# Semantic Web and Ontologies: A Survey

Vandana Mohan Patil
Assistant Professor
Department of Information Technology,
R.C. Patel Institute of Technology,
Shirpur, Maharashtra, India

## ABSTRACT
In this paper, we give an overview of relationship between the two potential areas of research, Semantic Web and Ontologies. In our survey, we will first describe the state-of-the-art of both the areas and then discuss their relationship. The Semantic Web has taken today's World Wide Web one step ahead with the integration of semantics into the underlying domain knowledge. The Semantic Web relies on one or more ontologies for describing this semantic knowledge of the domain of interest. An ontology is a formal explicit description of i) concepts in a domain of discourse i.e. classes, ii) properties of each concept describing various features and attributes of the concept i.e. slots or roles and iii) restrictions on slots i.e. facets or role restrictions. Ontology together with a set of individual instances of classes constitutes a Knowledge Base. Recent research works are trying to improve the existing approaches of Web log mining by integrating the semantics into the underlying domain knowledge using Ontologies.

## General Terms
Semantic Web, Ontologies, Web log mining.

## Keywords
Semantic Web, Ontologies, WWW, knowledge base, semantic annotation, Web log mining.

## 1. INTRODUCTION
Today's World Wide Web (WWW) is growing exponentially in the size and use. However, in today's Web, Web pages are designed to be read by people, not by machines. Whereas, in Semantic Web, information is understandable by computers, so computers can perform more of the tedious work involved in finding, combining, and acting upon information on the web [1]. From the perspective of Tim Berners Lee, the inventor of the WWW, the Semantic Web is a new form of Web content that is meaningful to computers. 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]. For the Semantic Web to function, computers must have access to structured collections of information and sets of inference rules that they can use to conduct automated reasoning. Ontology is one of the basic components of the Semantic Web. Ontologies are collections of information. In philosophy, ontology is a theory about the nature of existence. From the perspective of Artificial intelligence and Web researchers, ontology is a document or file that formally defines the relations among terms. The most typical kind of ontology for the Web has taxonomy and a set of inference rules [1, 2]. Semantic Web use Ontologies to provide vocabulary and to describe the background domain knowledge [3]. The ontologies give meaning of data in a format that machines can understand. The data derive its semantics from ontology. It acts as knowledge base to support interpretation, reasoning, inference, mapping, and integration of heterogeneous data across communities [4]. Thus, now-a-days, "Ontology" has become a very potential area for research. Most of the researches on Ontology focus on matching ontologies, automated ontologies extraction from web sites, ontological concept extraction from Web pages, data integration etc.etc. [5]

## 2. SEMANTIC WEB
From the perspective of Tim Berners-Lee, the inventor of the WWW, the Semantic Web is a new form of Web content that is meaningful to computers. 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 [1]. In general, semantics is the study of meaning [2]. If a computer understands the semantics of a document, it doesn't just interpret the series of characters that make up that document: it understands the document's meaning [3]. Semantic Web is an efficient way to represent data on the World Wide Web, or is a database that is globally linked, in a manner understandable by machines, to the content of documents on the Web. Semantic technologies represent meaning using ontologies and provide reasoning through the relationships, rules, logic, and conditions represented in those ontologies. [2, 3, 4] Semantic Web provides the ability to tag all content on the Web, describe what each piece of information is about and give meaning to the content item. Thus, search engines become more effective than they are now, and users can find the precise information they are hunting [5]. Organizations that provide various services can tag those services with meaning. Using Web-based software agents, users can dynamically find these services on the y and use them to their benefit [6, 7].

## 3. ONTOLOGIES
### 3.1 Definition
Ontology is a conceptualization of a domain of interest that can be used in several ways to model, analyze and reason upon the domain [8]. Ontologies form the backbone of a whole new way to understand online data [5]. Ontologies are the basic infrastructure for the Semantic Web. Main role of ontologies in Semantic Web is to describe the semantics of various terms in the shared vocabularies of Semantic Web.

### 3.2 Need of Ontology
The rapid growth in the development of new technologies and massive increase in the use of computers in all areas of the human activities, have resulted in generation of very huge amount of data on the Web. This data needs the proper management in such a way that it can be efficiently accessed,

used, navigated and reasoned about the individuals. The efficient and proper management of Web data can be achieved through the means of one or more Ontologies [8, 9].

