Exloring Semantic Web using Ontologies

Digivjay Singh M.Tech-CSE Moradabad R. K. Mishra

Department of CSE DBGI, Dehradun **Chandrashekhar**
Department of CSE SKIET, Kurukshetra

ABSTRACT: Virtually unlimited potential of data is available on web as time is passing away. But it is more difficult to extract the required data because keyword search in being used which is based on text of the data not on the semantic of data. Recently, researchers are trying to explore the potential associating with web content with explicit meaning in order to create a Semantic Web to minimize the manual efforts in search results. The semantic web is the next generation of syntactic web enriched by meaningful relations and tagging in the data management. World Wide Consortium (W3C) also trying to develop the standards creating and maintaining the semantic web like RDF, OWL and SPARQL etc. In those standards power of XML is being utilized to connect documents by creating the ontologies. This paper gives an overview of where the syntactic web is lacking and the standards which are currently used in semantic web.

[Digvijay singh, r.k. mishra, chanrashekhar. ***Exloring Semantic Web using Ontologies.*** *N Y Sci J* 2012;5(10):8-12]. (ISSN: 1554-0200). http://www.sciencepub.net/newyork. 2

Keywords: syntactic web, semantic web, ontologies, WWW, W3C, rdf.

INTRODUCTION

The amount of information on the Web and within organizations has grown at an astounding rate; however, the technology used to accommodate and process this information has not. Computers still cannot automate this information or perform complex tasks. Syntactic web can only provide a structure for text-based information in a document; it cannot do anything with it.

Different formats and structures make it difficult for computers or individuals to exchange data. Additionally, requirements change faster than we can solve the problems. Frequently, users are overwhelmed by too much unrelated data rather than the specific data required solving their problem.

In order to meet this requirement, the technology must be capable of:

- Retrieving large amounts of textual data quickly.
- Allowing users to add annotations so that a reasoning capability exists.
- Making text retrieval more specific.
- Allowing conclusions to be drawn by data on the Web and across organizations.

The Semantic Web is the next generation of the current Syntactic Web in which computers can interpret the meaning of web content because of explicit semantics provided in markup. It can be thought of as an infrastructure for supplying the web with formalized knowledge in addition to its actual informal content. The Semantic Web components are deployed in the layers of Web technologies and specifications.

Structure of Semantic Web (Layered Approach)

The semantic Web identifies a set of technologies, tools, and standards which form the basic building blocks of an infrastructure to support the vision of the Web associated with meaning. The semantic Web architecture is composed of a series of standards organized into a certain structure that is an expression of their interrelationships. This architecture is often represented using a diagram first proposed by Tim Berners-Lee (Berners-Lee, Hendler et al. 2001). Figure 4 illustrates the different parts of the semantic Web architecture. It starts with the foundation of URIs and Unicode. On top of that we can find the syntactic interoperability layer in the form of XML, which in turn underlies RDF and RDF Schema (RDFS). Web ontology languages are built on top of RDF(S). The three last layers are the logic, proof, and trust, which have not been significantly explored. Some of the layers rely on the digital signature component to ensure security.

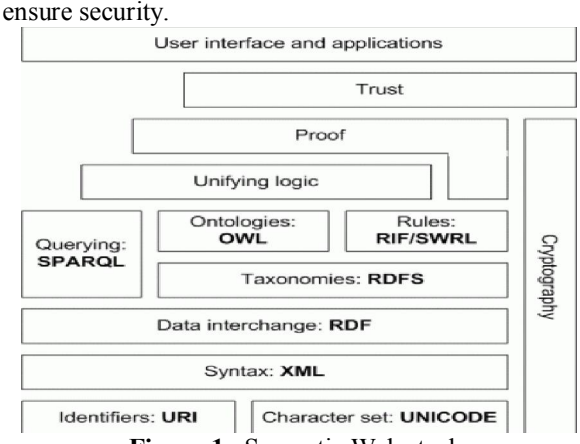

Figure 1. Semantic Web stack

a) URI (Identifiers) and Unicode (Character set)

A Universal Resource Identifier (URI) is a formatted string that serves as a means of identifying abstract or physical resource. A URI can be further classified as a locator, a name, or both. Uniform Resource Locator (URL) refers to the subset of URI that identifies resources via a representation of their primary access mechanism. A Uniform Resource Name (URN) refers to the subset of URI that is required to remain globally unique and persistent even when the resource ceases to exist or becomes unavailable.

