

“Is Semantic Web Real or Hype?”

Hamid R. DARVISH¹

Abstract

In this paper, first, I describe semantic web and second, I elaborate on four blocks of semantic web (XML, Resource Description Framework (RDF), RDF Schema, and Ontology) and finally some example of current semantic web tools and application. Semantic web is an extension of HTML (Hyper Text Mark up Language) and XML (Extensible Markup Language). XML and HTML play big part on semantic web implementation. The semantic web has been supported by W3C (www.w3c.org) organization, an international organization body for World Wide Web standardization.

Key words: *Semantic Web, Ontology, XML, HTML, RDF*

Introduction

In late 1980s, Tim Berners-Lee invented HTML (hypertext mark up language), while he was working for CERN (the European Particle Physics Laboratory) as an independent contractor [1]. Mr. Lee reveals in his book “Weaving the web” that Web pages are understandable by human beings but they are not comprehensible by machines. HTML tags are untreatable by only human readers but not computers. For example, for a chiropractics doctor office web page look like:

```
<h1> Agiliates physiotherapy center </h1>
```

Welcome to the home page of the Agiliyates physiotherapy center.

Do you feel pain? Let our (lovely secretary) and Steve Matthews take care of your body and soul.

¹ Computer Engineering Department, Çankaya University, Ankara, Turkey
E-mail: darvish@cankaya.edu.tr

```
<h1> consultation hours </h1>
Mon 11 am – 7pm <br>
Tue 11 am – 7pm<br> and so on.....
```

The format above is understandable by humans, but key-word searches may only find the word physiotherapy and consultation hours [2]. Semantic web can solve this problem by adding new meaningful tags to the web site. Semantic web adds Metadata tags to the website. Metadata is data about data which describe data in the web page. For example,

If we redefine the tags above such in XML format, as in tabel1 shows.

Table 1.

<pre><company> < treatmentofferd>physiotherapy </treatmentofferd> <company name> Agiliates physiotherapy Center </company name> <staff> < therapist> steve mattews </therapist > <secretary> Kelly Towsend </Secretary> </staff> </company></pre>

XML is an extension of HTML that gets the HTML in a structured format. But still does not add semantic flavor to the tags. RDF (Resource Description Framework) comes into rescue the XML. In other word, RDF interprets the eternal presentation of the web. Semantic web consist of seven layers: (figure.1) [3]

XML Schema : A language used to define the structure of specific XML languages.

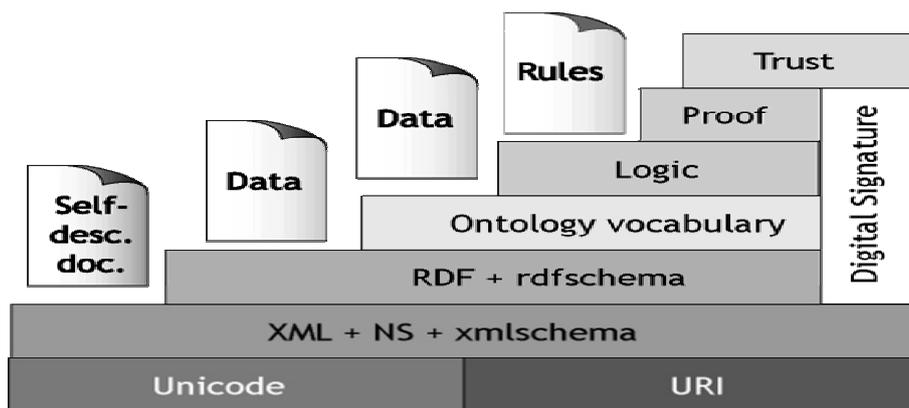


Figure 1. architecture diagram from Tim Berners-Lee's XML 2000 keynote address

Resource Description Framework (RDF): Resource Description Framework is an XML-based language to describe resources. In other words, instead of marking up the internals of a document, RDF captures meta data about the “externals” of a document.

