ISSN:0975-9646

A Novel Ontology Processing System Using Information Filter and Knowledge Repository for Semantic Web

Sudha J Vishwakarma

Department of Information Technology
Patel College of Science & Technology (PCST), Indore (M.P), India

Abstract: **Web based information processing systems had exponentially grows due to their technological advancements. It is having billions of web pages which store numerous information which can be retrieved by a simple user queries. Sometimes these queries depend on the content relativity and generate the responses according to the web network. Such network is comes under the semantic web. Here the web process the queries and based on the relative pages holding the similar type of data. The data mapping rules and the relationships are defined using web ontologies. Ontologies increases the relevance of the information which is searched by the user. They primarily create a meta-data model using schema designing with RDF and XML. Both uses tagging process to separates the data. Sometimes the result generated is not related to demanded queries. This is because the information is uncertain and unstructured. There must be some mechanism which can even process the raw data as inserted by the user. For serving this control with traditional ontologies there will be information filter and rules repository for knowledge management. This paper suggests interoperable ontologies definitions with heterogeneous information supports for knowledge management systems. The mechanism is having wide variety of tagging which improves the data classification and rules formation. At the initial stages of work, the analytical results are showing that the work is proceeding in the correct direction and its futuristic proof of concept will shows the same.**

Index Terms- **:** Semantic Web, Ontology, RDF, XML, Knowledge Repository, Information Filter, Logics Representation;

I. INTRODUCTION

In today's world the information and knowledge is the most effective asset of the user. The data which is continuously flowing on the internet is mainly resides behind the World Wide Web. It serves the widely known standards which assures the interoperability between the various levels of the information exchanges. It holds the data of different types which is having some structure or relation between the fields. This is known as structured data. Another is of unstructured type which can't be processed directly. Mostly the data on the internet is unstructured which unnecessary consume the resources. The next generation web must be designed in such a way which provides the machine processible information. That is what the main objective behind the semantic web. It enables the intelligent services over the information processing or knowledge extraction mediums. Some of those services are information brokers, search agents, and information filters. All they provide the robust and interoperable system. Formation of Semantic Web is only possible when the system is been established with high interoperability. It is the new generation of web based network used for effectively representing the knowledge packed in the form of object attributes called triplets [1].

Internet holds the data for web content onto some servers which specifies the searches involve in the knowledge mapping. But in future the data access patterns must be modified in such a way that the knowledge can be delivered as a service using web 2.0. The Semantic Web was made through incremental changes, by bringing machine-readable descriptions to the data and documents already on the Web. Thus by using semantic web the information can be defined in such a way that it can be used for computer in making the web content interoperable and integrated. One way to enable machine-to-machine exchange and automated processing is to provide the information in such a way that computers can understand it. The Semantic Web is not a separate Web but an extension of the current one, in which information is given well-defined meaning, better enabling computers and people to work in cooperation [2].

Operational Levels of Semantics

Semantics is the field of study for deriving the meaning and knowledge using signs, words or terms. Here the degree of semantic can be varied according to the content relativity or terms used for describing the word meanings. While developing the operational levels of semantics it could be separated into four areas: Controlled Vocabulary, Taxonomy, Thesaurus and Ontology. Here the controlled vocabulary is a list of term which is enumerated explicitly with an unambiguous definition. Taxonomy is an subject based classification of controlled vocabulary. A thesaurus is a networked collection of controlled vocabulary terms with conceptual relationships between terms. A thesaurus is an extension of taxonomy by allowing terms to be arranged in a hierarchy and also allowing other statements and relationships to be made about the terms. Last and the most important is Ontologies. They are similar to taxonomies but use richer semantic relationships among terms and attributes, as well as strict rules about how to specify terms and relationships. Ontologies have generally been associated with logical inferencing and recently have begun to be applied to the Semantic Web [3]. Ontologies consist of definitional aspects such as high-level schemas and assertion aspects such as entities, attributes, interrelationships between entities, domain vocabulary and

factual knowledge which are connected in a semantic manner.

This paper covers the various aspects of ontology study and presents some of the work related with the knowledge management using ontologies for semantic web. The next section 2 covers the ontology portion. Later on a literature survey is given in Section 3. Section 4 gives the problem which still remains to be solved for better results. Finally a improved solution to the given problem is given with benefits, results factors and conclusion.

