INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 1 (2018) PAGES 2742 - 2746
RECEIVED : 15.01.2018 PUBLISHED : 13.02.2018

Feb 13, 2018

AN INTEGRATED APPROACH TO IMPROVE ONTOLOGY MAPPING PROCESS IN SEMANTIC WEB

[1]Sivasankar.C, M.Phil, Research Scholar, K.M.G College Of Arts & Science, Gudiyattam.

Prof.P.Daniel Sundarraj, Head & Assistant Professor, PG & Research Department Of Computer Science & Applications, K.M.G College Of Arts & Science, Gudiyattam.

Abstract:

Ontology Mapping is an important issue, widely recognized in a research community of Semantic Web. Large number of ontologies, developed across the World Wide Web in a distributed manner demands for automatic or at least semi-automatic ontology mapping system to integrate information from different ontologies and to make the vision of Semantic Web reality. The integrated approach here means two things it integrates several techniques from different computational area such as Computational Linguistic, Information Retrieval, and Machine Learning in order to provide semi-automatic ontology mapping process; and It takes care of ontology mapping process right from the creation of the ontologies. The algorithm performs these steps in iterative and interleaved manner depending on its execution configuration. The language processing activities such as Tokenization, Lemmatization, Abbreviation Expansion, Spelling Correction, Elimination of Stop words, etc. are performed whenever required. It uses domain specific thesaurus for abbreviation expansion and synonym. It also uses Word Net, an online lexical database, to strengthen the Linguistic Matcher. The AI - ATOM has been evaluated using the measures precision, recall, and F- measure on two small real world data sets. The system is tested with different algorithm configuration to decide the best possible default configuration for the domain under consideration. The preliminary case studies show encouraging result.

Keywords: Ontology Mapping, ATOM, wordnet, Tokenization, Semantic Web, F-Measure Algorithm, XML/ RDF Framework.

1. INTRODUCTION

The present WWW has huge collection of pages, but majority of them are in human readable format only. As a consequences software agent cannot understand and process this information, and much of the potential of the Web has so far remain untapped. To overcome this problem, delivered at a meeting of the World Wide Web Consortium (W3C), Tim Berners-Lee presented his vision for a new Web, Semantic Web, where machines The Semantic Web, as the name implies, is a Web focused on the conveyance of meaning. It facilitates modeling, sharing and reasoning with knowledge available on the Web through the formal representation of knowledge domains with ontologies The present HTML based WWW defines Web pages syntactically and were intended only for human consumption. The HTML defines how content of the Web pages should be displayed on browser, but does not tell anything about the subject and nature of the content. Hence, these Web pages cannot be read and processed by machines without human intervention to derive any meaning out of it. With the Semantic Web, information on the Web can be defined semantically in such a way that it can be used by machines, not only for display purposes, but also for interoperability and integration The Semantic Web is consisting of a philosophy (idea), a set of design principles, collaborative working groups, and a variety of enabling technologies. Some elements of the Semantic Web are expressed as prospective future possibilities that have yet to be implemented or realized. Other elements of the Semantic Web are expressed in formal specifications and notations such as RDF Schema (RDFS) and the Web Ontology Language (OWL), all of which are intended to provide a formal description of concepts, terms, and relationships within a given knowledge domain. Key technologies in Semantic Web include explicit meta-data, ontologies, logic and inference, and intelligent agent. It uses the information such as names, data types, and domains of element. It also uses tokenization for compound element names before computing element level score.

2. RELATED WORK

For name matching it uses morphological normalization, categorization, string based techniques such as common prefix and suffix; and a thesauri lookup. This component transforms input schemas into trees that enrich the structure by augmenting referential constraints. The linguistic similarity of node and similarity of their leaf nodes are combined to decide similarity between two elements. Finally, it calculates weighted mean of linguistic and structural similarity.

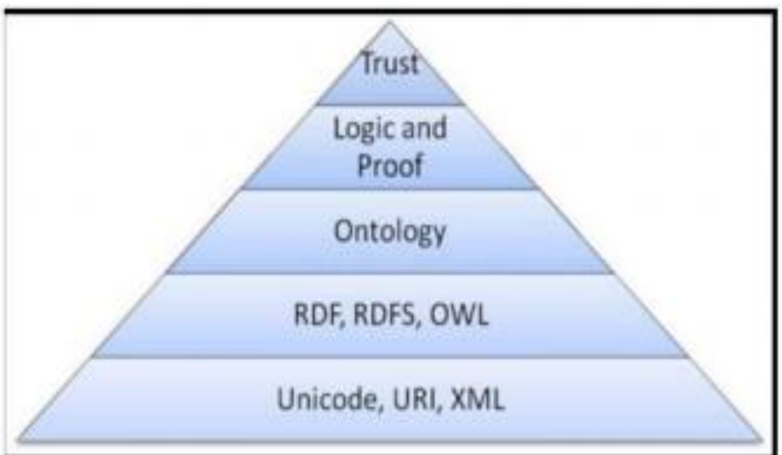

Fig.1.Architecture

Cupid represents the input schemas internally as trees. These trees are processed by name matchers assisted with auxiliary thesaurus to get linguistic similarity at element level and then by structural matchers that uses name and data type information of leaves in an iterative manner with knowledge propagation until the desired threshold is exceeded. It combines linguistic and structural similarities using a weighted sum. Finally, it generates the alignments based on this weighted similarity crossing some threshold values COMA represents the schema internally as a directed acyclic graph where elements are the paths. This is done to capture the contexts in which elements occur. It allows the user interaction to improve the accuracy of matching based on approval of suggested matches or mismatches. It may also be used to evaluate the different combinations of matchers. The main components of COMA are: the Repositoryto persistently store all match-related data, the Modeland Mapping Poolsto manage schemas, ontologies and mappings in memory, the Match Customizerto configure matchers and match strategies.

