

Semantic Web Search

Research Paper



Submitted to

Mr Santosh Khanal

Department of Computer Science and Engineering
Kathmandu University
Dhulikhel, 5 Kavre

Submitted by

Milan Thapa (52)
Sem 7
milans.thapa78@gmail.com

Contents

1	Introduction	1
1.1	Search History before Semantic Web	1
1.2	Semantic Web —an Intelligent Web	1
1.3	Why Semantic Web Search ?	2
2	Available Semantic search portals	3
3	Vision for Future	4

1 Introduction

1.1 Search History before Semantic Web

From the invention of the internet and web, the informations and data are being accumulated, refined and distributed all over the world. Up to par the web have been thorough its different versions. **Web 1.0** was all about the *information web*, where people just collected the data over the network. It gets advanced to **Web 2.0** as a *social web* when people started getting connected through social networking like facebook, linkedin. Meanwhile the data accumulation rate over the web gets its exponential hit and the process is still being growing progressively. The search engines during the days of web 2.0 were just programmed to serve the user's query with some algorithms of page ranking like *Goolge PageRank in old days to predicate relevant result, before its knowledge graph*. With those chaos of the information all over the web, a user using some searching techniques like: the user provides a phrase which is intended to denote an object of which the user is trying to gather information about. But the poor search engines can't exactly resolve the phrase user is looking for, it is because the engines really don't have any idea(clue) of what they are dealing with. They were just programmed without any intelligence to analyse the meaning of a simple phrase. So with no alternatives, the user tries to locate a number of documents or pages which together will give him/her the information s/he is trying to find. *i.e. every time a user needs a information about something, s/he becomes an explore, exploring the web, filtering the contents, merging all together and finally drawing some desired meaning out of it.*

So that is how **semantic web** emerged as an extension of the current web by standards and technologies that help machines to understand the information on the Web, to support richer discovery, data integration, navigation, and automation of tasks.

1.2 Semantic Web —an Intelligent Web

Semantic web —an intelligent web is a web designed to solve all the chaos of informations and data spreading all over the current internet. It's an evolving collection of knowledge, built to allow anyone on the Internet to add what they know and find answers to what they want to know. Information on the Semantic web, rather than being in natural language text, is maintained

in a structured form which is fairly easy for both computers and people to interact and work with. The mechanism of linking the search engine to the semantic web give rise to the new terminology called semantic web search or so called semantic search. Its a search engine designed to improve search accuracy by understanding searchers intent and the contextual meaning of terms as they appear in the searchable dataspace (on the web), to generate more relevant results. As an explanatory example

When searching for "laptop", then one is looking for laptops or synonyms / related concepts (such as "notebook"), but also for special kinds of laptops that are not synonyms / related concepts, such as e.g. IBM/Lenovo ThinkPads. That is where web needs an intelligence and reply the desired results.ie semantic search.

One important goal of the semantic web is to make the meaning of information explicit through semantic mark-up, thus enabling more effective access to knowledge contained in heterogeneous information environments, such as the web. Semantic search plays an important role in realizing this goal, as it promises to produce precise answers to users queries by taking advantage of the availability of explicit semantics of information. Another example could be like this: when searching for news stories about "UNG students", with traditional searching technologies, we often could only get news entries in which the term "UNG students" appears. Those entries which mention the names of students but do not use the term "UNG students" directly will be missed out. Such news entries however are often the ones that the user is really interested in. In the context of the semantic web, where the meaning of web content is made explicit, the semantic meaning of the keyword (which is a general concept in this example of UNG students) can be figured out. Furthermore, the underlying semantic relations of metadata can be exploited to support the retrieving of information which is closely related to the keyword. Thereby, the search performance can be significantly improved by expanding the query with instances and relations.

1.3 Why Semantic Web Search ?

The effective exploration of the web content involves with the numerous challanges like:

Useability The user on a web searching for something does't care how the structured query is specified. The end user often does not know the

query language and the underlying data scheme. For example searching pdf,docs,tex in google one merely don't know that use of tag file-type:pdf/doc/tex could be handy.

Scalability As the amount of available data is ever growing, the ability to scale the most desired results is essential.

Imprecise Information Needs The information needs expressed by the user might be imprecise. An effective search solution should be able to consider this aspect to deliver relevant results. For example: a user might search phrases like vdo for video,fbk for facebook,anime for animation,b4 for before, bday for birthday, etc.

Data Change The Web of Data is continuously changing. Thus, efficient mechanisms for index update at the Web scale are needed when data changes.

So understanding the user's intent by analysing the web resources knowledge scattered over the internet is the first...

2 Available Semantic search portals

Some available cool semantic search portals are tabulated below:

Search Portals	Description
Bing	particularly its new Knowledge Graph feature first semantic search engine on the Internet(2002)
Google	
GoPubMed	
Hakia	semantic search engine with realtime annotator plugin/bookmarklet which adds a smart layer to every website
iGlue	
Kosmix	social media semantic search
Lexxe	beta in early 2011
Swoogle	food and recipe semantic search image semantic search
Yummly	
Rendipity	

3 Vision for Future

A number of semantic search tools have been recently developed. Present semantic search tools like are enhancing the performance of traditional search technologies, they are however not suitable for naive users, i.e. ordinary end users who are not necessarily familiar with domain specific semantic data, ontologies, or SQL-like query languages. The semantic search engines should be implemented in such way that it could address these issues:

- the search engines should tackle the problem of knowledge overhead by supporting a Google like query interface such that the proposed query interface provides a simple but powerful way of specifying queries and also help avoid the undesired typing mistakes.
- it should address the problem of existing semantic-based keyword search engines by supporting complex queries. It should provide comprehensive means to make sense of user queries and to translate them into formal queries, thus being able to produce precise results that on the one hand satisfy user queries and on the other hand are self-explanatory and understandable by end users. For example support for two or more search phrases separated by space or some characters: *poverty Nepal NorthAmerica SouthKorea*
- it should be properly designed and tested so that it serves promising results that is so quick, clean and clear.

We have given a brief overview of approaches to semantic search on the Web, 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 semantics to Web queries and content when performing Web search. It is through this integration that the reasoning capabilities envisioned in Semantic Web technologies are coming to Web search and the Web. As we have seen, the formulation of queries and their results in 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 very beginning, and many

open research problems still persist. Some of the most pressing research issues are maybe:

1. how to automatically translate natural language queries into formal ontological queries, and
2. how to automatically add semantic annotations to Web content, or alternatively how to automatically extract knowledge from Web content.
3. how to create and maintain the underlying ontologies.

and the answer to last question may be done either:

- manually by experts, e.g., in a Wikipedia like manner, where different communities may define their own ontologies.
- 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 ontology fragments, and/or from existing annotations of Web pages in microformats or RDFa), or
- semi-automatically by a combination of both above.

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 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/or soft (weighted) constraints and (conditional) preferences (e.g., similar to [24]), 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 desired enviroment.