II. BACKGROUND

Semantic web must be defined along with the knowledge management which is a big challenge for organizations. Information is processed to extract the knowledge which could be resides in many forms of World Wide Web. Knowledge management is a strategic task for the organization which involves the rules of extraction. This process is performed in five phases [4]:

(i) Acquisition
(ii) Creation,
(iii) Storage,
(iv) Validation, and
(v) Utilization

Here the semantic web offers more intelligent service of machine understandability of content. Ontologies serve the basic building block of Semantic Web. They provide a common platform for analysing the knowledge resides behind the applications and people's communications. Ontologies themselves can be classified into the different groups: Knowledge Representation ontologies, General/Common ontologies, Meta-ontologies, also called Generic Ontologies or Core Ontologies, Domain ontologies, Task ontologies. In Semantic Web the tools used for understanding the human vocabulary is provides by Ontologies. Mainly it starts with getting the web syntax knowledge and the schematic mapping behind them. The basic tools for providing the above facilities are eXtensible Markup Language (XML) and Resource Description Framework (RDF)**.** XML is having a wide tag library for describing the data and are invisible to peoples who read the documents but are accessible to the web which provides the searches for them. RDF does exactly what its name indicates -- using XML tags, it provides a framework to describe resources [5]**.** A web resource is simply any identifiable information on the web. Uniform Resource Identifiers (URIs) uniquely identify resources of any kind. They are also known as Uniform Resource Locaters (URLs). RDF uniquely identifies the web resources and their location from where the content is initiated. To do this, RDF uses triples written as XML tags to express this information as a graph**.** These triples consist of a subject**,** predicate **and** object, which are like the subject, verb and direct object of a sentence. In order to understand what words mean and what the relationships between words are, the computer has to have documents that describe all the words and logic to make the necessary connections. In the Semantic Web, this comes from schemata and ontologies - tools for helping a computer understand human vocabulary.

Ontologies are made working with the help of schematic definition given with the knowledge representation systems. They organize the information with the RDF tags and XML. They process the metadata related with the descriptor for the document or content. The trouble with ontologies is that they are very difficult to create, implement and maintain. Depending on their scope, they can be enormous, defining a wide range of concepts and relationships. During the last few years various tools is came in the market which provides ontologies of various systems. Each tool represents certain functionalities. The layering allows for a modular design of applications that bundle some or all of the functionalities provided [6]. Some of them are covered here as given below:

- RDF Vocabulary Description Language schema (RDFS) - RDFS adds classes, subclasses and properties to resources, creating a basic language framework.
- Simple Knowledge Organization System (SKOS) - SKOS classifies resources in terms of broader or narrower, allows designation of preferred and alternate labels and can let people quickly port thesauri and glossaries to the Web.
- Web Ontology Language (OWL) - OWL, the most complex layer, formalizes ontologies, describes relationships between classes and uses logic to make deductions. It can also construct new classes based on existing information. OWL is available in three levels of complexity -- Lite, Description Language (DL) and Full.

.Some specified product based ontologies tools are RDF*ferret,* OntoShare, Spectacle and OntoEdit.

III. LITERATURE SURVEY

During the last few years various researches had been suggested towards improving the ontologies processing structures. Ontologies serve the logical relation of deriving the knowledge from the given inputs. Some of the respective work is given below as surveyed literature.

In the paper [7], ontology alignment is performed using multiple ontologies using semantic web based RDF documents in a heterogeneous manner. Mainly it allows the different ontologies to operate interdependently. It derives the knowledge form the relation using more than one set of rules. For getting the work in right direction, the queries must return some values of knowledge, dataset terminologies and their alignment information. This alignment matching involves uncertainties of knowledge rule and their interpretations. This paper serves the objective using distributed inference and probabilistic reasoning to allow efficient ontology alignment.

In the paper [8] ontologies are used for knowledge integration for the different organization using their rule base repository. Here the representation of knowledge structure must follows the standardization rules of semantic web. Mainly the knowledge extraction matches certain section rules from which the relativity is measured. Thus such system is known as knowledge management system. The KMS system deploys the networked knowledge workers which work towards processing the entire

information throughout the lifecycle of knowledge. It serves the dynamic requirements of the organizational learning and other assets using semantic web ontologies. But there is a limitation with ontologies that it can't be used for all the web domains.

