INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY

WINGS TO YOUR THOUGHTS.....

An Approach to Semantic Search Engine

Palak Shah[1], Bhargavi Kongari[2], Ronak Dedhia[3]

[1]Student of Computer Engineering Department
[1]K. J. Somaiya College of Engineering
Mumbai, Pin no.400077
palak.shah@somaiya.edu
[2]Student of Computer Engineering Department
[2]K. J. Somaiya College of Engineering
Mumbai, Pin no.400077
bhargavi.k@somaiya.edu
[3]Student of Computer Engineering Department
[3]K. J. Somaiya College of Engineering
Mumbai, Pin no.400077
ronak.dedhia@somaiya.edu

Abstract: *Search engine is the way to find out any kind of information on web. This paper provides the model of a semantic search engine through which desired information can be located on the web. This model of the search engine makes use of semantics that is meaning of queries, to retrieve relevant results. Model is divided into four layers, namely Semantic Web, Data Layer, Application layer and User Interface layer. This paper presents an approach to semantic web engine in a specific domain that is restaurants in Mumbai. The Semantic web provides semantics through application of semantic web technology –Ontology. Data layer represents information retrieved about restaurants. Application layer gives the logic to retrieve data using ontology and database. User interface allows user to enter query and displays result of searched query.*

KEYWORDS*: Semantic web, Ontology, Sematic search engine, Information retrieval.*

1. INTRODUCTION

The internet comprises of various kinds of services, web page repositories, resources. It is hence very difficult to find the desired information from this huge amount of information. Search engines like Google, yahoo uses indexing. It indexes the key words of the web pages and information is searched based on these keywords.[1] The issue in this way of search engine is that it has no semantics associated with it. Semantic Web is a collection of elements which includes languages representing formal and informal knowledge, tools and standards and methods for semantic web applications. Ontologies are used to store knowledge about some domain of interest example this paper presents the cuisine ontology merged with location ontology. Anontology describes the concepts and its relationships in that domain. Ontology hence provides the semantics. The use of semantics, the machines can perform automated search of the required query. Hence this kind of semantic search engine provides a better method to locate, refer, retrieve data and provide result. To achieve this, a number of methods, tools and standards are researched and proposed in last few years.

2. WHAT IS SEMANTIC WEB?

Tim Berners-Lee, the inventor of the World Wide Web and director of the World Wide Web Consortium ("W3C"), defines the Semantic Web as "a web of data that can be processed directly and indirectly by machines"[3]. According to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries".

The content for human usage is primarily consisted by WWW. The Universal Resource Locator (URL) links various contents on WWW. Usually the semantics are inferred by the users. Web content normally consists formatting instructions for an agreeable presentation which again for human use. There is no kind of formal logical constructs in WWW content. Consistently, the Semantic Web mainly involves the statements for application usage. The statements are linked together via constructs that forms semantics which is meaning of the link. Thus, well-defined *meaningful* path is delivered by the link semantics, rather than a user-interpreted one. Semantic Web's main purpose is to drive the development of the current Web. It enables users to search, share, and combine information more easily. Human beings are skilled to use the Web to carry out tasks such as finding the French translation for "seven days" or reserving a library book or searching for the lowest price for an electronic device. As web pages are designed to be read by people and not by machines, the machines cannot accomplish all these errands without human guidance. The semantic web is an idea of information that can be readily inferred by machines. Machines can perform more tiresome work that is involved in combining, finding and acting upon information on the web. Semantic web makes use metadata.

"The Semantic Web is an extension of current web in which information is given well-defined meaning, better enabling computers and people to work in co-operation." A Web with a meaning is called semantic web. The description of things is done in a way that computers can understand it. It is an extension to the normal Web. [6] Conventional Web consists of human operator and uses computer systems for various tasks like searching and aggregating but Semantic Web is understood by computers and it searches, aggregates and combines the information without a human operator. It is easily processible by machines, on a global scale. It efficiently represents data on the World Wide Web.

3. WHAT IS ONTOLOGY?

"In computer science and information science, an ontology is a formal naming and definition of the types, properties, and interrelationships of the entities that really or fundamentally

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

exist for a particular domain of discourse. It is thus a practical application of philosophical ontology, with a taxonomy" as defined by Wikipedia.[4] Ontology compartmentalizes the variables needed for some set of computations and establishes the relationships between them. Ontology not only helps in understanding the concepts of the domain but also helps the machine to interpret the definitions of concepts in the domains and relations between them. It characterizes information in a manner such that this data can be used by machines not only for display but also for automation, integration and reprocessing the same data across various applications. A common vocabulary for researchers who need to share information or data in a domain is defined by ontology. It consists of machine-interpretable explanations or definitions of basic theories in the domain and relationship amongst them.

