Semantic search engine: A survey

Anusree.ramachandran, R.Sujatha
School of Information Technology and Engineering, VIT University
Vellore-632014, T.N, India
anusree.ramachandran@gmail.com
r.sujatha@vit.ac.in

Abstract

The semantic search engines have some advantages on the web search engines from the users view. In this fast life everybody need the answer for their queries very fast. In this scenario semantic searching engines will be helpful. It deals with the actual meaning of the queries. The tremendous growth in the volume of data or the information lead the traditional search engines to get the answers syntactically correct but large in amount. That might be the reason to get into the semantic search engines which gives the selected results which the user searching for. So here in this paper, a survey is done about the semantic search engines to revel the promising features of the semantic search engines (SSE). It deals about the description of some of the best semantic search engines.

1. Introduction

The new generation prefers the semantic searching. In this fast life nobody have time to keep searching in the links to get the exact meaning or to know the extract of the searching word. Here comes the exact use of semantic search engines. In ordinary search engines it gives the links, and that also may not be giving the appropriate result searching for. But in the semantic search engines it gives the extract of the content (searching for) as a small paragraph and followed by the link. To just know what the searching word is, that information must be enough which we get in the first page while we search a word in the semantic search engine.

Semantic search engines differ from the ordinary search engines. Even in semantic search engines also the results given by each and every search engines differs from each other. The most prominent part is that how they differ from the traditional searches and their results. The way in which the semantic search engines helps the user. The upcoming of the semantic search engines one or the other way helping the user to get the prominent data in a short period time. The paper deals about the conventional searching and the semantic searching along with methodologies and some of the semantic search engines to support the need for highly effective search engines.

2. General Search Engines

It includes the search engines Google yahoo etc. which provides a number of links when search the user for a query. It became a vast collection of information for the searches. It may not be containing the exact fact but it searches the query related all items which syntactically matches for the searched query. As far as users are concerned they need relevant and precise results.

Conventional Searching

Conventional searching helps the user to have the links of the searched query. It gives all the possible urls. In conventional searching it is not considering about the different meanings the words can have in fact it will show all the matches possible. By clicking or going through the links only we can have the clear picture about the query we are searching for. But it is not the case of the semantic search engines.

It is time consuming process if we goes through each and every links one by one. Any it may not contain the contents which we are searching. That may also happen in this type of search engines. While comparing this with the semantic search it is giving a difficult way to the user to get the result of the specified query.
The conventional search engines always provide the links for the user to go through to reach the results. It will also have the short keys to search in the web, pictures, videos, news, shopping etc... but the user will not get the answer for the query. In all these searches the search engine will provide the list of links by which the user can reach the destination. There will not be any answer as such in the search result the must have to navigate through the pages to get the intended thing. If the user are in search of a particular document that is searching for a specified document, this type of searching will be useful. If the users are doing it for attaining the knowledge it will not be a helpful search. In conventional search engines it is not sure that the search thing is the same what we are getting or the other possibilities of the same query. It will be cleared by the semantic search engines. In semantic search engines the small picture about the content will be given along with the description but it is not done in conventional searches. While considering the human beings the concepts will be more clear by seeing the pictures rather than seeing the or going through the notes and documents so semantic search is helpful in this way also.

3. Semantic Search Engines

Semantic search engines include the searching of the query related to the entered data by the user in the data space. But the thing is that it gives a short review of the commonly used related word description for the convenience of the users, it makes the job easier-the search can be done easily. Semantic search engine differs according to the user. For the ordinary users no need of analyzing the data which is in high level. The very depth accessing of the data is not needed. So for the research purpose and for the students also for the technical professional the semantic search engine differs.

Semantic Searching

The semantic search engine reveals us the data based search result for the query we used. It gives the exact meaning definitions which is very common and it gives the other meanings also. Compared to the other search engines the semantic search engines help us to find the answer of the searching queries very fast and accurately. Rather than the word matching it gives the exact data and their reference links. The search engines are able to compare or to extract the data and should be giving very relevant results for the queries. The semantic search highly improves search accuracy of the query related data and the search engine delivers the exact content, the user intent to know.