In the paper [9] some of the artificial intelligence based ontologies are developed for high end software reusability and knowledge sharing. The paper discusses some of the well known tools in the field of ontologies and knowledge extraction. The paper also describes the military ontology operations using a novel tool COMO (Comprehensive OWL Military Ontology). An application proof of concept is also implemented using the above tool. These applications are Planning for Urban Terrain Operations (PLUTO) system and Military Analogical Reasoning System (MARS). Ontology-based applications increase knowledge sharing by combining machine learning, knowledge extraction, and linguistic techniques. Today's AI technologies and tools facilitate effective real time knowledge management by allowing large-scale machine-mediated web-enabled knowledge sharing and reuse.

The paper [10] gives another ontology based system Customer Knowledge Management (CKM). It deals with customer record managements and their behaviour analysis towards a particular system. Here the system automatically processes the users record and derive the relation between them. The system uses Ontology Web Language (OWL) over the source code analysis for showing the semantic relation of ontologies.

The paper [11] a new model RDFKB (Resource Description Knowledge Base) is proposed which is a complete semantic web knowledge case. It solves the problem for managing and querying the knowledge sets. It provides various benefits over the other tools such as performance, uncertain information processing, inferred triples, probability information, and lineage information. RDFKB provides a complete and efficient RDF data repository and knowledge base.

The paper [12] work towards making an standard ontology process as a basic building block of semantic web. The creation of ontologies involves various complexities such as their integration, mapping, translation, reuse and consistency checks. As there is not any standard process of ontologies development, heterogeneous mixing and extraction is not possible with traditional systems. Moreover, the issues of duplicate information across documents and redundant annotations are major challenges of automatic ontology creation as the automatically populating ontology from diverse and distributed web resources poses significant challenges.

Thus after studying the various works related to ontology construction, there are so many possibilities with ontology development for semantic web for knowledge management systems.

IV. PROBLEM STATEMENT

After studying the different research articles related to Ontologies for effective Semantic Web knowledge management there are some of the problems identified which needs to be resolved for further improvement. First of all the evolving Semantic Web must support the knowledge management with model based application supports. Thus it processes the information having relation between the objects and their attributes. The knowledge is presented from the different sources and the processed values are served to some third party as a service. Though there are some of the problem associated with such functionalities are:

(i) Ontologies creation from heterogeneous sources is very tedious process due to lack of predefined rules for extraction. There must be some mechanism which verifies the rules for knowledge representation [13].

(ii) Types of tagging is insufficient, it must be according to the schematic categories of knowledge sources. Here the query optimization might be used with dynamic tags.

(iii) Change management with ontologies structure s not covered with RDF and OWL. During the dynamic updates the changes must be processed in such a way so as the accuracy of the analyzed content is improved [14].

Thus the ontologies categorization, their formation and dynamic nature must be handled with effective tagging schemes for further improving the above problems. It could be made feasible by using a meta data model, relation management and transformation of Ontologies.

V. PROPOSED WORK

This paper proposes a novel ontologies for representing the knowledge based systems in Semantic Web. Mainly it works towards improving the traditional web and forming it in such a way that it supports the recent Web 2.0. It provides the context oriented web data similarities based on object constructions. It focuses on the search meaning rather that the keyword mapping. The most related sentences with similar terminologies or meaning is separated out and processed according to the given ontologies. Once the mapping is performed the knowledge can be easily derived. It can be associated with the domain with a given standard. The work also develops the ontologies which could be further shared and reused. It saves the resources and provides additional search capabilities. While making the ontologies which can be shared the common vocabulary have to be developed along with the relation between them. The ontologies developed with the system define the structure of the knowledge and the associated assumptions which is commonly accessible to all the resources.