b) RDF (Data Interchange)

The World Wide Web Consortium (W3C) has developed the Resource Description Framework (RDF) language to standardize the definition and use of metadata. RDF is a simple general-purpose metadata language for representing information in the Web and provides a model for describing and creating relationships between resources. Resources have properties associated to them. Properties are identified by property-types, and property-types have corresponding values. Property-types express the relationships of values associated with resources. The basic structure of RDF is very simple and basically uses RDF triples in the form of subject, predicate, object.

subject: a thing identified by its URL
predicate: the type of metadata, also identified by a URL (also called property)
object: the value of this type of metadata

c) RDF Schema (Taxonomies)

The RDF Schema (RDFS 2004) provides a type system for RDF. The RDFS is technologically advanced compared to RDF since it provides a way of building an object model from which the actual data is referenced and which tells us what things really mean. The RDF Schema (RDFS) allows users to define resources with classes, properties, and values.

The rdfs:Class is similar to the notion of a class in object-oriented programming languages. When a schema defines a new class, the resource representing that class must have an RDF:type property whose value is the resource rdfs:Class. Anything described by RDF expressions is called a resource and is considered to be an instance of the class rdfs:Resource. Other elements of RDFS are illustrated in Figure 2 and described below.

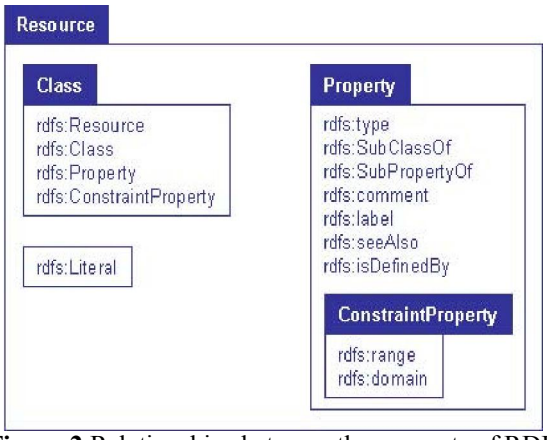

Figure 2. Relationships between the concepts of RDF Schema

d) Ontologies (OWL)

Ontology is an agreed vocabulary that provides a set of well-founded constructs to build meaningful higher level knowledge for specifying the semantics of terminology systems in a well defined and unambiguous manner. Ontologies are usually expressed in a logic-based language, so that detailed and meaningful distinctions can be made among the classes, properties, and relations. Ontologies can be used to increase communication either between humans and computers. The three major uses of ontologies are:

1. To assist in communication between humans.
2. To achieve interoperability and communication among software systems.
3. To improve the design and the quality of software systems.

e) SPARQL (Query Language)

SPARQL has been designed to meet the requirements and design goals as described in the W3C RDF Data Access Working Group (DAWG) document "RDF Data Access Use Cases and Requirements".

SPARQL provides functionalities to

- extract information represented as literals, blank nodes, and URIs
- gather RDF sub graphs, and
- build new RDF graphs upon information achieved from the queried graphs.

SPARQL offers many of the basic features desired in an RDF based query language. It provides a subset of operations on plain literals, XSD integers and XSD floats and 26 operators such as comparison of numeric values, functions on string values casting, comparison of duration, time and date values.

g) Logic, Proof and Trust

The purpose of this layer is to provide similar features to the ones that can be found in First Order Logic. The idea is to state any logical principle and allow the computer to reason by inference using these principles. For example, a university may decide that if a student has a GPA higher than 3.8, then he will receive a merit scholarship. A logic program can use this rule to make a simple deduction: "Shyam has a GPA of 3.9; therefore he will be a recipient of a merit scholarship."

Inference engines, also called reasoners, are software applications that derive new facts or associations from existing information. Inference and inference rules allow for deriving new data from data that is already known. Thus, new pieces of knowledge can be added based on previous ones. By creating a model of the information and relationships, we enable reasoners to draw logical conclusions based on the model. The use of inference engines in the semantic Web allows applications to inquire why a particular conclusion has been reached, i.e. semantic applications can give proof of their conclusions. Proof traces or explains the steps involved in logical reasoning.

A Practical Scenario for the explanation of working

In a real scenario, there is a site Fashion ADS which a web service that maximizes precision of advertising. In particular, it mediates in meaning negotiation between sellers and prospective buyers by intelligent use of pragmatic pattern matching. The framework is explained using a situation in which there is a company BABYWEARS that manufacture clothes for kids. It wants to advertise in Fashion ADS for promotion of its new collection of frocks for baby girls and exploring new market.