Guha et al categorize the searches into two: navigational and research. The navigation type the user uses the search engine as a navigational tool to navigate through the intent document. But in research type of search the user will be giving an phrase which be denoting the object which the user searching for. As far considering the semantic search engines and the semantic search it will not come under the navigational search because the searching page itself it will be delivering the necessary data to understand about the searched query. That means there will not any search for a particular document. The search is done to get the data or the knowledge about the particular thing, for that a number of documents will be analyzed and suggested to the user by the semantic search.

The semantic web is the framework which helps the user for publishing, reusing the data and sharing the data across applications, community boundaries and enterprises. It’s a W3C approach which is a layered structure. The two layers at the bottom give the foundation, using XML for syntax and URI for standards. The middle layers represent the concepts, properties and the individuals based on the RDF, RDF schema and OWL. The topmost layer extents to represent the inference rules and the logic frame work and tools. The semantic web vision and its core ontology overcome the defects of other searches by having the query related searching. It considers the queries which we can relate by the semantics which means meaning.

Way to represent the difference between the traditional search engines and the semantic search engines. DuckDuckGo is one of the semantic search engines. Here we are searching for the result of the word apple and the result is [9]
It is giving all possible answers with its short description. By seeing this the user can have the basic idea about the apple and which meaning is he searching for. It tells that the apple can be the fruit bank, Software Company or the multimedia corporation. All these possibilities are there. To get this much data in the conventional search engines we have to go through at least three or four link. And that links might be bit confusing. At first it giving the possible things the apple can mean to. Then it comes to the description of those things in one sentence.

4. Methodology

Resource Description Framework (RDF)

It is the foundation of the semantic web. It is a standardized language by W3C. Semantic web consists of web of data that is the data collection is done rather than the document collection. The RDF is the main concept which helps for this. It rates the data given in the different resources and by comparing the hyperlinks[10].

Mapping between keywords and concepts

It is a common approach in the semantic search engine. It has got several advantages that are generally the available data may not be formally encoded. For example in fuzzy keyword to concept mapping deals about the mapping of the textual materials to the well-defined information. In other way is the natural language system, natural languages are in the form of expressions which is mostly acceptable by the humans. As far as considering the mapping patterns as we map from graph to sentence the visual query tool like SEWASIE gives most accurate picture to the humans about the relationships. The humans also will be more comfortable with the natural language since that is more comfortable to them[7].

Graph patterns

It is an important concept in semantic search. In semantic search it is used in multiple varying roles. It is used to solve or encode the complex constraint queries given by the user, it is solved by locating the corresponding graph in the RDF network. The common RDF pattern Anyanwu and sheth is also used to have connection between the paths and the common names. Graphs patterns also gives the idea that where to collect or to fetch the data for particular item.

Logics

It is internally very much tied with semantic web. Even the standard web ontology language (OWL) is based on the description logics. Still the application built on the top of the logical frameworks with the wine agent is an exception than the common ordinary example. In most of the cases the applications take few entailments for the base and create functionality according to their requirements over that. The examples GRQL, ODESeW, SHOE, and SEWASIE do the same.

Combining uncertainty with logics

While augmenting the text search with the ontology technique, to combine the uncertain annotation some formalization could be needed. For this probability logics or the fuzzy logic experiments are under taken in this field.

5. Some of Semantic Engine

Hakia

It is a common search semantic search engine in use. It is well organized by the tabs Web results, crediblesites, images and news. Credible site includes the sites which are vetted by the librarians and other information professionals. For some of the user queries it produces the resumes. They are the portals which gives all information related to that subject. For each resumes there will be an index of links which refers to the corresponding pages, it helps for a quick reference. According to the query the content of the resume will vary. Resumes are one the most impressive feature of hakia. Hakia will also provide the related queries also, which help the user to reach the target or to get the thing very easy [5].

SenseBot

It gives a summarized accurate search result according to the query given. The search engine itself tries to understand the concept of the query, actually what it contains and will give an appropriate result. To do this it
make use of the text mining on the web pages which results on the queries to categorize the actual semantic concept[2]. This summarized result helps the user to get the result of the content very fast. The most important fact is that the answer will be relevant and precise. The effort put by the user can be lessen, not need to go through so many web pages to get the results. It will be containing the tag cloud relating to the query to the remaining relevant concepts[2]. These sentences will be describing the query and each of the sentences are followed by a related link.