4. ARCHITECTURE

On the basis of the current Web search engines and the features of the Semantic Web, we present the architecture of Semantic searching in hotel website in Fig.1. The system should be able to run some reasoning engines to identify the necessary facts and rules so as to achieve the desired conclusions.

4.1 Ontology Creation:

Ontology is a representation of knowledge which defines set of concepts using classes, their properties and relationship among them. These concepts provide the semantics for the application.[3]

Reasons to develop an ontology are (1) Sharing common understanding about structure of information among people or software agents, (2) Reusing domain knowledge, (3) Making domain assumptions explicit,(4) Separating domain knowledge from the operational knowledge, and (5) Analyzing domain knowledge.

Ontology is created using Protégé 4.1.2 which is a free, open-source ontology editor. Two ontologies food ontology and location ontology is created.

4.2. Ontology merging:

The single merged ontology would include all the knowledge the merged ontologies i.e. Food and location ontology had. Protégé has a built-in feature to merge ontologies which is used to merge the ontologies.[5]

4.3. Restaurant database:

The database contains information about restaurants such as address, contact, photos, etc. The answer to user query is a list of restaurants from the database which is retrieved using the semantics from the ontology.

4.4. User query:

The user enters a query related to the domain in natural language. The query is transformed into keywords by eliminating stopwords and stemming it. The keywords are then processed and the result is returned to the user.

4.4. Processing:

The keywords are first searched in the ontology using Jena API to obtain semantics. Jena API makes use of SPARQL i.e. *SPARQL Protocol and RDF Query Language* query to retrieve semantics from ontology. The result of SPARQL query is then fed to JDBC program. The program searches it in the database using SQL i.e. *Structured Query Language* and list of restaurants satisfying user query is retrieved. The result is then displayed to the user through GUI.

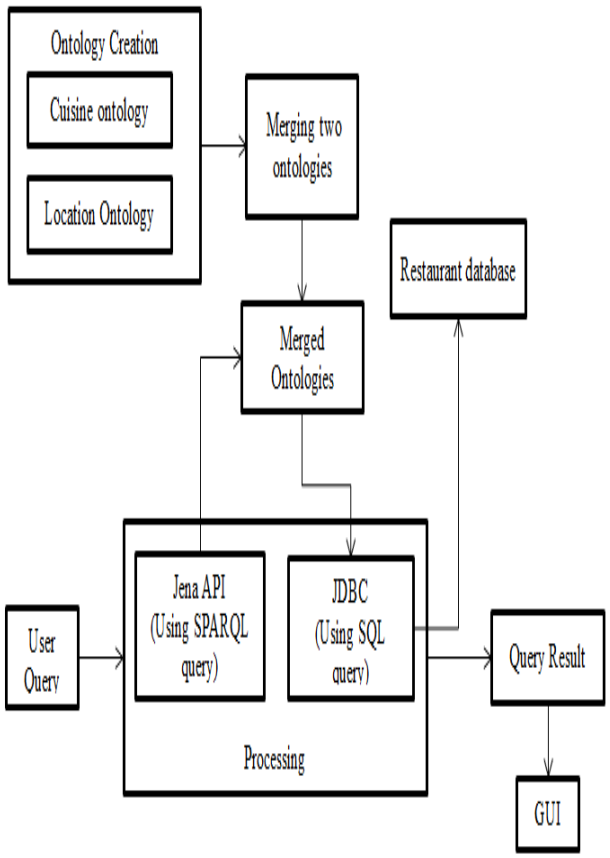

Figure 1: Architecture

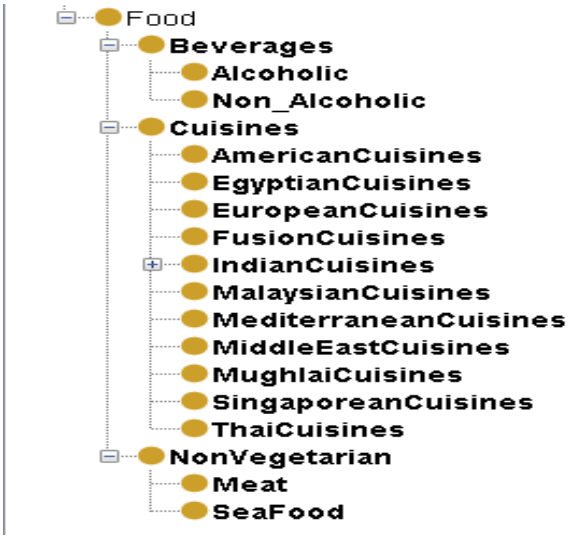

Figure 2: Food Ontology

INTERNATIONAL JOURNAL FOR ADVANCE RESEARCH IN ENGINEERING AND TECHNOLOGY
WINGS TO YOUR THOUGHTS.....