RDF Schema : A framework that provides a means to specify basic vocabularies for specific RDF application languages to use.

Ontology : An ontology formally defines a common set of terms that are used to describe and represent a domain.

Software Agents : SW agents are programs that collect Web content from diverse sources, process the information and exchange the results with other programs.

Logic and Proof : Logical reasoning is used to establish the consistency and correctness of data sets and to infer conclusions that are not explicitly stated but are required or consistent with a known set of data. Proofs trace or explain steps of logical reasoning.

Trust : A means of providing authentication of identity and evidence of the trustworthiness of data, services and agents.

Why Do We Need the Semantic Web?

The Semantic Web is not just for the World Wide Web. It represents a set of technologies that will work equally well on internal corporate intranets. This is analogous to Web services representing services not only across the Internet but also within a corporation’s intranet. So, the Semantic Web will resolve several key problems facing current information technology architectures [4].

Describing relationships between data items

Ontology sets formal and explicit specification of domain models – and their relationships. For example, figure 2 shows ontology for a class animal [5].

Semantic Web Technology on Use

Oracle new pressroom runs on a semantic web search engine since December 2006. As Justin Kestelyn notes “Oracle Secure Enterprise Search serves as the eyes and ears of the Seamark navigator Semantic Web engine by crawling the available content and metadata and the Oracle RDF Store (which is fully integrated with Oracle Database out

```
class-def animal                % animals are a class
class-def plant                 % plants are a class
    subclass-of NOT animal      % of things that are not animals
class-def carnivore             % carnivores are a class
    subclass-of animal          % which is a subclass of animals
    slot-constraint eats        % that eat animals
        value-type animal
class-def herbivore             % herbivores are a class
    subclass-of animal          % which is a subclass of animals
    slot-constraint eats        % that eat plants
        value-type plant
class-def springbok             % springboks are herbivores
    subclass-of herbivore
class-def lion                  % lions are carnivores
    subclass-of carnivore
    slot-constraint eats        % that eat herbivores
        value-type herbivore
```

Figure 2. Ontology models the domain in terms of classes, sub classes and their properties.

of the Box) serve as its memory by sorting RDF data[6]. Another example, HaKia“ is a semantic web search engine based on New York, USA [7], which uses fuzzy Logic and natural Language methods for its search engine. HAKIA“ suppose to release the Turkish version in future. Protégé is an open source Ontology editor and Knowledge-base framework that is widely used in academia designed by Stanford University [8]. Protégé is based on Java, is extensible, and provides a plug-and-play environment that makes it a flexible base for rapid prototyping and application development.

Conclusion

Since its inception, Semantic Web technology has had its pros and cons. One of advantages is *Semantic Web* improves the search technique since it uses well structure web domain using well design Ontology. Nevertheless, the opponents argue that a well structure Ontology is hard to design. Therefore, it is difficult to achieve a real operational semantic web application. Semantic Web technology is in infancy; it matures as more research is on the way. Semantic Web activity is a kind of AI(Artificial Intelligent) that span on several multidisciplinary fields such as cognitive science, computer science, logic, mathematics and information studies and content analysis.

REFERENCES

Berners-Lee's original proposal to CERN. World Wide Web Consortium (March 1989).

A Semantic Web Primer, MIT Press, 2004, Grigoris Antoniou and Frank van Harmelen <http://mitpress.mit.edu/catalog/item/default.asp?ttype=2&tid=10140>

<http://www.w3c.org> (w3c vision for Semantic web)

Michael C. Daconta, Leo J. Obrst, Kevin T. Smith, "The Semantic Web-A Guide to the Future of XML, Web Services, and Knowledge Management", Wiley Publishing, Inc.

<http://www.hp1.hp.com> - Introduction to Semantic Web Technology

www.oracle.com/oraclemagazine (Page 18, December 2006)

www.hakia.com

http://protege.stanford.edu/doc/tutorial/get_started/why.html.