The process starts with generating the interface based request from the client machine or some server. The server after getting the search request tries to access the predefined ontologies which can match the most relativity rules. These rules are used for showing the relation between the terms of given ontologies. Here the designed Semantic Web built along the syntax which is defined by the URL or URI for giving the data rules. Syntax is given with the Resource Descriptor Framework (RDF). Here the combination of RDF and XML can be used for defining the logical schema of the ontology. The RDF manages the meta-data repository

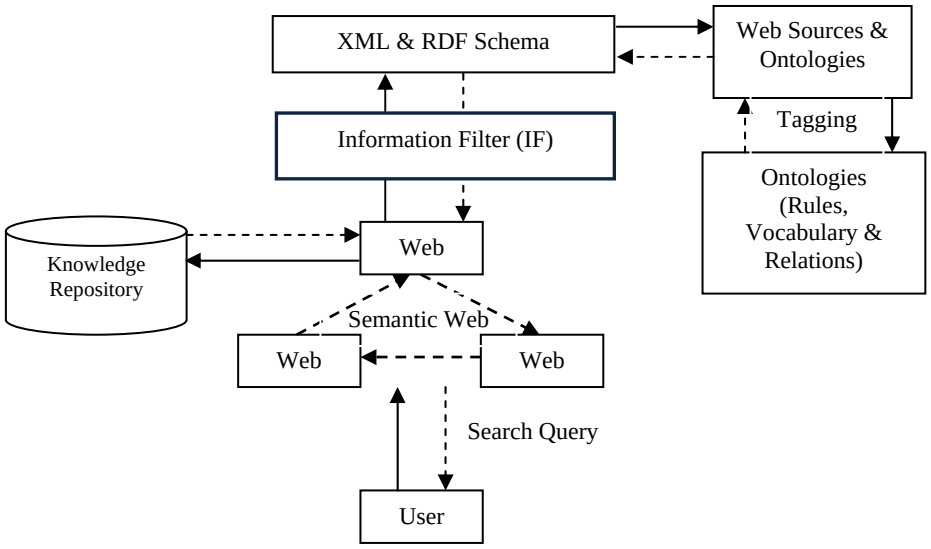

FIGURE 1: PROPOSED ONTOLOGY MECHANISM

and let their exchanges open between the communicating parties. Mainly the proposed work provides the improved RDF mapping using the discrete model which supports the Meta data changes. It also holds the conceptual relationship between the various ontologies and handles their transformation according to the requirements of resources. The data model used with RDF will captures the knowledge about the resources and their predicates or relationships. The suggested mechanism uses the triples data model with a dynamic combination of the subject, their relationship and the object.

The suggested system works on making the dynamic query updates using knowledge source modifications. As the source or its knowledge base changes the way ontology deals with the query is also changed. The users request is passed on to the Semantic Web network which separates the query according to their URI's. Mainly the separation involves partition of the data for forming the triples which contains the information o the subject, their relationships and the objects which is a data values. The segregated information is stored in the knowledge repository from where it is further processed. Now if the information which is passed for comparing and mapping with the ontologies doest matched with the required formats the processing systems the request for processing. It removes the problems associated with uncertain information and legacy ontologies processing's. The meta-data model and logics data is defined in the hierarchies given with the RDF and XML schemas. They are further used for holding the various types of user's oriented models. The rules defined for the ontologies formation is based on the domain knowledge and the logics. The rules can be dynamically updated by the ontology editor which later on be guided by the previously assigned logics. Knowledge repository is a relational database organized in a way that enables efficient storing and access to RDF metadata. This repository can be seen as a RDF repository. Knowledge processing component enables efficient manipulation with the stored knowledge, especially graph-based processing for the knowledge represented in the form of rules, e.g. deriving dependency graph or consistency checking Knowledge sharing is realized by searching for rules that satisfy the query conditions. In the RDF repository rules are represented as reified RDF statements and while in RDF any statement is considered to be an assertion, we can view an RDF repository as a set of ground assertions in the form (subject, predicate, and object). Rules are also related to domain ontology, which contains domain axioms used for deriving new assertions. Therefore the searching is realized as an inferencing process.

After analysing the complete process it seems that the suggested solution will met all its requirements and goals after the implementation or the proof of concepts. Still we are having a strong proof which shows the suggested approach is effective that its competitors.

BENEFITS OF WORK

- It provides effective processing of uncertain data which can't be processed form the previous ontologies.
- Semantic web is used for further exploring the knowledge base solutions.
- It shares the common understanding of the structure of information between the multiple software agents.
- It enables the reusability of the ontologies developed for the similar domains.
- It also allows the assumption to derive the explicit knowledge.
- It increases the operational and analysis capabilities of the system
- Annotation or Tagging improves the search accuracy, precision and recall.

