

A Survey On Meta Search Engine in Semantic Web

G.Sudeepthi¹, Prof. M.Surendra Prasad Babu².

¹Assistant Professor, ²Professor.

sudeepthigovathoti@yahoo.co.in, drmspradbabu@yahoo.co.in.

Abstract

The Search engines plays an important role in the success of the Web, Search engines helps any Internet user to rapidly find relevant information. But the unsolved problems of current search engines have led to the development of the Semantic Web. In the environment of Semantic Web, the search engines are more useful and efficient in searching the relevant web information., and our work shows how the fundamental elements of the meta search engine can be used in retrieving the information resources in a more efficient way. Meta search engines that utilizes the power of a traditional search engine and enriches the search result using the knowledge base to produce better results, In this paper we made a brief survey on the concept of meta search engines, and focused on the architecture and their key technologies involve., finally summarized various semantic meta search engines developed so far such as SavvySearch, Metacrawler, hot bot, harvest42, dogpile e.t.c, in their own strategy.

A search engine will go through the contents of the web based on the behaviour of the user and tries to provide all the “answers” based on the known content. In the first stage, when the World Wide Web has just started to grow, search engines only focused on the ability to index the web resources and then answer user query based on keyword search a powerful search engine that is capable to return a high precision result is required. The emergence of Google has led to a new standard in search engine, the importance of quality in searching. Google, with its unique way of ranking pages in the web has become the de facto standard for searching information on the web. Google still has some limitations since it is a searching mechanism that is based on keyword. Both the indexing of web pages and the matching of user query do not consider the meaning within it. But Semantic search engine searches are based on the meaning of the input query rather than considering a query as a group of string.

1. Introduction

The World Wide Web (WWW) allows the people to share the information (data) from the large database repositories globally. We need to search the information with specialized tools known as generically search engine. There are many search engines available today, but retrieving meaningful information is difficult. However to overcome this problem in search engines to retrieve meaningful information intelligently, semantic web technologies are playing a major role. The Semantic Web is an extension of the current web in which information is given well-defined meaning. Semantic search has the power to enhance traditional web search, but it will not replace it.

A metasearch engine is a search tool that sends user requests to several other search engines and/or databases and aggregates the results into a single list or displays them according to their source. Metasearch engines enable users to enter search criteria once and access several search engines simultaneously. Metasearch engines operate on the premise that the Web is too large for any one search engine to index it all and that more comprehensive search results can be obtained by combining the results from several search engines. This also may save the user time from having to use multiple search engines separately.

Metasearch engines are used to describe the paradigm of searching multiple data sources in real time. It has its own virtual database. They do not compile a physical database or catalogue of the web.

Instead, they take a user's request, pass it to several other heterogeneous databases and then compile the results in a homogeneous manner based on a specific algorithm.

No two metasearch engines are alike. Some search only the most popular search engines while others search lesser-known engines, like newsgroups, and other databases. They also differ in how the results are presented and the quality of engines that are used. Some will list results according to search engine or database. Others return results according to relevance, often concealing which search engine returned which results.

Search engines that search for meaning than words will make the web more manageable. Search engines frequently have different ways they expect requests submitted. For example, some search engines allow the usage of the word "AND" while others require "+" and others require only a space to combine words. The better meta search engines try to synthesize requests appropriately when submitting them.

2. Architecture of semantic meta search engine:

The architecture of the semantic web meta search engine³ are mainly composed by user interface, member search engine selection module, query forwarding module and result integration module. Fig.1 shows their basic architecture.

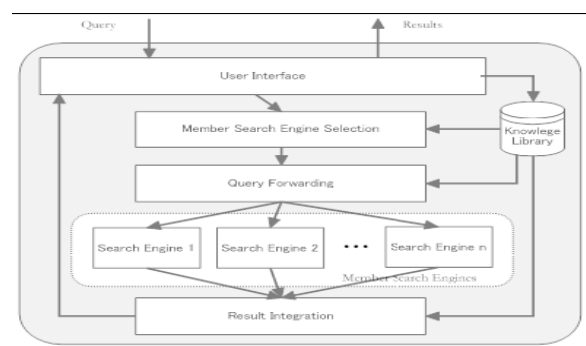


Fig 1: Architecture of semantic meta search engine

(a) User interface

In this module users post their query through a form, while search engines feedback to users a list of

target entries users choose the entries possibly meeting their information needs¹. It provides customized settings such as the number of entries displayed per page, the search scope of language, etc to fully meet users' preferences.

(b) member search engine selection module

They assess their member search engines and choose the ones performing well on a specific subject, reducing the response time and improving user experience. But it is not easy to choose appropriate member search engines⁴. The main strategies for selecting member search engines are User Selecting, Fixed Members, Intelligent Selecting.

(c) query forwarding module:

Each member search engine has its own system of grammar rules and retrieval parameters, therefore meta search engines separately transform the query request received from the user interface to formats supported by each member search engine. But query forwarding should be carried out in parallel to enhance the work efficiency².