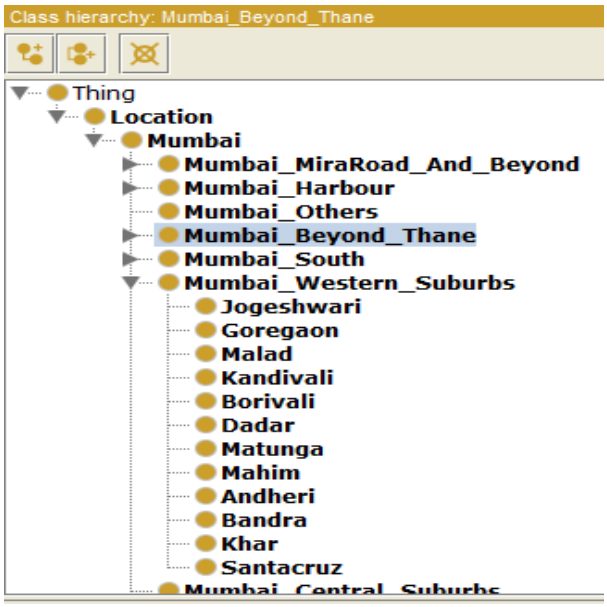

Figure 2: Location Ontology

```
SPARQL query:
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX owl: <http://www.w3.org/2002/07/owl#>
PREFIX xsd: <http://www.w3.org/2001/XMLSchema#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT ?subject ?object
        WHERE { ?subject rdfs:subClassOf ?object }
```

Figure 3: Example of SPARQL query

5. CONCLUSION

The quality of web is enhanced with the use of semantic web. Information retrieval is enhanced by Semantics enhances. In this paper, we have investigated the meaning of semantic web, use of its application that is ontology through protégé software. Our observation that the traditional World Wide Web doesn't incorporate the semantics became an inspiration to survey for semantics. Survey had been done to find out the limitations of the traditional web and architecture of the process to build semantic search engine is proposed. Through the paper characteristics of semantic web are presented along with the attributes of semantic search engine. The process of building semantic search is proposed in such a way that it optimizes the search in terms of time and speed both. This proposed search engine mechanism based on semantics allows knowledge derivations and retrieves results based on user's queries by utilizing the semantics and relation between the information sources. Currently our focus is to develop the semantic search engine for a particular domain which is restaurants in Mumbai, keeping in mind the survey information. We would acknowledge the guidance we received from our professor Mrs. Pradnya Gotmare.

REFERENCES

[1]. Kamath S.S., Piraviperumal D.,Meena G., Karkidholi S. ,Kumar K.: "A semantic search engine for answering domain specific user queries" Communications and Signal Processing (ICCSP), 2013 International Conference.

[2]Esa A.M., Taib S.M. ,Hong Nguyen Thi: "Prototype of semantic search engine using ontology " Open Systems (ICOS), 2010 IEEE Conference .

[3]. Yi Jin, Zhuying Lin, Hongwei Lin: "The Research of Search Engine Based on Semantic Web ".

[4]. Lipyeow Lim, Haixun Wang, Min Wang "Semantic Data Management: Towards Querying Data with their Meaning", vol 1438-1442, IEEE 2007.

[5]. Noman Islam1, Abu Zafar Abbasi2 and Zubair A. Shaikh "Semantic Web: Choosing the Right Methodologies Tools and Standards", vol 1-5, IEEE2010.

[6]. Hazber Mohamed, Yang Jincai, Jin Qian "Towards Integration Rules of Mapping from Relational Databases to Semantic Web Ontology", Computer Science and Technology Department Huazhong Normal University Wuhan, China, 2010 International Conference on Web Information Systems and Mining.

[7] Ponnada, Sharda N "Model of a Semantic Web Engine for Multimedia Content Retrieval" Computer and Information Science, 2007.