DESIDOC Journal of Library & Information Technology, Vol. 31, No. 2, March 2011, pp. 116-120

Role of Ontology in Semantic Web

Kaushal Giri

Don Bosco Institute of Technology, Mumbai-680 651
E-mail: kaushal.giri@gmail.com

ABSTRACT

The present generation of computers is changing from single isolated devices to entry points into a worldwide network of information exchange. Therefore, support in the exchange of data, information, and knowledge is becoming the key issue in computer technology today. The increasing volume of data available on the Web makes information retrieval a tedious and difficult task. Researchers are now exploring the possibility of creating a semantic Web, in which meaning is made explicit, allowing machines to process and integrate Web resources intelligently. The vision of the semantic Web introduces the next generation of the Web by establishing a layer of machine-understandable data. The success of the semantic Web crucially depends on the easy creation, integration and use of semantic data, which will depend on building an ontology. This paper states the role of ontology in supporting information exchange process, particulary with semantic Web.

Keywords*:* Ontology, computer technology, information retervial, Internet, semantic Web

1. INTRODUCTION

Today we are living in the information age and society can be labeled as information society. Knowledge, a predominant element, pervades and dominates the day-to-day activities in information society. There is a sea change in information generation, distribution, and access. In recent years, finding information from the Web is becoming more and more complex. Searching on the Internet can be compared to dragging a net across the surface of the ocean.

The traditional search engines are unable to provide satisfactory solutions to this. Research is in progress to build a Web, which is semantically richer than the current one. The ability to translate knowledge from different languages is an important ingredient for building powerful artificial intelligent (AI) systems, by easing the difficult and time-consuming task of knowledge base construction[1].

The vision of semantic Web proposes an environment where the data and services on the Web can be semantically interpreted and processed by machines to facilitate human consumption. The semantic Web relies heavily on the formal ontologies that structure underlying data for the purpose of comprehensive and transportable machine understanding. Semantic Web technology relies on ontology as a tool for modelling an abstract view of the real world and contextual semantic analysis of documents[2]. Therefore, the success of the semantic Web depends predominantly on the proliferation of ontologies, which requires fast and easy engineering of ontology and avoidance of a knowledge acquisition bottleneck. This paper discusses about the development of semantic Web, semantic Web technologies and role of ontology in the semantic Web.

2. LIMITATION OF THE PRESENT WEB

Current World Wide Web (WWW) is a huge library of interlinked documents that are transferred by computers and presented to people. It has grown from hypertext systems, but the difference is that anyone can contribute to it. This also means that the quality of information or even the persistence of documents cannot be uniformly guaranteed. HTML's simplicity has seriously hampered more advanced Web applications in many domains and for many tasks. This was the reason for defining another language, Extensible Makup Language (XML), which allows arbitrary domain and task specific extensions to be defined. Semantic Web is an XML application, developed to make the present Web semantically richer.

3. SEMANTIC WEB

The idea of the semantic Web was conceived by Tim Berners-Lee, the founder of the WWW. He envisions that in future, the vast amount of information on the Web will bear machine readable metadata, resulting in computers being able to manipulate the contents automatically, without human intervention. Therefore, the semantic Web is imagined as an extension of the Web, in which information is given a well defined meaning. It is the application of advanced knowledge technologies to the Web and distributed systems in general. It describes methods and technologies to allow machines to understand the meaning or "semantics" of information on the WWW. To accomplish this, the provided information should be structured, accompanied by sets of inference rules that can be used by computers to conduct automated reasoning.

According to Tim Berners Lee, "The Web has developed most rapidly as a medium of documents for the people rather than of information that can be manipulated automatically. By augmenting webpages with data targeted at computers and by adding documents solely for computers, we will transform into semantic Web. Computers will find the meaning of semantic data by following hyperlinks to definitions of key terms and rules for reasoning about them logically"[3].

3.1 Development of the Semantic Web

The vision of extending the current human-focused Web with machine processable descriptions of Web content was first formulated in 1996 by Tim Berners-Lee, the original inventor of the Web. The semantic Web has been actively promoted since then by the WWW Consortium, the organisation that is chiefly responsible for setting technical standards on the Web. At this point, the field of knowledge representation and reasoning took center stage, but outcomes from other fields of AI have also been put to use to support the move towards the semantic Web. For example, natural language processing and information retrieval have been applied to acquire knowledge from the WWW[4].