(d) Result integration module:

Each member search engine will return a list of references. These returned references return them to users in a uniform format. Duplication removing and reference resorting are the main two occupations of this module.

3. Technology involved in meta search engine:

Reference resorting², which is considered as one of the key technologies of meta search engines. The performance of a meta search engine can be determined to large extent by this technology. Three typical resorting algorithms are shown in the figure 2.

| Algorithm type | Strategy involved | Drawback |
|--------------------|---|------------------------------------|
| Simple Combination | It simply appends one member search engine returned eferences | It overlooks the low-ranking ones. |

| | | |
|---------------------|---|--|
| | to another one's. | |
| Round Robin | It is an interval combination strategy that sorts the member search engines by their performance. | It requires more attention to repetitious pages when taking this algorithm |
| Relevance Resorting | It calculates the global relevance degree based on each reference's evaluation score, location in reference lists, occurrence number, member search engine's performance or other factors, and resorts all the references by this degree. | It needs to take into various factors into consideration which consumes more amount of time. |

Fig 2: Types of restoring algorithms

4. Advantages of meta search engine:

There is a new trend in the market where people seem to be slowly understanding the power of meta search engines over the traditional search engines. No search engine in the world can retrieve information from the entire web and using multiple search engines³ simultaneously isn't easy. Meta search engine solves this problem as they send queries to different search engines and aggregates the results into a single list. Some of the advantages of the meta search engine

- Meta search engine instantly derive and compile listings in response to your keyword enquiries from other search engines. This saves time than searching them in individual search engines.
- It benefits the user by eliminating duplicate hits and grouping the most relevant ones at the top of the list and it will save our precious

time when we are looking for results in different search engines.

- They take the best out of the individual search engines which makes the combined result much more comprehensive than searching them individually adding richness of the data retrieved.
- Many meta search engines gives us the option to choose the list of search engines where we want our query to be saved.
- They provide high on privacy when compared to normal search engines like Google which keep on tracking Internet activity if you are logged into a browser using your Google account.
- Meta searching is an excellent tool when our purpose is to search any topic and get it's overview

5. Disadvantages of meta search engine:

- Translation of query syntax and fields are not exact.
- Variable quality of search engines.
- We Must keep up with the changes in member search engines.

6. Summary of various meta search engines

A tabular form is given in the below table which summarizes some important meta search engines that are developed so far are shown in the figure 3:

| Type of Meta Search engine | Description |
|----------------------------|---|
| SavvySearch | It accesses around a dozen different search engines and returns all the results. It offers search forms in over 25 languages. |
| Brainboost | It provides specific answers to questions asked in natural language. It is best suited if questions that have been already discussed somewhere on the Web in one form or another. |

| | |
|-------------|--|
| DeeperWeb | It allows navigating through search results. The technique involved is tag cloud technique. |
| MetaCrawler | It provides users the option to search for images, video, audio. |
| Yolink | It provides highly structured data in a highly unstructured environment and the results can be utilized immediately to make decisions. |
| Clusty | The technique is to organize numerous search results into several meaningful categories called clusters. We can arrive at interesting search results that may be too far down in the ranking. We can view similar results together in folders rather than scattered throughout a list. |
| HotBot | It searches the Internet for user documents stored on a hard drive. |
| Harvester42 | It integrates several distributed data sources. It queries over 30 major search engines in parallel and presents a large result page with the individual search engine results. |
| Excite | It is a collection of Internet sites and services and offers online service for a variety of content. |
| dogpile | It fetches results from several other popular search engines, including from audio and video content providers. It filters for duplicates and then presents the results to the user. |
| Ixquick | It is the first search engine to delete private details of its users like IP addresses and other personal information are deleted within 48 hours of a search. It does not share its users' personal information with other search |

| | |
|--|---|
| | engines or with the provider of its sponsored results. It uses a "Star System" to rank its results - by awarding one star for every result that has been returned from a search engine. |
|--|---|

Fig 3: Summary of various meta search engines

7. Conclusion and Future Work

In this paper we made a brief survey on meta search engine and the key technologies of meta search engines through which the performance of a meta search engine can be determined. We summarized various meta search engines that are developed so far. We hope, this paper gives a broad overview of meta search engine. We concluded that Meta search engine solves the problem for user as they send queries to different search engines and aggregates the results into a single list. As a future work, improvement over the existing search engines with better results can be achieved by removing the limitations particularly in terms of time, accuracy of result and relevancy of result. An efficient semantic web search engines should meet out these challenges efficiently and allow end users to ask complex queries and provide comprehensive means to handle them and provide quick response to user queries.

8. References

- [1] Wenliang Zheng, "An Intelligent Agent-based Meta Search", 2010 International Conference on Future Information Technology and Management Engineering.
- [2] Lin Guoyuan, Tang Jiutao "Studies and Evaluation on Meta Search Engines" 2011 IEEE.
- [3] Dr. M. Rajaram, "Web Caching in Semantic Web Based Multiple Search Engines", 2010 IEEE.

