ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 9, September 2014

Enhancing Social Personalized Search Based on Semantic Search Log using Ontology

K.Maheswari [1], Dr.S.Kirubakaran [2]

Assistant Professor, Department of CSE, SNS College of Technology, Coimbatore, India[1]

Assistant Professor, Department of CSE, INFO Institute of Engineering, Coimbatore, India[2]

ABSTRACT: As the Information available in the Internet is vast, the search engine provides search results based on page ranks. But the search results are not related to one particular user's environment. But it is possible to provide customized search to each user with semantic technologies. Semantic Web is to add semantic annotation to the Web documents in order to access knowledge instead of unstructured material, allowing knowledge to be managed in an automatic way. A system called as Semantic Search log Social Personalized Search would be able to provide results for search query that relates to a particular user's environment, the data that the user might have found to be useful while searching. In this system, supervised learning technique is used to learn about the user. Semantic web search is applicable for each and every registered user in this application. In the proposed work, ontology search logs are used, which will be used for providing customized search logs according to the user defined input.

KEYWORDS: Semantic Web Ontology; Semantic web search; Semantic search log.

I. INTRODUCTION

Ontology is the study of the categories of things that exist or may exist in some domain. The product of such a study, called as an ontology, is a catalog of the types of things that are assumed to exist in a domain of interest D from the perspective of a person who uses a language L for the purpose of talking about D. The types in the ontology represent the predicates, word senses, or concept and relation types of the language L when used to discuss topics in the domain D. 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 co-operation. In our proposed work one of the semantic web technologies such as ontology is used to provide customized search to the users. Social networks are the domain in which we could obtain user oriented information, which we use for providing personalized search results. We use supervised learning technique to learn about the user based upon his interactions inside the system. This process can be able to make applicable for each and every registered user in this application. User can give their basic information in their profile and get benefits from their each and every search.

II. REVIEW OF LITERATURE

Pedrinaci C, Domingue J, Brelage C,Van Lessen T,Karastoyanova D, Leymann F have proposed that semantic BPM, that is, the enhancement of BPM with semantic Web services technologies, provides further scalability to BPM by increasing the level of automation that can be achieved. They have described the particular SBPM approach which enhances BPM solutions in order to achieve more flexible, dynamic and manageable business processes. Web services technologies, provides further scalability to BPM by increasing the level of automation that can be achieved [9].

Ko R.K.L, Ko K L, Lee S.G proposed Genesis - a Web application which formulates business process definitions dynamically, given a user business goal and underlying business criteria. Genesis is a standalone module which provides dynamic capabilities using hierarchical ontology. In this paper they have demonstrated how a business user can, via our proposed Genesis application, rapidly and dynamically formulate abstract business process definitions based on High-level business goals (e.g. buy, sell) and Business criteria (e.g. total order cost, type of sourcing methods, etc). There are three main components in the Genesis application architecture such as Web-based Graphical

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 9, September 2014

User Interface (GUI) and Server. Web-based Graphical User Interface (GUI) is a Java Server Page (JSP) page which interacts with the user and passes information to the server. The components of Genesis are housed in a Java Servlet which comprises Genesis algorithm and BOWL. Genesis Algorithm is an extension of hierarchical task network (HTN) planning methodology in artificial intelligence planning, this algorithm dynamically chains tasks required for particular business collaboration. Business-OWL (BOWL) is hierarchical task network ontology of business tasks rooted in OWL (Web Ontology Language). This ontology stores a hierarchy of compound and primitive tasks, and the possible permutations of decomposing a compound task. The algorithm traverses the hierarchies of tasks to find the tasks which match the particular business collaboration. Before understanding the roles of each component in Genesis, it is beneficial to first discuss the roles of and interactions between individual components [10].

V. M. Hoang and H. H. Hoang proposed a new approach called Ontological Hierarchical Task Network (O-HTN) based on HTN Planning and Web Service Modeling Ontology (WSMO) for forming collaborative business processes dynamically for the cross-enterprise collaboration. They have followed Semantic Web-based approaches for BPM which acts as a promising solution with taking advantages of Semantic Web technologies such as ontologies, semantic web services [6].

