations of the current Web.

2.1 Limitations

1. The web content lacks a proper structure regarding the representation of information.
2. Ambiguity of information resulting poor interconnection of information.
3. Unable to deal with enormous number of users and content ensuring trust at all levels.
 4. Incapability of machines to understand the provided information due to lack of a universal format.
 5. Automatic information transfer is lacking.

3. CRITICAL ANALYSIS

Research in the field of semantic search engines is focused on various approaches and classification theories. Miller et al. [1] talked about Navigational Searches which points to the classification of documents based upon the intention of the user. Mangold [2] focused on architecture, coupling, user context, query modification, transparency, structure of ontology and relevant technology as parameters to realize semantic search. In another critical research on semantic search engine [3], it is pointed out that augmenting traditional keyword search with semantic techniques is considered as the important parameters to implement the semantic search engine. Hildebrand et al. [4] suggested a search system based upon query construction in section with custom search algorithms. Dietze and Schroeder [5] suggest a new classification approach based on 9 criteria which include structured/unstructured file, text mining type, type of documents, number of documents, Ontologies, clustering, result type, highlighting, scientifically evaluated. Dong et al. [6] present a extended classification with semantic search algorithm based on the Graph, methodology on distributed hash tables and logics-based Information retrieval.

4. PROPOSED METHODOLOGY

In this proposed method accessing the web using semantic search engine highly improves search accuracy of the query related data and the search engine delivers the exact content with their hyperlinks to the user. There's no denying the power and popularity of the Google search engine. But there are other ways to search the web content, using semantic search engines. By using our proposed algorithm semantic search engine will ensure that it results in more relevant and smart results. Using domain ontology we will create clusters using OWL languages. Each cluster will combine all related hyperlinks of the query for that user is looking for. Instead of giving different search results related to the keyword, semantic search will give all different classes of the search query and each result directly will show the hyperlink of the query as shown in the Figure 1. Many semantic web languages have been developed like RDF, OIL, DAML+OIL and OWL.

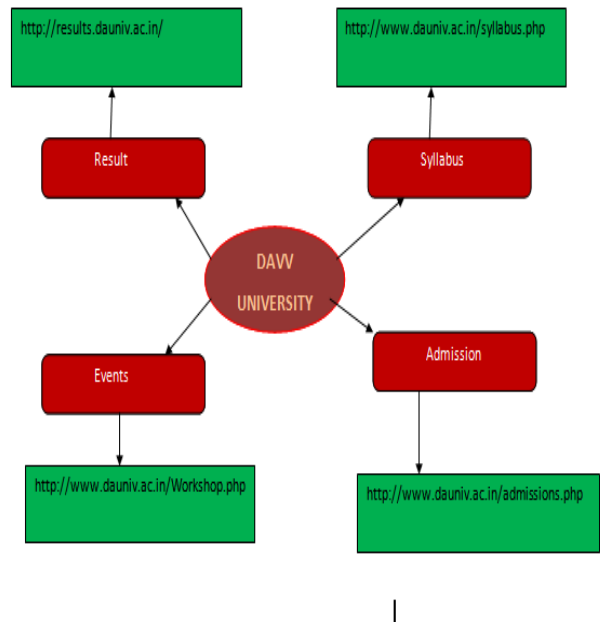


Figure: 1

In the above diagram *ellipse* shows the user query in the search engine. It will contain all different classes of the search keyword as displayed in the figure by *rounded rectangle*. Each class contains the hyperlink for the query result. Instead of giving different search result for the user query, our approach will directly show the resultant hyperlink for the user query as depicted by *rectangle*. This process will take less time for searching comparatively traditional search method.

Table 1

Primary_Word	Secondary_Word	Hyperlinks.
DAVV	Result	http://results.dauniv.ac.in/
	Syllabus	http://www.dauniv.ac.in/syllabus.php
	Events	http://www.dauniv.ac.in/Workshop.php
	Admission	http://www.dauniv.ac.in/admissions.php

5. PROPOSED ALGORITHM

```

AlgorithmSearch(SearchKeyword)
int a = 0;
Boolean Value= True
While(value)
{
    If(SearchKeyword=
primary_word+Secondary_Word+.....)
    {
        a=1;
        value= false;
    }
}
if (a= 0)
{
    Hyperlink not available
}
else
{
    Display hyperlink
}
Exit.
    
```