## 3.3 Ontology Levels

Ontologies are typically divided into:

(1) Foundational or top level ontologies
These ontologies cover the most general categories those can be expected to be common to all domains. Therefore, they are also called as "domain independent".

(2) Domain ontologies
These ontologies are tailored for a specific area of human activity. e.g. medicine, electrical engineering, biology, business etc. Domain ontology can be task specific, company specific and so forth.

(3) Application ontologies
These ontologies are restricted to a particular activity in a domain. e.g. the diagnosis of lung diseases in medicine or a computer based order handling system in business. The application ontologies are necessary because a general ontology of the business domain cannot be used directly in any concrete business application.

(4) Personal level ontologies
This ontology level takes into account, for example, the way in which an individual uses particular information system for a particular task, which is often different from the way others use the same system for the same or different task [10].

## 3.4 Ontology Development phases

**Basically, Ontologies are developed in an iterative fashion [10]. Development of ontology involves following phases:**

(1) Requirements and Analysis

(2) Design and Implementation

(3) Testing and Validation

(4) Maintenance [11]

In practical terms, developing an ontology includes:

1. Defining classes in the ontology
2. Arranging the classes in a taxonomic (subclass-superclass) hierarchy
3. Defining slots and describing allowed values for these slots
4. Filling in the values for slots for instances [10].

We can then create a knowledge base by defining individual instances of these classes filling in specific slot value information and additional slot restrictions.

## 3.5 Reasons of Ontology Development

Ontlogies are developed for the purpose of-

(1) Sharing of common understanding of information structure among not only people but also software agents.

(2) Enabling reuse of knowledge of a particular domain.

(3) Explicit domain assumptions.

(4) Isolation of domain knowledge from operational knowledge

(5) Analysis of domain knowledge, etc [10].

## 3.6 Ontology Components

Typically Ontologies have following components:

1. Concepts in a domain of discourse (sometimes called as classes).
2. Properties of each concept describing various features and attributes of the concept (sometimes called as slots, roles or properties).
3. Restrictions on slots (facets (sometimes called role restrictions)). [8]

## 3.6 Applications of Ontology

Following are some of the applications of Ontology:

(1) Information retrieval procedure

(2) Knowledge representation/sharing

(3) Semantic Digital Libraries

(4) Software engineering

(5) Natural-Language processing

(6) Multi-agent systems

(7) Information systems in general

(8) Ontology based reasoning

(9) Policy language specification

(10) Database design, etc [8,9,10].

# 4. SEMANTIC WEB AND ONTOLOGIES

The explosive growth in both the range and quantity of Web content has highlighted some serious shortcomings in the hypertext paradigm. In the first place, the required content becomes increasingly difficult to locate using search and browse. Second is answering more complex queries along with more general information retrieval, integration, sharing and processing can be difficult or even impossible [1, 8, 12].

Existing approaches are very useful, but they do not solve the general problem of how to locate and integrate information on demand and without human intervention. These problems can be overcome using Semantic Web. The Semantic Web is developed with the ultimate goal being to allow data to be shared effectively by wider communities and to be processed automatically by tools as well as manually" [3, 4]. A major difficulty in realizing this goal is that most Web contents are primarily intended for presentation to and consumption by human users. HTML markup is mainly concerned with layout, size, colour and other presentational issues. Moreover, Web pages increasingly use images, often including active links, to present information, and even when content is annotated, the annotations typically take the form of natural language strings and tags [2]. Human users are usually able to interpret the significance of such features, and thus understand the information being presented, but this may not be so easy for a software agent. A key idea behind the Semantic Web is to address this problem by giving machine accessible semantics to annotations [1]. This is achieved by using ontologies, rich conceptual schemas to give formally defined meanings to the terms used in annotations, transforming them into semantic annotations [13]. This vision of a Semantic Web is clearly extremely ambitious, and its full realization would require the solution of many very hard and long-standing research problems in areas such as knowledge representation and reasoning, databases, computational linguistics, computer vision and agent systems.

One such problem is the tradeoff between conflicting requirements for expressive power in the language used for semantic annotations and scalability of the systems used to process them [14]. Another is that integrating different ontologies may prove to be at least as hard as integrating the resources that they describe [15]. New problems include the need to create suitable annotations and ontologies, and the variable quality of Web information sources [1].