The community using the Fashion ADS service distinguishes two types of communication processes inquiring about objects for sale, initiated by customers, and advertising objects, initiated by producers.

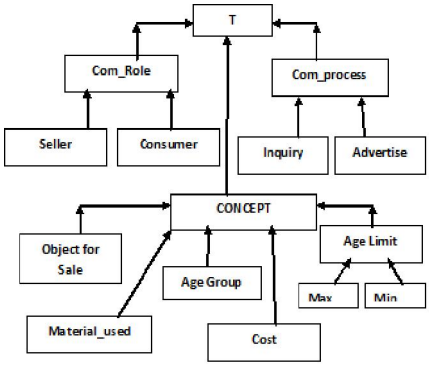

Figure 3.Fashion ADS community context ontology

Baby wears own corporate ontology specifying the Frock and age limit concepts. This is imported into the individual context ontology of Baby wears for the Fashion ADS service. It also adds the Gender (girl) concept, since that is what he wants to focus his particular potential customer search on. Since the frocks are for baby girls, it adds to his individual context ontology the required pattern that to be of interest for an advertisement girls for whom the frocks are purchased should be in infant age group. The Ritswears is a well-known outlet that provides various dresses of different brands for children.

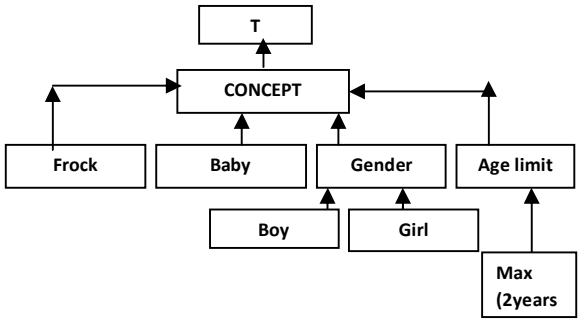

Figure 4. Individual Context Ontology of Babywears for the Fashion ADS Service.

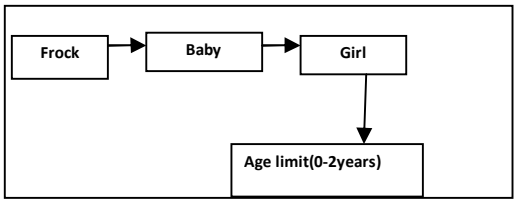

Figure 5.Required Pattern. This figure shows the pragmatic pattern of Baby wears for advertising its product and get potential customers

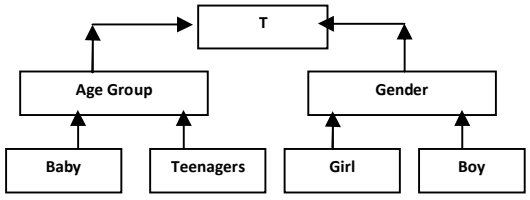

Figure 6. Individual Context Ontology of Ritswears for the Fashion ADS Service.

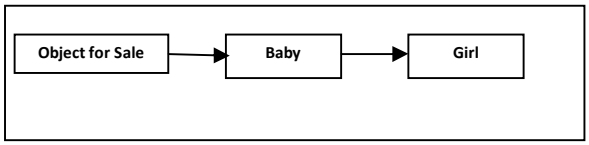

Figure 7. Required Pattern. This figure shows the pragmatic pattern of Ritswears for advertising its requirement to get potential providers

Now when Baby wears uses Fashion ADS for getting potential customer, it gives query as shown in figure 6.

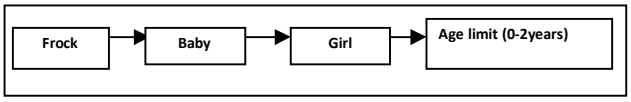

Figure 8. Query. This figure shows the query that is given by Baby wears agent in Fashion ADS search engine to find potential customers for its product.

Then Fashion ADS searches all the individual context ontology of the registered services to get the correct customer. When it matches with Ritswear context ontology, it found that it's not matched properly and so gives a null result. When agent of Baby wears finds the result is not coming appropriately, he modifies query to b general as shown in figure 9

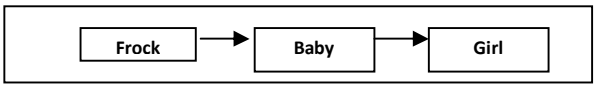

Figure 9. Generalized Query. This figure shows the simpler query that is given by Babywears agent in Fashion ADS search engine to find potential customers for its product

Now Fashion ADS match it with Ritswear's ontology and gets a match. Then to give more results it mines semantic web by using different synonyms and found that baby and infant are synonyms of each other and modifies search for infants. Then it found a new customer Kukoon that is especially for kid's wear, therefore demand that any sales offer in an advertisement concerns with kids only.