Ko R.K.L, Lee S.G , Lee E.W, Jususf A proposed Hierarchical task network (HTN) planning which is used for dynamically creating CBP task sequences ideal for direct Web service execution. In this paper they have established the rationale behind modeling business-to-business (B2B) collaboration tasks as hierarchical Web ontologies. To demonstrate the achievability of dynamic cBP formulation, they developed the genesis methodology, which consists of (1) business-OWL (BOWL) - a B2B hierarchical task Web ontology, and (2) the genesis algorithm - an extension of the hierarchical task network (HTN) planning algorithm to handle business criteria and control flows commonly found in business processes. In this paper, author introduced an algorithm for Dynamic formulation of cBPs which is an extension of the Hierarchical Task Network (HTN) planning algorithm. This algorithm dynamically formulates cBP definitions on the fly given criteria such as cost of product, type of B2B collaboration. And also, a decomposable BP modeling technique is used which is a novel method to model collaborative business processes (cBP's), and to decompose high-level compound tasks into low-level (operational) primitive tasks. Business-OWL (BOWL) is introduced. BOWL is an ontology which stores the cBP hierarchical task decompositions, actors, and control flows in B2B collaborations. HTN's store relationships between compound tasks, primitive tasks, and methods in ontology. Many traditional techniques of HTN planning usually store knowledge, the problem and domain descriptions in the form of text-based files (e.g.the Planning Domain Description Language (PDDL). In contrast to XML-based files like OWL, these formats are usually not adept for current Web-based processing and may not be directly suitable for the increasingly Web-dependent business environments. And the current techniques adopting classical HTN tools (e.g. SHOP2, XPlan) require prior translation into the classical domain descriptions before planning can happen. Therefore, the knowledge captured in HTN's needs to be modelled in Web-friendly formats that facilitate Web based applications. OWL is the ideal candidate in this respect. The information cannot be merged with other files to create new knowledge in the older representations of HTN's. Modelling HTN's as ontologies in OWL will allow dis-similar ontologies to be merged to create new knowledge [8].

Yakup Yildirim, Adnan Yazici, Turgay Yilmaz have proposed a semantic content extraction system that allows the user to query and retrieve objects, events, and concepts that are extracted automatically. They have introduced an ontology-based fuzzy video semantic content model that uses spatial/temporal relations in event and concept definitions. In addition to domain ontologies, they have used additional rule definitions (without using ontology) to lower spatial relation computation cost and defined some complex situations. The proposed framework has been fully implemented and tested on three different domains. There are basically three levels of video content which are raw video data, low-level features and semantic content. Raw video data consist of elementary physical video units together with some general video attributes such as format, length, and frame rate. Second, low-level features are characterized by audio, text, and visual features such as texture, color distribution, shape, motion, etc. Third, semantic content contains high-level concepts such as objectsand events. The first two levels on which content modeling and extraction approaches are based use automatically extracted data, which represent the low-level content of a video, but they hardly provide semantics which is much more appropriate for users. The proposed model is a Meta Model and represented with VISCOM=<V C;DII>. V C is the set of VISCOM classes and DII is the set of domain-independent V ISCOM class

DOI: 10.15680/IJIRSET.2014.0309061

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 9, September 2014

individuals. Each VCx inVCis represented as VCx =< VCxname; VCxprop>, where VCxname is the name of the class and VCxprop is the set of relations and properties of class VCx: VISCOM has a number of classes representing semantically meaningful components of video, where VCxname = { fComponent; Object; Event; Concept; Similarity; . . .}. The linguistic part of VISCOM contains classes and relations between these classes. Some of the classes represent semantic content types such as Object and Event while others are used in the automatic semantic content extraction process. Relations defined in VISCOM give ability to model events and concepts related with other objects and events. VISCOM is developed on an ontology-based structure where semantic content types and relations between these types are collected under VISCOM Classes, VISCOM Data Properties which associate classes with constants and VISCOM Object Properties which are used to define relations between classes. In addition, there are some domain independent class individuals [13].