Notwithstanding the above mentioned problems, considerable progress has been made in the development of the infrastructure needed to support the Semantic Web. In particular, there has been impressive progress in the development of languages and tools for content annotation and for the design and deployment of ontologies [16].

## 5. CONCLUSION

The explosive growth in both the range and quantity of Web content has highlighted some serious shortcomings in the hypertext paradigm such as difficulty in locating the required content using search and browse, and answering more complex queries along with more general information retrieval, integration, sharing and processing etc. Existing approaches are very useful, but they do not solve the problems without human intervention. This is the ultimate goal of the Semantic Web to allow data to be shared effectively by wider communities, and to be processed automatically by tools as well as manually by giving machine accessible semantics to annotations. This is achieved by using ontologies which are rich conceptual schemas to give formally defined meanings to the terms used in annotations, transforming them into semantic annotations.

## 6. REFERENCES

[1] T. Berners-Lee, J. Hendler, and O. Lassila, “The semantic web," Scientific American, vol.0501, May 2001.

[2] N. Balani, “The future of the web is semantic search," www.developerworks.com.

[3] E. T. O Neill, B. F. Lavoie, and R. Bennett, “Trends in the evolution of the public web," D-Lib Magazine, vol. 9, no. 4, Apr 2003.

[4] S. Decker, S. Melnik, F. van Harmelen, D. Fensel, M. Klein, J. Broekstra, M. Erdmann, and I. Horrocks, “The semantic web: the roles of xml and rdf," Internet Computing, IEEE, vol. 4, no. 5, pp. 63 - 73, Sep-Oct 2000.

[5] A. Swartz, “The semantic web in breadth," http://logicerror.com/semanticWeb-long, last Access: Aug 2012.

[6] S. B. Palmer, “The semantic web : An introduction," http://infomesh.net/2001/swintro/#simpleData, last Access: Aug 2012.

[7] G. Stumme, A. Hotho, and B. Berendt, “Semantic web mining," Web Semantics: Science, Services and Agents on the World Wide Web, vol. 4, no. 2, pp. 124-143, Jun 2006.

[8] I. Horrocks, “Ontologies and the semantic web," in Magazine of Communications of the ACM - Surviving the data deluge, vol. 51, no. 12. New York, NY, USA: ACM, , pp.58 – 67,Dec 2008.

[9] D. Fensel, F. van Harmelen, I. Horrocks, D. McGuinness, and P. Patel-Schneider, “Oil: an ontology infrastructure for the semantic web," Intelligent Systems, IEEE, vol. 16, no. 2, pp.38-45, Apr 2001.

[10] D. L. Noy, N. F.and McGuiness, “Ontology development 101: A guide to create your first ontology," http://www.ksl.stanford.edu/people/dlm/papers/ontologytutorial-noy-mcguinness.pdf, last Access:May 2012.

[11] H. O. Nigro, S. E. G. Csaro, and H. X. Daniel, “Data Mining with Ontologies: Implementations, Findings, and Frameworks. 701 E. Chocolate Avenue, Suite 200, Hershey PA 17033:Information Science Reference, 2008.

12] H. Dai and B. Mobasher, “Using ontologies to discover domain-level web usage profiles," in Proceedings of the Second Semantic Web Mining Workshop at PKDD 2001, km.aifb.unikarlsruhe.de/semwebmine2002/papers/full/bamshad.pdf, Aug 2002.

[13] D. Vallet, M. Fern_andez, and P. Castells, “An ontology-based information retrieval model," The Semantic Web: Research and Applications, pp. 103-110, 2005.

[14] M. Eirinaki, D. Mavroeidis, G. Tsatsaronis, and M. Vazirgiannis, “Introducing semantics in web personalization: The role of ontologies." vol. 4289. Springer, , pp. 147-162, Oct 2005.

[15] D. Antoniou, M. Paschou, E. Sourla, and A. Tsakalidis, “A semantic web personalizing technique: The case of bursts in web visits," in Semantic Computing (ICSC), 2010 IEEEFourth International Conference on, Sep 2010, pp. 530 -535.

[16] J. Hoxha, M. Junghans, and S. Agarwal, “Enabling semantic analysis of user browsing patterns in the web of data," arXiv preprint arXiv:1204.2713, 2012.