INTERNATIONAL RESEARCH JOURNAL IN ADVANCED ENGINEERING AND TECHNOLOGY (IRJAET) E - ISSN: 2454-4752 P - ISSN : 2454-4744
VOL 4 ISSUE 1 (2018) PAGES 2742 - 2746
RECEIVED : 15.01.2018 PUBLISHED : 13.02.2018
Feb 13, 2018

A GUI based system AI-ATOM is developed as a prototype for the proposed integrated approach. The system consists of five horizontal components, viz., Ontology Management, Ontology Project Management, User Management, System Configuration and Ontology Mapping Engine; and six vertical components, viz., Language Processing, VSM Engine, Label Matcher, Linguistic Matcher, Structure Matcher, and User Interaction. The brief description of each component is given below. It allows the user to create new ontology, maintain an existing ontology, delete existing ontology, and to import ontology from file containing parenthesized tree. While user creates new ontology, the system suggests the concept labels along with their meaning from the user specified context dictionary (Domain Specific Labels), based on its usage frequency, that are best matching with partially entered labels by user. The context dictionary is maintained with the help oftable. This helps the user from typing and thinking different label from intention of the community of users in the same domain.

3. PROPOSED SYSTEM

This helps a lot when two such on tologies need to be mapped later. Thus, system proposed here tries to improve ontology mapping process right from its creation.This amendment needs a user to deal with better ontology mapping solutions. The proposed technique suggests the layered architecture and elaborates the functionality of lexical, semantic lingual and contextual similarity. SM will be completed with the help of super concepts name. Semantically, overlapping of concepts name will be evaluated by using domain vocabulary. Flow of the proposed technique is when user interacts with system to find some specific concept. The given input consists of source and target ontology.

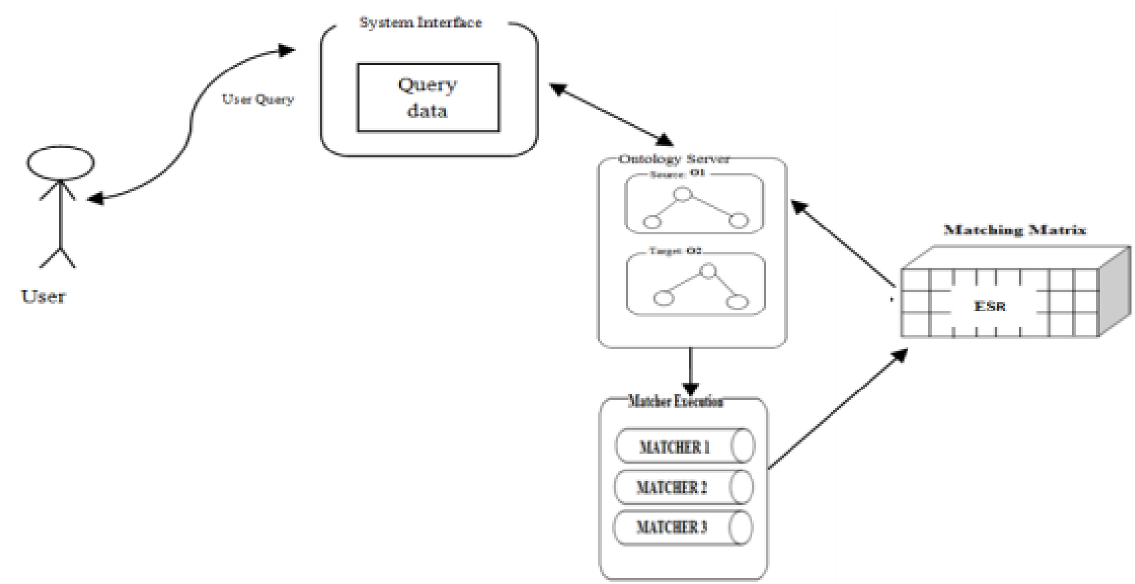

Fig.2.Proposed System

The proposed technique Concept Similarity Mapping (CSM) is used to discover, structure, and demonstrate correlations among concepts, which incorporate properties and roles. The working of CSM technique will be follow layered method. In the start, user pass a query which analyses the similarity testing among concepts and manipulates the refine similarity metric that root on the lexical, semantic lingual and contextual similarity levels. Consider an instance of web-based application of reservation. A user who makes a room reservation for three days and affords to spend 500 dollars .An application of web offers services to explore for hotel reservation including dataset. The user acknowledges a query to explore for reservation is less than 500 dollar. Then web application may requisite to explore the sources of data throughout the web. The proposed system classify into a layered approach where each matching level is performed an individual task and produce highly cohesive .