As semantic Web is a relatively new and dynamic field of investigation, it is difficult to precisely delineate the boundaries amongst these. Semantic Web communities have defined their community by including those researchers who have submitted publications or held an organising role at any of the past International Semantic Web Conferences, or the Semantic Web Working Symposium[5]. Today semantic Web has created its own importance by developing meaningful results.

3.2 Semantic Web Technologies

The semantic Web contains resources corresponding not just to media objects (such as Webpages, images, audio clips, etc.) as the current Web does, but also objects such as people, places, organisation and events. Further, the semantic Web will contain not just a single kind of relation (the hyperlink) between resources, but many different kinds of relations amongst the different kinds of resources. Various technologies for the semantic Web are shown in the Fig. 1

XML is a simple language that permits users create their own tags in order to annotate Web documents. It provides an elemental syntax for content structure within documents. XML Schema is a language for providing and restricting the structure and content of elements contained within XML documents. While an XML document is a tree, an Resource Description Framework (RDF) document consists of sets of triples. Each triple contains a subject, predicate and an object. These triples can be written using XML tags. RDF is a simple language for expressing data models, which refer to objects ("resources") and their relationships. An RDF-based model can be represented in XML syntax. In RDF, a document makes assertions that things have properties, and this is how most of the data can be described and further processed by the computers.

RDF schema extends RDF and is a vocabulary for describing properties and classes of RDF-based resources, with semantics for generalised-hierarchies of such properties and classes[6]. XML and RDF are two complimentary technologies used to build an intelligent Web. RDF model uses vocabulary defined by the terms of ontology. The combination of an RDF model and the associated XML gives the computer enough information to discover the meaning of data. Data about other data is

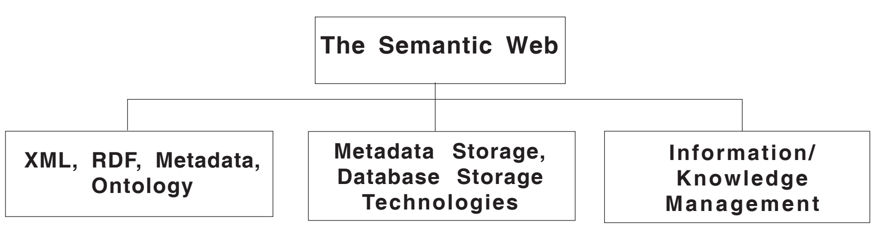

Figure 1.Various semantic Web technologies.

often called metadata. XML and RDF deal with metadata, i.e., they deal with the description of the information available on the Web. But, if machines are expected to interact with each other or share data in the true sense of the word, then semantic interoperability is essential. For this, a formal specification is required to explicitly define various terms and their relationships. Ontology was thus developed in AI to facilitate knowledge sharing and reuse, and can be built using XML and RDF. But during recent past, many specific ontology-development languages have been developed. Knowledge management from the field of AI provides the necessary means to fill the knowledge gap. Information that is hard to access for our machines can be made accessible using an ontology. Ontologies will play a major role in supporting information exchange processes in the Semantic Web.

4. ONTOLOGY

Since the beginning of the 1990s ontology has become a popular research topic investigated by several AI research communities including knowledge engineering, natural language processing, and knowledge representation. More recently, the notion of ontology is also becoming widespread in the fields like intelligent information integration, cooperative information systems, information retrieval, electronic commerce, and knowledge management.

The reason ontology is becoming so popular is because it promise a shared and common understanding of some domain that can be communicated between people and application systems. Because ontology aims at consensual domain knowledge, its development is often a cooperative process involving different people, possibly at different locations. According to Wikipedia, "In computer science and information science, ontology is a formal representation of the knowledge by a set of concepts within a domain and the relationships between those concepts. It is used to reason about the properties of that domain, and may be used to describe the domain[7].

