Semantic Search: A Brief Survey
Gladys Gnana Kiruba. B
1Department of Computer Science and Engineering
1VIT University, Tamilnadu, India

Abstract—The success of the web depends on the amount of information it provides to the user. The Semantic web is the extension of the World Wide Web that facilitates users to share content beyond the limitations of applications and websites. This paper proposes a survey on semantic search to support the researchers and designers in understanding the advances in search technology with that of classical web search. It also provides some technologies, features and evaluation related to the semantic search.

Key words: Semantic Search, Traditional Search

I. INTRODUCTION

In this age of internet, information and communication systems play a key role in both the internal and external control systems as they form the basis of whole WWW. Nowadays digitally stored information is growing and the communication grows not only between the computer and human but also between computers. The information needs to be integrated and the knowledge needs to be managed. The semantic gap [6] during the communication must also be bridged. The technologies related to information retrieval to form a knowledge base and communication is shown in fig 1.

The semantic web describes the relationships between things and the properties of things. Tim Berners-Lee who is the originator of www defines semantic web [19] as “The Semantic web is an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation”. The semantic technology functions ranges represent knowledge and meaning, discover, organize, interoperate, integrate, interpret, infer, reason, provision, present, communicate, act using semantic etc. The technology of semantic includes ontology, RDF, RDF Schema which are all used to define the formal description of the given domain knowledge. The ontology [15] is a formal, explicit specification of a shared conceptualization. RDF [5] is a language for representing information in the World Wide Web. Search the web is not easy. The World Wide Web provides a large junk of information from which the required information has to be obtained. Semantic search seeks to improve search by understanding searcher intent and the contextual meaning as they appear in the searchable data space, whether on the web or within a closed system, to generate more relevant results. Taking the current research semantic search forms a wide range of usage which uses semantics.

This paper provides a brief overview in challenges in traditional search in first section followed by functions of semantic search and categorical features of semantic search engine as second and third sections. The fourth sections talks about the technologies of semantic search followed by the evaluation and applications.

II. CHALLENGES IN TRADITIONAL SEARCH

Nowadays search forms the basis of both the research and business. The web search like Google is most widely used. They form a dominant role among the broad spectrum of promising ideas starting from the algorithms to retrieving information. But the question is “does it provide the relevant information what the user actually wants”. Mostly from lots of search results obtained for the query, only the first top ten results are analyzed and used. But it may also be irrelevant from the user’s expectation. Also number of problems [10] is encountered in enhancing the quality and maintenance of the web details which may be specific to the domain. The problems may include no definite formulation, considerable uncertainty, complex interdependencies, and incomplete, contradictory and changing requirements. The web search is also faced by algorithmic challenges which are the main constraints in indexing the page result. To all it’s the semantic search is widely attracted by large community of people as it increases the search accuracy and reasoning, and works on the contextual meaning of terms. The relevancies of search results are increased by semantic intelligence by changing the natural language processing to disambiguation queries.

Semantics is the meaningful computing. Search and semantics rely on similar technologies and have similar goals. To make information more findable and usable both apply the data structuring techniques. Join the search and semantics and you get semantic search, in essence, which make search easier, smarter, accuracy and understanding of context. Taking the variety of work published recently into account, the semantic web have the five following core functional parts as: Semantic tagging [1], Linguistics Natural language Processing [20], Statistics [8], Semantic Web Metadata / Ontology [17][16][13], and Artificial Intelligence[7]. The step by step procedure for semantic search is given as 1) Syntax for knowledge representation (RDF) [5]. 2) Ontology description (OWL) [13]. 3) Web-service description (OWL–S) [3]. 4) Tools for reading and
III. CATEGORICAL FEATURES OF SEMANTIC SEARCH ENGINE

A broad range of techniques and approaches for semantic document retrieval has been built in the framework of the Semantic Web. In [2] Christoph Mangold introduces an ordering scheme for semantic search engines and clarifies terminology. Based on this comparison, he identifies not only basic concepts and outstanding features, but also opens issues. The Mangold classifications are listed below.

A. Architecture:
As standalone maintain a concept index of document and for Meta search use subroutine search engines.

B. Coupling between documents and ontologies:
Tight coupling is obtained by data of documents refer explicitly to the concepts of the specific ontology where as in Loose coupling they are not committed to any ontology.

C. User interaction:
They are transparent by semantic capabilities are invisible to the user on the other hand they are interactive by asking for clarification and recommendation. In some cases both can be combined and used.