APPLICATION AREA

(i) Web Service & monitoring
(ii) Web performance logger
(iii) Knowledge Processing System
(iv) Semantic Web
(v) Web 2.0 Mining and Evaluation
(vi) Online services marketing

VI. EVALUATION PARAMETERS

With this work the aim is to perform binary classification to predict ontologies based data files. For evaluating the ontologies based search queries classification models, the work uses recall, precision, inspection rate, and vulnerability rate.

- **Recall (R)** is defined as the percentage of files found: R=TP*100/(TP+FN).
- **Precision (P)** is defined as the percentage of correctly predicted files:

P=TP*100/(TP+FP).

VII. CONCLUSION

In recent world the focus is mainly resides behind the data oriented mechanism. It aims towards processing of information and knowledge exchanges between the heterogeneous parties in an effective manner. It directs the searching provided with the nominal related key values and the results must be having higher accuracy and relativity. For achieving its aim the ontologies provide a common understanding of a domain that can be communicated between people and different application systems and will play a major role in supporting information exchange processes in various areas. The use of a single ontology for all application will never be possible. Ontology will never be convenient for all subjects and domains or for a large and varied community such as the Web community. This paper presents an novel mechanism of ontologies processing using semantic web, RDF, XML schema and information filters. The approach is serving all its goals and giving the better results at initial phases of research. The futuristic implementation in correct direction will definitely improves the ontologies processing and will serve uncertain and dynamic information processing's.

REFERENCES

[1] Stefan Decker And Sergey Melnik, Frank Van Harmelen, Dieter Fensel, Michel Klein,Jeen Broekstra,Michael Erdmann,Ian Horrocks, " The Semantic Web" The Roles of XML and RDF",IEEE Internet Computing, October 2000, vol. IS, nr. 3, pages. 63-74.

[2] " Semantic web Technology, Layered Architecture, RDF and OWL representation"

[3] Asta back, sari vainikainen, caj Si)dergard and helene juhola,"Semantic web technologies in knowledge Management",

[4] York Sure et. al, "On-To-Knowledge: Semantic Web Enabled Knowledge Management", Institute AIFB, University of Karlsruhe

[5] Nenad Stojanovic, Siegfried Handschuh,"A Framework for Knowledge Management on the Semantic Web", ACM Press, pages498- 99,ISBN:IS81139128, 2002.

[6] Zeeshan Ahmed & Detlef Gerhard, "Web to Semantic Web & Role of Ontology", in Mechanical Engineering Informatics and Virtual Product Development Division, Vienna University of Technology, Getreidemarkt, Austria.

[7] James P. McGlothlin and Latifur Khan, "A Semantic Web Repository for Managing and Querying Aligned Knowledge", in The University of Texas at Dallas Richardson, TX USA

[8] Napoleon Alexandru & Constantin Florin Sîrbu, "Semantic Integration of Knowledge Management Systems", Communications of the IBIMA Volume 3, 2008

[9] Kamal Gella et. al., Effective Knowledge Management Using An Ontology Based Approach", Issues in Information Systems, Volume X, No. 2, 2009

[10] Yalan Yan & Xianjin Zha, "Applying OWL to Build Ontology for Customer Knowledge Management", in Academy Publisher, doi:10.4304/jcp.5.11.1693-1699, VOL. 5, NO. 11, 2010.

[11] James P. McGlothlin, Latifur Khan & Bhavani Thuraisingham, "RDFKB: A Semantic Web Knowledge Base", in Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence, 2011.

[12] Imran Sarwar Bajwa, "A Framework for Ontology Creation and Management for Semantic Web", in International Journal of Innovation, Management and Technology, Vol. 2, No. 2, April 2011.

[13] A.Rabiyathul basariya & O.S.Jannath Nisha, "Semantic Based Query Optimization using Ontology for Knowledge Management", in IEEE-International Conference On Advances In Engineering, Science And Management (ICAESM -2012), ISBN: 978-81-909042-2-3, March, 2012.

[14] Kaushik Manavadhariya et. al., "An Architecture for Efficient Data Retrival Through Semantic Web", in Indian Journal of Applied Research, Vol 3, Issue 4, 2013.