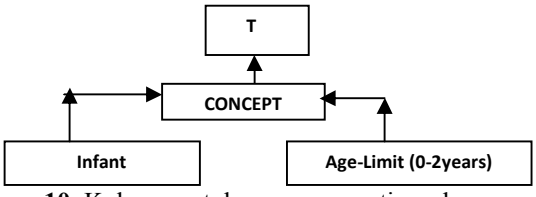

Figure 10. Kukoon ontology on semantic web

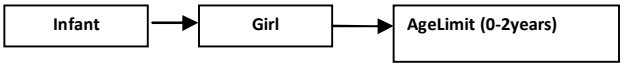

Figure 11. Semantic pattern query. This figure shows the semantic pattern of Kukoon in Semantic Web

Mining the semantic web, Fashion ADS concludes that there is common pattern, which should be added to common context ontology. And thus negotiation process is undertaken. And in return the Baby wears gets two customers from the fashion ADS. This meaning negotiation regularly updates the common context ontology and makes it more powerful to mine the search and also creates a virtual social network where all users of common concept, interest are grouped and provided with their desired products. And this updated ontology again mines the web to find better results.

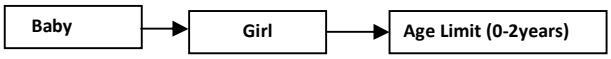

Figure 12. Common pattern added to the common context ontology. This figure shows the updating of the common context ontology of Fashion ADS after meaning negotiation.

CONCLUSION AND FUTURE SCOPE

Internet provides a strong ability for advertisers to target users on the internet with the aid of information technology. Targeted advertising refers to delivering the appropriate advertisements to the users and is considered as the trends of internet advertising. In the study, we present an intelligent agent based targeted advertising model. Our target advertisement model adopts pragmatic pattern and semantic web to mine user's request.

We summarize the future direction as follows: A target advertising system needs to collect user context ontology. Yet, in general, the websites are not developed using RDF/OWL and so do not have ontologies. The vision of the Pragmatic Web is thus to augment human collaboration effectively by appropriate technologies, such as systems for ontology negotiations, for ontology based business interactions, and for pragmatic ontology-building efforts in communities of practice.

References

[1] Berners-Lee, T., Hendler, J. and Lassila, O. (2001), The Semantic Web, Scientific American, May 2001: pp 35-43

[2] McCool, R. Rethinking the Semantic Web, Part 1. IEEE Internet Computing (Nov.–Dec.2005), 86–88.v

[3] de Moor, A., Keeler, M. and Richmond, G. (2002), Towards a Pragmatic Web. In Proc. Of the 10th Intl. Conference On Conceptual Structures (ICCS 2002), Borovets, Bulgaria, July 2002, LNAI 2393. Springer, Berlin, pp. 235-249.

[4] Repenning, A. and Sullivan, J. (2003), The Pragmatic Web: Agent-Based Multimodal Web Interaction with no Browser in Sight. In Human-Computer Interaction - INTERACT'03. IOS Press, IFIP, pp. 212-219.

[5] de Moor(2005),Patterns of the web. Invited paper, Proc. of the 13th International Conference on Conceptual Structures (ICCS 2005), Kassel, Germany, July 2005, LNAI. Springer Verlag, Berlin, pp.1-18

[6] S. Thaddeus, S.V. Kasimir Raja: Ontology-driven Software Engineering Environment. Proceedings of the Eighteenth International Conference on Software Engineering and Knowledge Engineering, San Francisco, California, July (2006), 337-342.

[7] Singh, M. P. (2002b), The Pragmatic Web: Preliminary Thoughts. In Proc. of the NSF-EU Workshop on Database and Information Systems Research for semantic Web and Enterprises, April 3-5, Amicalolo Falls and State Park, Georgia.

[8] Spyns, P. and Meersman, R. A. (2003), From Knowledge to Interaction: from the Semantic to the Pragmatic Web. Technical Report STAR-2003-05, STARLab , Brussels.

7/30/2012