VOICE ENABLED ONTOLOGY BASED SEARCH ENGINE ON SEMANTIC WEB FOR BLIND

N.Karthik,

Assistant Professor – Department of Information Technology,
Adhiparasakthi College of Engg., Kalavai, Vellore, Tamil nadu, India.
Email:nkarthikapce@gmail.com

M.Ashwini,

B.Tech (IT),
Adhiparasakthi College of Engg., Kalavai, Vellore, Tamil nadu, India.
Email:ashwinichinna03@gmail.com

K.Anitha,

B.Tech(IT) ,
Adhiparasakthi College of Engg., Kalavai, Vellore Tamil nadu, India.
Email:anigowsh@gmail.com

Abstract---With the development of web services, the retrieval of relevant services has become a challenge. Keyword based discovery mechanism is insufficient, due to retrieval of large amount of irrelevant information. Here we propose a semantic search engine to conquer this problem. Semantics with meaning and use of data brings information closer to human thinking and decision making. In this work, we have described voice enabled ontology based search engine for blind. Here we use conceptual based search engine. This search receives the user query through voice and converts them into text with the help of software; this text is progressed in the ontology list and produces the exact result using semantic web. The displayed text is converted into voice.

Keywords: search engine, ontology, semantic web, RDF

I. INTRODUCTION

Semantic search helps the user queries to be understandable for electronic agent searching. In this way, ontologies play main role to define the semantic and relationship between the user queries. Therefore, by adding paradigm ontology in order to interpret user queries and their relevant documents in safety and security domain, we enable user with using fewer number of terms to gain information. The main objective of this study is to improve the usability and efficiency of semantic search using ontology in order to enrich the user queries and gain user satisfaction in resultant search with input and output as voice. Semantic simply means meaning. Meaning enables a more effectual use of fundamental data. Semantic web is a web of data illustrated and linked in ways to establish framework that hold to defined grammar and vocabulary constructs.

The Semantic Web [1] allows data to be shared and reclaimed across applications. It predicts a globally interlinked network of machine- processable data, made possible by the distribution of semantic data models, which is also known as ontology. The Semantic is based on the Resource Description Framework (RDF). Semantic Web technologies are still in the early process of development. One of the most important issues is locating suitable information using existing ontologies for capturing the user-required information from the Semantic Web. For example, if one wants to publish their top ten colleges in Semantic Web, they would like to find some ontologies representing things like "area", "college name", and "courses offered". Otherwise, they should build their ontologies by themselves.

The Resource Description Framework (RDF) is a unit of W3 prototype that normally enables the exchange of records and metadata. RDFs are mainly used in describing resources in semantic web activity. RDF based data model is more suitable for ontological models. RDF is a standard model for data interchange on semantic web.

In this work, a semantic search engine is designed especially for blind. This search is an optimized search which is more essential for blinds. The input is got in form of voice and the voice is converted to text thus the converted text is searched through ontology list. The relevant results are displayed and they are converted in form of voice for blinds.

A . ORGANIZATION OF THE PAPER

The rest of the paper is organized as follows: section 2 deals with related works in semantic web. The proposed model is described in section 3. Finally some concluding remarks and future works are given in section 4.

II. RELATED WORKS

Swoogle [2] is an implemented system that analyzes, discovers and indexes knowledge programmed in semantic web document. It reasons about document and records meaningful metadata about them. It is an ontology project undergoing constant development. It also provides a customizable algorithm inspired by google's page rank algorithm. One of the interesting properties computed for each semantic web document is its ontology rank- a measure of document's importance on semantic web.

In [3] author makes a brief survey of existing literature regarding intelligent semantic search technologies. Semantic web search is a search engine for semantic web. This intelligent search engine for semantic web technologies has combined description of reasoning inference system and digital annals ontology to complete intelligent search engine.

In [4] they describe that, while many techniques are adopted by Personalized Abstract Search Services (PASS), which is used to search abstract of research papers. This paper only describes the construction of fuzzy ontology and its use for query refinement. Fuzzy ontology provides information about sets of terms with broader and narrower meaning. Fuzzy ontology is built in two stages, using information obtained from system's collection.

Stage- 1: Creating full ontology from fuzzy narrower term relations.

Stage- 2: Eliminating unnecessary terms from the above created full ontology.

In this paper [5] a semantic search engine featuring multi –scoping is introduced. The proposed architecture has several parts,

1. Domain construction
2. Automatic ontology creation for the domain.
3. Annotation generation for document and queries.
4. Customization of general ontologies for specific domain and context- aware expansion.

In this paper [6] has presented a semantic based user model called user ontology, for supporting personalized application in semantic web. This model utilizes idea, taxonomic and non- taxonomic relations in a given specific domain ontology to confine the user's attention. The proposed user ontology model has been incorporated into semantic search engine and applied to document retrieval. Thus relatively specific user ontology can be described within a shorter time.