**Powereset**

This helps for a comprehensive view of the thing we search for. We can type our need in the search box as queries. It will be giving the answers. The best thing is that the way it aggregates the information provided by the different resources. It offers another option that is “Factz” it will be there in the search results. It provides a set of suggestions about the query given and also the related queries.

**DeepDyve**

It is one of the most powerful research tools for professionals. It helps in retrieving the content from the web in depth.[6]. This is one of the things it differs from the traditional search engines. Rather than the other users some users are in need of information in depth –most of them will be the students, technical professionals etc. This gives us an option that we can have one single letter to 25000 characters in a single query. Since it is a search in depth it will be giving a complex answer but it will give option to sort, save and refine the searches. Still it gives easy navigation.

**Factbites**

The logo itself contains the sentence “where result make sense”, as all the semantic search engines the result of the things will make the sense of knowing the things very closely and accurately. It gives the meaningful relevant sentences to the user at the first page itself. It used to search for matches on the whole topic area, will not stick on to specific keywords. It’s fine in filtering out the sites which are not relevant.

**Lexxe**

It is also a widely used search engine. In this most of the answers are taken from the other sites on the internet and the unstructured texts. It is giving the results on the meaning basis rather than any other basis. A fully automated third generation search engine which is now in alpha version. It is also one of featured language processing search engine.

6. Effectiveness of Semantic Search Engines

- **Cares about the concept matching**

Conceptually the results should match with the query searching for. The web contains almost all the data in the world it should be categorized properly to get the exact result for the search. Not even in the spellings but by concept also it should match with the searched query. it is done in semantic search engine. As we consider the search engine duckduckgo…………………. 
Generalization

It is another important functionality the search engines must do. The query must be so general and the users need specific centralized answer as answer.

Care about morphological changes

It deals with the change in the term while typed by the users but the meaning is same so it should give the exact result if we take an example the term ‘recruit’ the result should not differ for recruit-recruits-recruited-recruitment and recruiting. This should be taken care by the semantic search engines.

Considers about the known items

It also takes care about the popularity of that word. That means the result is arranged according to the relevance[1]. For example if we search “duck” we get a small description about the bird which commonly known ,second preference goes to Donald duck,then comes the choice of animals, transportation, computer software etc..

Care of the synonyms

It will categorize the synonym in the correct sense. It should not be necessary the user is aware of particular synonyms in that cases the semantic search engine helps the user to get the maximum available relevant result according to the searched query.

Cares about the query format

The semantic search engines sensibly respond to the queries raised by the users. This made easy search for the users. It is able to give exact answers to the queries starting with why,who, what, when, where, how many,how much etc...

It deals about the knowledge matching

It updates or gets the related thing about the query searched. It reveals the related-connected-embedded knowledge about the query the user searching for.

Able to have self-testing or the performance

The semantic search engines are able to deliver the performance score about the web pages. It exactly tells who popular the page is.

No specific format for the queries

The user is having the freedom to write the knowledge according to his/her wish .The search engine is able to categorize logically and deliver the result. The engines are able to process and find logically matching contents.

It does not depend on the user behavior, statistics or any other artificial things.

Semantic search engines are capable of giving the output or search results by analyzing the input query on the basis of data. It is capable of deliver information according to the user behavior and the previous searches.

7. Semantic Search Engines Traffic

Alexa ranking helps to study about the traffic and which obviously reflects the usage of semantic search engine is growing rapidly and needs to be made still more effective. It is depicted in the graph below.
8. Conclusion

By this we can conclude that the semantic search engines more convenient from the user view. Nobody have time to lose by searching the needed content in this fast life that is the area the semantic search engine gets the chance. There are so many semantic search engines available. A survey based on the semantic search engine is done. Still there are areas the differences which are to be measured. One of the main works will be based on the architectural differences between the conventional and semantic search engines.

9. References


[3] The Anatomy of a Large-Scale Hypertextual Web Search Engine, Sergey Brin and Lawrence Page, Computer Science Department, Stanford University, Stanford, CA 94305, USA


[7]SemSearch: A Search Engine for the Semantic Web, Yuangui Lei, Victoria Uren, and Enrico Motta, Knowledge Media Institute (KMi), The Open University, Milton Keynes


[10]Context Path Traversal in RDF Graphs 5-24-06 Notes Glenn Takanishi