III. PROBLEM DESCRIPTION

A. Existing system

Cross Enterprise Collaboration is one of challenges on the B2Bi research. Semantic Web-based approaches for BPM have been foreseen as a promising solution with taking advantages of Semantic Web technologies such as ontologies, SWS. A semantic framework can be used for forming a dynamic collaboration of business processes within BizKB systems for across-enterprise collaboration. In the existing system ontology is implemented for Business integration. In the existing system, basic search engine should not produce a customized search. Denial of search attacks may be possible. The Search engine used here produces more junk or advertisement websites at the top level of the search. Here, updates will be possible only through the admin web server of the search engine. And no ontology profile is implemented. The existing system involves with a simple search engine. In a basic search engine, the web crawling is done by several distributed crawlers. There is a URL server that sends lists of URLs to be fetched to the crawlers. The web pages that are fetched are then sent to the store server. The store server then compresses and stores the web pages into a repository. Every web page has an associated ID number called a doc ID which is assigned whenever a new URL is parsed out of a web page. The indexing function is performed by the indexer and the sorter. The indexer performs a number of functions. It reads the repository; uncompressed the documents, and parses them. Each document is converted into a set of word occurrences called hits. The hits record the word, position in document, an approximation of font size, and capitalization. The indexer distributes these hits into a set of "barrels", creating a partially sorted forward index. The indexer performs another important function. It parses out all the links in every web page and stores important information about them in an anchors file. This file contains enough information to determine where each link points from and to, and the text of the link. The URL resolve reads the anchors file and converts relative URLs into absolute URLs and in turn into doc IDs. It puts the anchor text into the forward index, associated with the doc ID that the anchor points to. It also generates a database of links which are pairs of doc IDs. The links database is used to compute Page Ranks for all the documents. The sorter takes the barrels, which are sorted by doc ID, and resorts them by word ID to generate the inverted index. This is done in place so that little temporary space is needed for this operation. The sorter also produces a list of word IDs and offsets into the inverted index.

B. Proposed system

In our proposed system ontology search logs are used, which will be used to produce customized search results according to the user defined input. This application will be processed in any of the search engine. Semantic Search log for Social Personalized Search is used which would be able to provide results for search query that relates to a particular user's environment, user's area of interests, user's likes and dislikes, the data the he/she might have found to be useful for him while searching. Supervised learning technique is used to learn about the user based upon the information given by the user during registration. To provide the best search results to the user, data mining techniques such as classification and clustering is used to get the patterns of the data.

IV. MODULES DESCRIPTION

The modules used for accomplishing personalized search using ontology are

A. User profile creation module

B. Ontology Search engine update module

DOI: 10.15680/IJIRSET.2014.0309061

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 9, September 2014

C. Semantic web search module

A. User profile creation

Authentication of administrator and user is done. Initially, the registration page is created and it is connected with the back end server. And the user should create their account in order to register their details into semantic logs. The information about the user is stored in order to achieve social personalized search. While creating the account user is asked to provide their security information. So that, an unique username and password is generated for each user. Each and every time when the user logins in to their profile for searching, the authentication is done. A Social network is implemented for sharing. Along with that user can create their friends group. So that, the user can share their search with other users.

B. Ontology Search Engine Updates

It involves with admin process where the data centre is available with huge number of data sets consisting of the user profile information. An Ontology table structure is created in order to update the data which will be used when the user logins for searching and information to the user is provided according to the user profile information which is stored in the database. When the user logins for searching the information, the particular user's profile information is retrieved from the database and it is analyzed. After analyzing the data the search results are provided based on the users view.

C. Semantic Web search

The user is provided with the search page where the user can search for the information. And also the user can either do the general search or semantic web search. In case of general search general information which is common to everyone is provided to the user based on the hit ratio. Otherwise if the user is doing the semantic search, particular user must login to the system with their corresponding user name and password. And then the user can continue with the searching process. The search will be done according to the profile information of the user. An ontology interface process is implemented here in order to get the relevant feedback information. The search engine will search the information according to the Meta data as well as according to the hit ratio. Here the database is designed with the users additional information so that search results are based on the search logs as well as hit ratio and meta data. Here the user can get more efficient search results comparing to the general search. And retrieval will not be more complex.