In this article [7] presents an efficient and focused crawling strategy which considers a number of distributed crawlers which recover relevant pages to specific topics. Crawlers are functioning in favor of vertical search engine, which is providing higher accuracy. The experimental results show efficiency of our strategy. The relationship between the number of processors and time required to complete the crawling process is almost linear.

In this paper [8] describes, as the web is exponentially expanding, it is very hard to find interesting information directly from web sources using keyword- based search engine. One approach explored in certain research prototype or commercial system is to extract domain- specific information from relevant web sites and semantically integrate on structured or semi- structured ontology access. This system integrates wrapper text mining technology.

Text mining technologies are,

1. Wrapper technology: Extracting interesting information from various sites. A schema- guided toolkit is used for generating wrappers.
2. Text mining technology: Indexing, categorization, summarization and key word extraction.
3. Xml server technology: This displays the document in appropriate format.

In this paper [9] has presented an approach towards mining semantic web data, focusing on clustering objects described by ontology- based metadata. The overall clustering approaches have applied on real world data, namely the CIA world fact book. Therefore we will model country clusters within the CIA world fact book ontology and experiment to which degree the algorithm is able to discover these country clusters.

To bring the semantic web [10] that provide enhanced knowledge and information services, we need professional ways to access and bring knowledge from web document. Artequakt automatically extracts knowledge about artists from web, populates a knowledge base. We setup the knowledge extraction tool to select the first relevant pages for each artist, and analyze.

Oil [11] is properly grounded in web language such as, xml schema and RDFs, it offers different levels of complexity and at least the inner layers enable efficient reasoning support. This paper deals with precisely this necessity. OIL is a suggestion for facilitating the semantic web, i.e. information with machine processable semantics. It is based on existing application such as RDFs and XOL, and improving them with necessary features for expressing rich ontologies. The paper describes the modeling primitives, underlying rationale, syntax, motivation semantics, tool environment, and applications of OIL.

III. PROPOSED MODEL

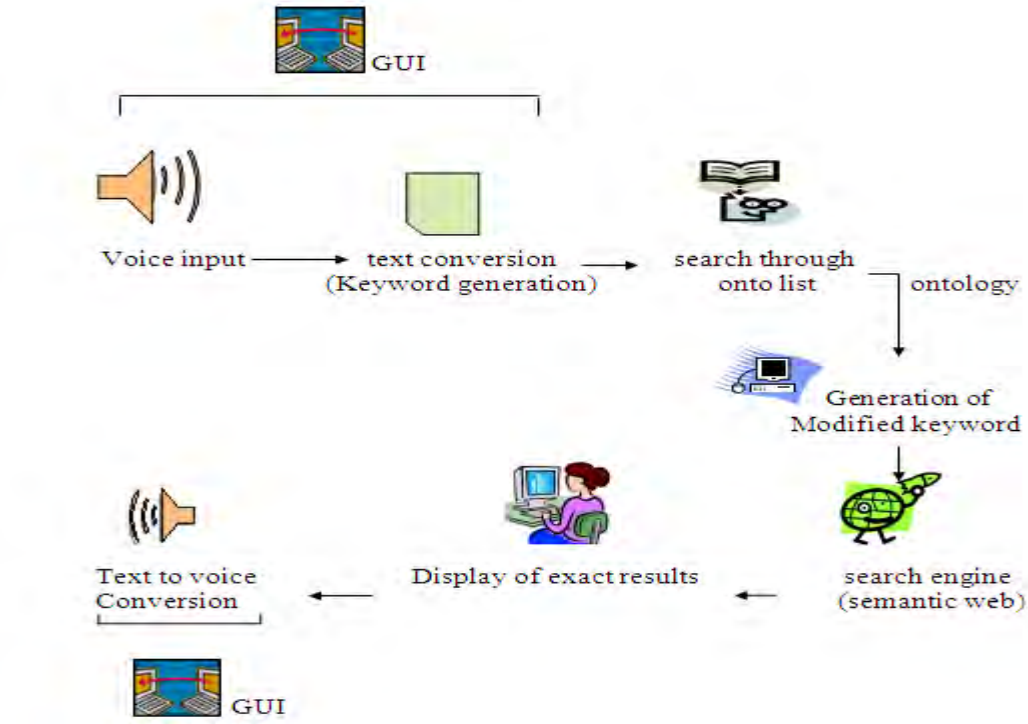

Fig 1. Work flow of proposed model

A. GRAPHICAL USER INTERFACE (GUI) CREATION

User interface is defined as an intermediate between the user and the machine. Using this interface the user can interact with machine. User interface is started by layout creation. Layout which is a type of resource defines what is drawn on the screen. This is simply a template for the interface screen sometime for entire portion of screen. Layout's primary purpose is organizing the controls of device. Here using mic, voice input is got from the user (blind) to process their query. Layout is designed using eclipse IDE. Views: The *Views* are the visible elements in the User Interface. Using views we can add text and button. Here textview attributes are used.