4.1 Components of Ontology

Contemporary ontologies share many structural similarities, regardless of the language in which they are expressed. Most ontology describes individuals (instances), classes (concepts), attributes, and relations. Common components of ontologies include:

- *Individuals*: Instances or objects (the basic or "ground level" objects)
- *Classes*: Sets, collections, concepts, classes in programming, types of objects, or kinds of things.
- *Attributes*: Aspects, properties, features, characteristics, or parameters that objects (and classes) can have
- *Relations*: Ways in which classes and individuals can be related to one another
- *Function terms*: Complex structures formed from certain relations that can be used in place of an individual term in a statement
- *Restrictions*: Formally stated descriptions of what must be true in order for some assertion to be accepted as input
- *Rules*: Statements in the form of an if-then (antecedent-consequent) sentence that describe the logical inferences that can be drawn from an assertion in a particular form
- *Axioms*: Assertions (including rules) in a logical form that together comprise the overall theory that the ontology describes in its domain of application
- Events: The changing of attributes or relations

4.2 Ontology Construction

The following steps are involved in the construction of Ontology:

- *Acquiring the domain knowledge*: This step consist of identifying and collecting appropriate expertise and information resources to present in a common language, all descriptive terms with consensus and consistency
- *Design the conceptual structure*: Identify the key concepts within the domain along with their associated properties. Further identifying the relationships among the concepts
- *Develop the suitable details*: Include concepts, relationships and various instances to achieve the level of detail to satisfy the given purpose of the ontology
- *Verify*: Ensure the structure for uniformity. If required, revise any logical, semantic or syntactic errors among the elements
- *Commit/submit*: Once domain expert ensures, the ontology is committed by publishing it within its planned deployment location. Then the ontology can be referred and used within the environment

4.3 Ontology Languages

An ontology language is a formal language used to encode the ontology. The advantage of formal languages is the reasoning mechanism which appears in every phase of conception, use and maintaince of an ontology. There are a number of such languages for ontologies, both proprietary and standards based, which are as follows:

4.3.1 DAML+OIL

DAML stands for DARPA Agent Markup Language. DARPA in turn stands for Defence Advanced Research Projects Agency and is the central research and development organisation for the Department of Defense. OIL stands for Ontology Inference Layer or Ontology Interchange Language. DAML+OIL are a successor language to DAML and OIL that combines features of both. In turn, it was superseded by Web Ontology Language (OWL). DAML+OIL uses description logic (DL) style model theory to formalise the meaning of the language.

This is a very important feature to reduce arguments and confusions, thus giving the language the ability to precisely represent the meaning of information. This ability is crucial for automatic reasoning, which is the goal of the semantic Web[8].

4.3.2 SWRL

Semantic Web Rule Language (SWRL) is a proposal for a semantic Web rules-language, combining sub-languages of the OWL, Web Ontology Language (OWL DL and LITE) with those of the Rule Markup Language (Uniary/binary datalog). SWRL adds rules to OWL+DL. The reason is that these rules provide more expressive power to description logic. SWRL plays an important role in ontology for the semantic Web.

4.3.3 Web Ontology Language

The Web Ontology Language (OWL) is a family of knowledge representation languages for authoring ontologies. The languages are characterised by formal semantics and RDF/XML-based serialisations for the semantic Web. The OWL was designed to add the constructs of DL to RDF, significantly extending the expressiveness of RDF schema both in characterising classes and properties. It has been designed in a way that it maps to a well-known DL with tractable reasoning algorithms[9]. It has classes, sub-classes, properties, sub-properties, property restrictions, and both class and property individuals. Standardised formal semantics and additional vocabulary makes OWL to represent explicit term descriptions and the relationship between entities.

4.4 Ontology Editors

Ontology editors are applications designed to assist in the creation or manipulation of ontologies. They often express ontologies in one of many ontology languages.

4.4.1 Protege

Protege is a free, open source ontology editor and a knowledge acquisition system. Protégé is being developed at Stanford University in collaboration with the University of Manchester. This application is written in Java and heavily uses Swing to create the rather complex user interface[10].

4.4.2 DOME

The DERI Ontology Management Environment (DOME) is developed by the Ontology Management Working Group (OMWG). The mission is to create a suite for the efficient and effective management of ontologies that provides an integral solution for the overall problem.