D. User content:
Learning can be done by interacting dynamically to the other end they are hard –code by query category.

E. Query modification:
It can be done manually by the user modification or by optimizing the query of the system through query rewritten or by graph based traversal algorithm.

F. Ontology construction:
It can be anonymous, standard property or Domain-specific property.

G. Ontology technology:
It includes RDF, OWL etc.

IV. TECHNOLOGY FOR SEMANTIC SEARCH

In [18] the author Seth Grimes categories the semantic search as related searches, reference results, semantically annotated results, full-text similarity search, search on semantic/syntactic annotations, concept search, ontology based search, semantic web search, faceted search, clustered search and natural language search. On the other hand in [14] the author Makela categorized the search methodology as:

Augmenting traditional keyword search with semantic techniques: More specific ontological techniques are used, i.e., Terms are expanded to their synonym and meronym sets. Direct ontological Browning is supported. The intention is to find related concepts as the writer of the document.

A. Basic Concept Location:
The main goal is to locate instances of the core semantic web formed by concepts, relationships and instance. Users can choose the class of instances by means of ontological navigation.

B. Complex Constraint Queries:
Many Semantic Searches with this approach are based on navigating the ontology as the last approach. One way is based on a global intersection of distinct selectors, constraining do not need to be ontological.

C. Problem Solving:
The Semantic search use the ontological knowledge to solve a problem; searching for solutions by inference and other reasoning techniques.

D. Connecting Path Discovery:
In this the semantic search are based on the ideas of a vast amount of varied semantic data will be available to be mined for semantic connections.

V. EVALUATION OF SEMANTIC SEARCH

The wide spread availability of machine understandable information has the potential to deeply impact on the semantic search. Hildebrand et al. [11] conducting a survey on the role of semantics in current end user search applications. The systems are evaluated on the basic phases as query construction, search algorithm and presentation of results.

<table>
<thead>
<tr>
<th>Search phase</th>
<th>Feature</th>
<th>Functionality</th>
<th>Interface</th>
<th>Components</th>
</tr>
</thead>
<tbody>
<tr>
<td>Query construction</td>
<td>Free text input</td>
<td>keyword(s) natural language</td>
<td>Single text entry</td>
<td>Property-specific fields</td>
</tr>
<tr>
<td>Operators</td>
<td>Boolean operators special purpose operators regular expressions</td>
<td>Application-specific syntax</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Controlled terms</td>
<td>Clear and concise input output restriction predefined queries are selected</td>
<td>Value list Faceted Graph</td>
<td></td>
<td></td>
</tr>
<tr>
<td>User feedback</td>
<td>Pre-query elucidation</td>
<td>List of suggestion Semantic auto completion</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Search algorithm</td>
<td>Syntactic matching</td>
<td>Exact, prefix, substring match Minimal edit distance Stemming</td>
<td>Not applicable</td>
<td></td>
</tr>
<tr>
<td>Semantic matching</td>
<td>Thesaurus expansion Graph traversal</td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Semantic search is a broad task as it is more than the simple document search and for its diverse application. It gives hand to multi-disciplinary research and development through natural language processing and text mining. At last it is highly user-friendly as it provides vertical search among domain ontologies, visualization and mobile searching.

VI. APPLICATIONS

The application of semantic search has been wide spread and nowadays lots of semantic search engines are available. In this era there are a number of differing approaches working towards the common goal of providing a search experience that benefits from meaning and context. Basically semantic search applications are classified as Bottom-up approach and Top-down approach. The bottom-up approach involves embedding semantical annotation (meta-data) right into the data. Powerset and Hakia are the examples of this approach. The Top-down approach is a system which analysis the existing information; the final top-down solution would be a fully blown natural language processor, which the people can understand like a normal text reading. Spock, a vertical search engine is the example of Top-down approach. Generally semantic search can be classified as General search (Freebase, Yahoo, Microsearch,...), Natural Language search (Powerset, Hakia, Askmenow,...), vertical Search (Kango, Adaptiveblue, Reportlinker,...), Social Networking Search (Semantinet, Delver,...) and personalized search (Twine, MavinIT ps,...). As common the semantic search can be used from any range of domain such a library application to the most complicated domain as medical field.

VII. CONCLUSION

Semantic search is a broad tasks as it is more than the simple document search and for its diverse application. It gives hand to multi-disciplinary research and development through natural language processing and text mining. At last it is highly user-friendly as it provides vertical search among domain ontologies, visualization and mobile searching.

REFERENCES


Table 1: Evaluation of Search Process