B. VOICE TO TEXT CONVERTER

Speech-to-text conversion is the development of renovating spoken lexis into written content. This process is also called vocabulary detection

Working principle

All speech-to-text systems contain two models: an auditory model and a speech model. In toting up large glossary systems use an articulation model. Universal vocabulary detection is understood that it is an important mania to the user. To get the finest transcription eminence, all of these models be capable of specialized for a given vocabulary, dialect, application province, type of speech, and communication conduit.

Auditory model: This model describes the probabilistic actions of the encoding of the linguistic information in a vocal signal. The most leading approach uses continuous mass hidden Markov models (HMM) to illustrate context dependent phones.

Speech model: A speech model captures the regularities in the spoken language and is used by the vocabulary detection to guesstimate the probability of words series. This model attempts to confine the syntactic and semantic restrictions of the language by estimating the frequencies of n words.

C. SEARCH THROUGH ONTO LIST

Ontology creation using protégé software

Ontology is a description of a domain. Ontologies can be written in various formats and can be used by computers to reason about the domain they describe. They are also useful as a common format that allows for exchange of knowledge across applications / platforms. The ontologies will be working which are written in OWL 2. OWL 2 is an ontology language that defines the concepts you can use to write ontology.

Onto search

It is a web-based system, which can offer online service. The user inputs keywords (in form of voice) to describe the nature of the required ontology. Then, OntoSearch will apply the Google Web APIs to search the Internet for relevant files (the file type is restricted as RDFs which later can be changed) and return all the relevant URLs on the screen. The displayed URLs (output) are turned out in form of voice to user.

D. TEXT TO VOICE CONVERTER (OUTPUT AS VOICE)

Text To Speech renovate, also called "Speech blending" is the artificial fabrication of human speech. A Text to Speech engine makes over any text into speech in real time. It factually reads out loud any given information with a smooth and natural swishing voice.

IV. CONCLUSION

Since the existing search engine are basically text-based. They cannot be used by the blind people. This search engine is not only efficient for blind people but can also be competent for illiterate people, since this project is based on voice.

According to the above premise, this study presents a prototype of semantic search engine based on ontology, used to optimize the search process. A domain of relevant document collection has been used in the process of constructing the semantic search engine, which is related to the area of safety and security. This ontology is so flexible such that domain of study can be adapted to different domains easily.

V. FUTURE WORK

The approach of this work is to design a prototype of semantic search engine. In this basis, the paradigm ontology needs to be improved based on domain in the future work. We have created the ontology list for a specific domain. Also the proposed ontology can be substitute to different domains, which is mostly helpful for blind.

VI. REFERENCES

[1] Sean B. Palmer, The Semantic Web: An Introduction, 2001-09
[2] Swoogle: A semantic web search and metadata engine. X L. Ding et al. In Proc. 13th ACM Conf. on Information and Knowledge Management
[3] Madhu, G., Govardhan, A. and Rajinikanth, T.V. (2011). “Intelligent Semantic Web Search Engines: A Brief Survey”. International journal of Web & Semantic Technology (IJWesT), 2(1).
[4] D.H.Widyantoro and J.Yen, “A Fuzzy Ontology-based Abstract Search Engine and Its User Studies”, In Proceedings of the 10th IEEE International Conference on Fuzzy Systems, 2001.
[5] Kyumars Sheykh Esmaili, Hassan Abolhassani, Mahmood Neshati, and Babak Bagheri Hariri,” mosse: a multi ontological semantic search engine ”.
[6] Xing Jiang, Ah-Hwee Tan, “learning and inferencing in user ontology for personalized semantic web search”, Information Sciences Volume 179, Issue 16, 20 July 2009, Pages 2794–2808, Elsevier.
[7] Rodrigo campus, Oscar Rojas, Mauricio Marin Mendoza, “distributed ontology- driven focused crawling” 21st Euromicro international conference on parallel distributed and network based processing, 2013.
[8] Georges Gardarin, Huaizhong Kou, KarineZeitouni, XiaofengMeng, Haiyan Wang: SEWISE: An Ontology-based Web Information Search Engine
[9] Maedche.A, Zacharias.V: “Clustering ontology- based metadata in semantic web” in proceedings of PKDD 2002.
[10] Harith Alani, Sanghee Kim, David E, Millard, Mark J, Weal, Wwendy Hall, Paul H Lewis, Nigel R, Shadbolt, IEEE intelligent systems 2003, “Automatic ontology- based knowledge extraction from web document”
[11] Fensel D., van Harmelen F., Ian Horrocks, Patel Schnieder P.F. and D. L. McGuinness, “OIL: An Ontology Infrastructure for the Semantic Web.”, IEEE Intelligent Systems, vol. 16 issue 2, pp 38-45.