4.4.3 Knoodl

Knoodl facilitates community-oriented development of OWL-based ontologies and RDF knowledge bases. It also serves as a semantic technology platform, offering a Java service-based interface or a SPARQL-based interface so that communities can build their own semantic applications using their ontologies and knowledge bases.

4.4.4 Onto Edit

Onto Edit is an ontology editor developed by the On-To-Knowledge project. It implements an ontology constructions process. The Editor stores the ontology conceptual model in the SESAME repository and produces RDF concrete representations.

5. ROLE OF ONTOLOGY FOR THE SEMANTIC WEB

The semantic Web relies heavily on the formal ontologies that structure underlying data for the purpose of comprehensive and transportable machine understanding. Therefore, the success of the semantic Web depends strongly on the proliferation of the ontology which requires fast and easy engineering of ontology and avoidance of a knowledge acquisition bottleneck.

Conceptual structures that define an underlying ontology are germane to the idea of machine processable data on the semantic Web. Ontologies are metadata schemas, providing a controlled vocabulary of concepts, each with an explicitly defined and machine processable semantics. By defining shared and common domain theories, ontology helps both people and machines to communicate precisely to support the exchange of semantics. Ontology language editors help to build semantic Web. Hence, the cheap and fast construction of domain specific ontology is crucial for the success of the semantic Web.

6. CONCLUSION

The semantic Web is not a separate Web but an extension of the current one, in which information and services are given a well defined meaning. The semantic Web is a Web for machines, but the process of creating and maintaining it is a social one. To make possible the creation of the semantic Web the W3C has been actively working on the definition of open standards, such as the RDF and OWL. Although machines are helpful in manipulating symbols according to pre-defined rules, only the users of the semantic Web have the necessary interpretative and associative capability for creating and maintaining ontologies. The principal benefit of semantics is that it provides a formal foundation for reasoning about the properties of systems that do automated knowledge translation based on sharing of ontology. Developers are vigorously building semantic Web services. In spite of this, as the standards are prepared and the demand for intelligent agents grows, expertise will be required in the field of semantic Web services.

REFERENCES

1 Yagge, F. & Akkemans, J.M. Decentralized markets versus central Control: A comparative study. *J. Artificial Intell. Res.,* 1999, **11**, 301-33.

2. Stabb, S. & Maedche, A. Knowledge portals—ontologies at work. *AI Magazine*, 2001, **21**(2).

3. Berners-Lee, Tim; Hendler, James & Ora, Lassila: The semantic web. *Scient. Ameri. Mag.,* 2002.

4. Maedche, A. & Staab, S. Ontology learning for the semantic Web, *IEEE Intel. Syst.,* 2001, **16**(2), 72–79.

5. Horst, H. Combining RDF and part of OWL with rules. semantics, decidability, complexity. *In* ISWC 2005, edited by Gil, Y., Motta, E., Benjamins, V.R., & Musen, M.A. Fourth International Semantic Web Conference, Galway, Ireland, 2005. **3729**. pp. 668–84.

6. Decker, S. The semantic Web roles of XML and RDF. *IEEE Internet Comp.*, 2000, **4**(5), 63–74.

7. Gruber, T.A translation approach to portable ontology specifications. *Knowledge Acquisition*, 1998, **5**, 199.

8. Horrocks, I. DAML+ OIL: A reasonable Web ontology language. *In* Lecture Notes in Computer Science (LNCS), Springer-Verlag, 2002, **2287**. pp. 2-13.

9. Smith, Michael K; Welty, Chris & McGuinness, Deborah L. OWL Web ontology language guide, W3C, 2008. pp. 7-15.

10. Knowledge modelling at the millennium the design and evolution of Protege. *In* Proceedings of KAW-99, edited by E. Grosso, H. Eriksson, R. Fergerson, S. Tu and M. Musen, Banff, Canada, 1999.

About the Author

Mr Kaushal Giri obtained Bachelor's degree in Commerce, and Library & Information Science from the University of Mumbai, and Masters in Library and Information Science from Documentation Research and Training Centre (DRTC), Bengaluru. He is working as Chief Librarian at Don Bosco Institute of Technology, Mumbai. His areas of specialisations are information retrieval, knowledge and content management, digital libraries, semantic web, and ontology.