The above algorithm description

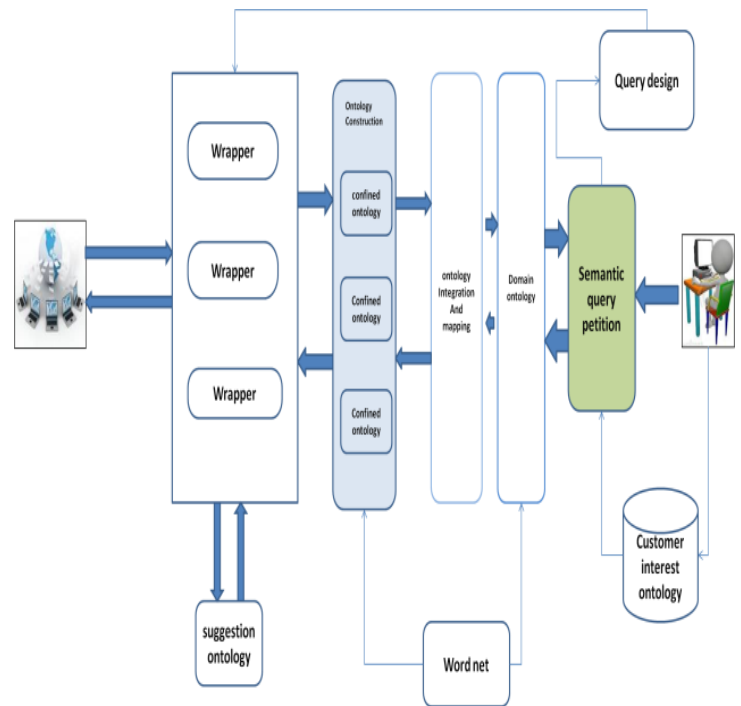
AlgorithmSearch(Search Keyword) :-Function has parameter search Keyword which is the input given by the client. With the help of the keyword the linear searching process will take place and scans the complete database and if the search Keyword matches with the database Value the exact hyperlink will be shown to the client on the same page.

6. DESCRIPTION

6.1 RETRIEVING THE DATA FROM THE CLIENT SIDE

Step-1 client will enter keywords in the search engine.

Step-2 Query will send to the specific domain related to the keywords



6.2 PROCESSING AT THE SERVER

Step-1 Customer interest ontology will suggest different ontology classes for the searching keywords.

Step-2 Match-making of search keywords with existing database keywords

Step-3 Linear searching process will take place in the database; If the perfect match is found then exact hyperlink will be displayed to the client.

Step-4 If searching keywords does not match with the database keyword, then hyperlink will not be shown to the client and message will be shown "enter the correct query".

7. CONCLUSION

Semantic web mining is relatively new sub-field of data mining. It has a vast scope for investigation keeping in view of the availability of tons of unstructured data on WWW. Lack of available global standards on this subject opens up enormous prospect for the research community to focus on this area in a big way. Non-availability of a rugged database managementsystem to manage semantic web mining opens up new avenues for the researchers to develop KIMS (Knowledge extraction management system) for unstructured data available on the web. Auser-

oriented semantic search engine is the need of the day. These fields if explored in a right manner will provide unlimited opportunities to extract knowledge from the goldmine of unstructured data available across the globe.

8. REFERENCES

[1] Miller, Guha, R., McCool, R. (2003). "E. Semantic Search". Proceedings of the WWW'03, Budapest

[2] C Mangold (2007). "A survey and classification of semantic search approaches".

International Journal of Metadata, Semantics and Ontologies, pp. 23–34.

[3] D. Buscaldi, P. Rosso, E. S. Arnal (2005). "A wordnet-based query expansion method for geo-graphical information retrieval". Working Notes for the CLEF Workshop.

[4] M. Hildebrand, J. Ossenbruggen, and L. Van Hardman (2007). "An analysis of search based user interaction on the semantic web". Report, CWI, Amsterdam, Holland.

[5] F. Figueira, J. Porto de Albuquerque, A. Resende, Geus, P. Lício de Geus, G. Olso (2009).

"A visualization interface for interactive search refinement". 3rd Annual Workshop on Human-Computer Interaction and IR, Washington DC. pp. 46-49.

[6] H. Dong, FK Hussain, and E. Chang (2008). "A survey in semantic search technologies". 2nd IEEE International Conference on Digital Ecosystems and Technologies.

REFERENCES

[1] S. M. Metev and V. P. Veiko, *Laser Assisted Microtechnology*, 2nd ed., R. M. Osgood, Jr., Ed. Berlin, Germany: Springer-Verlag, 1998.

[2] J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[3] 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 Device Lett.*, vol. 20, pp. 569–571, Nov. 1999.

BIOGRAPHIES



Ms. Jyoti Chaurasia
Master of Engineering
CSE Department
IET-DAVV University, Indore
MP



Professor Jagdish Raikwal
CSE Department
IET-DAVV University, Indore
MP