V. CONCLUSION

The design and implementation of efficient and effective Web Search Engines (WSE), is becoming more important as the size of the Web keeps growing. Furthermore, the development of systems for Web Information Retrieval represents a very challenging task whose complexity imposes the knowledge of several concepts coming from many different areas: databases, parallel computing, artificial intelligence, statistics, etc. Other than, business integration ontology can also be used in Search Engine. Thus we have implemented ontology in search engines for providing the best search results to the user.

REFERENCES

1. Akdemir U, Turaga P.K, and Chellappa R, "*An Ontology Based Approach for Activity Recognition* from Video," Proc. ACM Int'l Conf. Multimedia, A. El-Saddik, S. Vuong, C. Griwodz, A.D. Bimbo,K.S. Candan, and A. Jaimes, eds., pp.709712,http://dblp.unitrier.de/db/conf/mm/mm2008.html#AkdemirTC08, 2008.
2. Bai L, Lao S.Y, Jones G, and Smeaton A.F,"Video Semantic Content Analysis Based on Ontology," IMVIP '07: Proc. 11th Int'l Machine Vision and Image Processing Conf., pp. 117-124, 2007.
3. Jung J.J, "Semantic business process integration based on ontology alignment," Expert Systems with Applications, vol. 36, pp. 11013-11020, 2009.
4. Hoang H, Tran P.C, and Le T, "State of the Art of Semantic Business Process Management: An Investigation on Approaches for Business-to-Business Integration," presented at the Intelligent Information and Database Systems, 2010.
5. Minh Tuan Nguyen Hoang, Trung Van Nguyen, Hanh Huu Hoang, "A Ontological Collaborative Framework for Business Process Integration", IEEE 2013 Transactions on Knowledge and Data Engineering ,Volume 24,Issue:6
6. Ko K.L, Lee S.G , Lee E.W, Jususf A, "Dynamic Collaborative Business Process Formulation via Ontologised Hierarchial Task Network Planning", July 2009.
7. Pedrinaci C ,Domingue J, Brelage C,Van Lessen T,Karastoyanova D, Leymann F, "Semantic Business Process Management: Scaling Up the Management of Business process", Aug 2008.

DOI: 10.15680/IJIRSET.2014.0309061

ISSN: 2319-8753

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 9, September 2014

8. Ryan A.J , Ko K.L , Lee S.G, "Genesis – Dynamic collaborative business process formulation based on business goals and criteria," 2009.
9. Stency.V.S, Sureshkumar.R "Semantic Search Log for Social Personalized Search, International Journal of Computer Trends and Technology,volume 11 number3 ,May 2014.
10. Sumitha Thankachan, Dr.S.Kirubakaran, "An Intelligence Based Alert System for Farmers", International Journal of Innovative Research in Computer and Communication Engineering, ISSN(Online):2320-9801SSN(Print): 2320-9798 Vol. 2, Special Issue 3, July 2014.
11. S. G. L. Ryan K.L.Ko, E.W.Lee, Andre Jusuf,"Dynamic collaborative business process formulation via ontologised Hierarchical task network (htn) planning," 2009.
12. www.wisegeek.com/what-is-semantic-technology.htm.
13. Yakup Yildirim, Adnan Yazici, Turgay Yilmaz, "Automatic Semantic Content Extractionin Videos Using a Fuzzy Ontologyand Rule-Based Model" , IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 25, NO. 1, JANUARY 2013 47.
14. Yilmaz T, "Object Extraction from Images/Videos Using a Genetic Algorithm Based Approach," master's thesis, Computer Eng. Dept., METU, Turkey, 2008.
15. Yildirim Y, Yilmaz T, and. Yazici A, "Ontology-Supported Object and Event Extraction with a Genetic Algorithms Approach for Object Classification," Proc. Sixth ACM Int'l Conf. Image and Video Retrieval (CIVR '07), pp. 202-209, 2007.
16. Y. Yildirim and A. Yazici, "Ontology-Supported Video Modeling and Retrieval," Proc. Fourth Int'l Conf. Adaptive Multimedia Retrieval: User, Context, and Feedback (AMR), pp. 28-41, 2006.