4. ANALYSIS

Our focal point is to describe the technique of mapping among two ontologies, previous one declared to work with data from ontology to ontology, alike terms of these ontologies is confirm. The suggested method has used to direct the mapping similarity among super concepts along with its role. In which ontology server receives query from user and regain information related to equivalence ontology class. Results show that the matching metrics have exact similarity and increase performance of mapping. Therefore in future this method will enlarge to assemble the requisites of similarity of extraneous concepts.

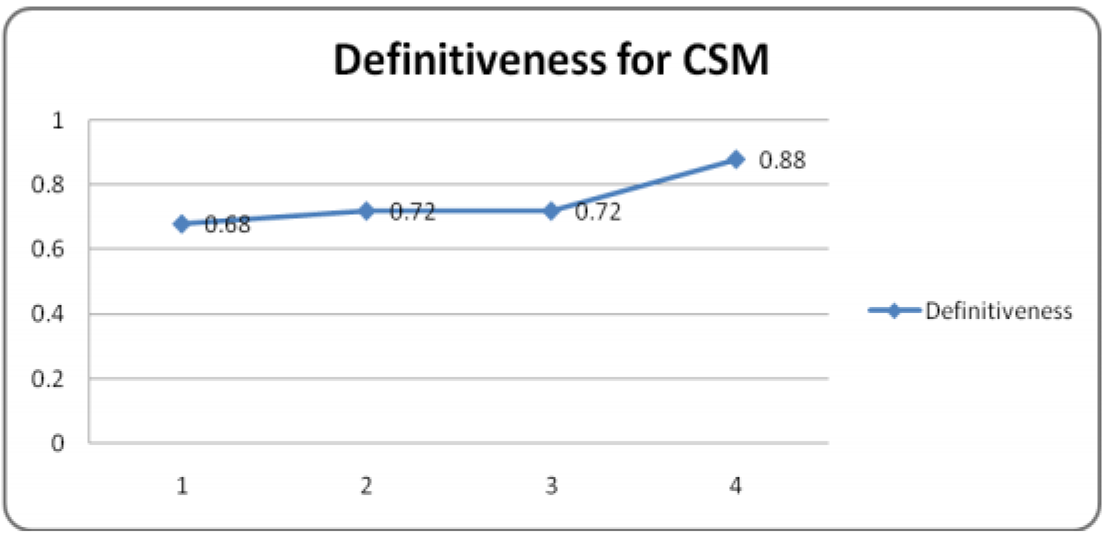

Fig.3.Analysis

So, there would be necessary to do work on its interpretation while merging two concepts into a rare concept. Furthermore, we will emphasis on how to upgrade the functionality of system, and also realize the many-to- many relations among concepts or properties to improve the solution of heterogeneity. With the formation of semantic web, the need of ontologies has significantly increased. Ontologies are well known necessary factor of semantic web, which are defined specification of shared conceptualization. Ontology consists of number of concepts, attributes and associations. Attributes represent the concept as well as association represents relation among concepts. So, Ontologies may refer to reuse, share the knowledge. In order to make the better use of ontologies, similarity matching process between various ontologies is used.

Another new advancement in SM is ontology mapping. Ontology mapping is a methodology in which similar pairs among various ontologies are matched through semantic associations. The mapping of ontology has become entirely applicable for testing the semantics among ontologies. However, users can give query data from various data sources transparently and applications can serve every data source despites of their varying representations. It is considered that various ontologies have alike components, so that there can be certainly used concepts, attributes and relations of components among two or more ontologies.

CONCLUSION

An Integrated approach is proposed that takes care of ontology mapping process from the very first step of ontology creation by allowing the user to set and use domain specific context dictionary and allowing the user to set other domain specific thesaurus such as abbreviation and synonym to provide context information to ontology mapping process that improves automation of ontology mapping process. It effectively combines the power of vector Space Model, used in information Retrieval, to generate potential candidate mapping elements to be selected for further processing. Thus it expedites the process by eliminating large number of weak candidate mapping elements. It proposes the novel approach of using previously rejected mappings to speed- up the process in addition to re-use of accepted mappings.

REFERENCES

1. Abels, S., Haak, L., & Hahn, A. (2005). Identification of Common Methods Used for Ontology Integration Tasks. IHIS'05. Bremen, Germany.

2. Ahmed, A., Volker, H., & Nematollaah, S. (2008). An Empirical Comparison of Ontology Matching Techniques. Journal of Information Science, 1-20.

3. Antoniou, G., &Harmelen, F. v. (2008). A Semantic Web Primer (Vol. 2e). Massachusetts London, England: The MIT Press Cambridge.

4. Beckett, D., &Broekstra, J. (2008). SPARQL Query Results XML Format. W3C Recommendation. World Wide Web Consortium.

5. Borst, W. N. (1997). Construction of Engineering Ontologies for Knowledge Sharing and Reuse. Enschede, The Netherlands:University of Tweenty.