Information Science and Technology
ISSN: 0976-917X & ISSN: 0976-9188, Volume 2, Issue 1, 2012, pp.-20-23.
Available online at http://www.bioinfo.in/contents.php?id=77

SEMANTIC WEB TECHNOLOGY

AWES SIDDIQUI AND RUCHA DESHMUKH

Computer Science, SGBAU, Sipna Coet Amravati, Maharshtra, India.
*Corresponding Author: Email- awes.siddiqui@gmail.com, ruchadeshmukh@gmail.com

Received: February 21, 2012; Accepted: March 15, 2012

Abstract- That we need a new approach to managing information is beyond doubt. The technological developments of the last few decades, including the development of the World Wide Web, have provided each of us with access to far more information than we can comprehend or manage effectively. The Semantic Web and Semantic Web technologies offer us a new approach to managing information and processes, the fundamental principle of which is the creation and use of semantic metadata. The Semantic Web initiative of the World-Wide Web Consortium (W3C) has been active for the last few years and has attracted interest and scepticism in equal measure. The term 'Semantic Web' is one which is widely used, often without much care or understanding of its origins and meaning. Using semantics our systems can understand where words or phrases are equivalent.
Key words-

Citation: Awes Siddiqui and Rucha Deshmukh (2012) Semantic Web Technology. Information Science and Technology, ISSN: 0976-917X & ISSN: 0976-9188, Volume 2, Issue 1, pp.-20-23.

Copyright: Copyright©2012 Awes Siddiqui and Rucha Deshmukh. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Introduction

The goal of the semantic web is to be “a web talking to machines”, i.e. in which machines can provide a better help to people because they can take advantage of the content of the Web. The information on the web should thus be expressed in a meaningful way accessible to computers.

This definition is easily related to what already exists on the web: wrappers for extracting data from regularly structured pages, natural language analysis for extracting web page contents, indexing schemes, syndication facilities for broadcasting identified web resources. Much of this is painful and fragile: the semantic web should make it smart and robust. The semantic web can also be thought of as an infrastructure for supplying the web with formalised knowledge in addition to its actual informal content. One of the challenges of the current semantic web developments is the design of a framework in which all these understanding can collaborate, because the full benefit of the semantic web can only be attained when computers relate resources from various sources.

The third common use of the term Semantic Web is to identify a set of technologies, tools and standards which form the basic building blocks of a system that could support the vision of a Web imbued with meaning. The Semantic Web has been developing a layered architecture, which is often represented using a diagram first proposed.

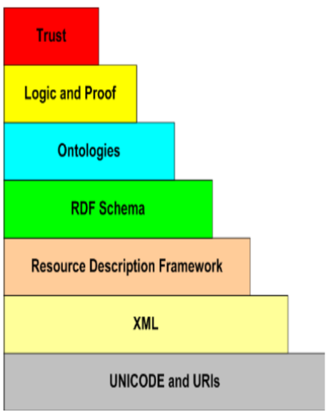


Fig. 1- Layers of Semantic Technology

We describe briefly these layers from the Fig 1.

A. Unicode and URI

Unicode, the standard for computer character representation, and URIs, the standard for identifying and locating resources (such as pages on the Web), provide a baseline for representing characters used in most of the languages in the world, and for identifying resources.

B. XML

XML and its related standards, such as Namespaces, and Schemas, form a common means for structuring data on the Web but without communicating the meaning of the data. These are well established within the Web already.

C. Resource Description Framework

RDF is the first layer of the Semantic Web proper. RDF is a simple metadata representation framework, using URIs to identify Web-based resources and a graph model for describing relationships between resources. Several syntactic representations are available, including a standard XML format.

D. RDF Schema

A simple type modelling language for describing classes of resources and properties between them in the basic RDF model. It provides a simple reasoning framework for inferring types of resources.

E. Ontologies

A richer language for providing more complex constraints on the types of resources and their properties.

F. Logic and Proof

An (automatic) reasoning system provided on top of the ontology structure to make new inferences. Thus, using such a system, a software agent can make deductions as to whether a particular resource satisfies its requirements (and vice versa).

G. Trust

The final layer of the stack addresses issues of trust that the Semantic Web can support. This component has not progressed far beyond a vision of allowing people to ask questions of the trustworthiness of the information on the Web, in order to provide an assurance of its quality.

Semantic web themes

Today, the W3C positions the Semantic Web into a number of areas:

- Linked Data
- Vocabularies
- Query
- Inference
- Vertical Applications

Lets briefly look at these five areas with this simple scenario:
"A traffic accident at the corner of Garden St and Central Ave in Eveleigh, Sydney"

A. Linked Data

The web is full of data and in order for the Semantic Web to operate fully, the "web of data" needs to organised in such a way that it is available in a standard format and relationships between data are visible. Linked Data is the ability to make explicit links between data and applications mayexploit the extra knowledge from other data sources by virtue of integrating information from several data sources, thereby providing a much better user experience. Linked data is fundamental to the Semantic Web and works based on the use of unique identifiers (URIs) for entities.

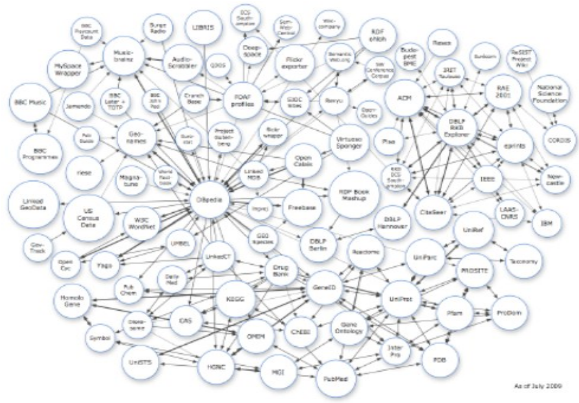


Fig. 2- Cloud of interlinked data

Fig 2. shows the growing community of Linked Data sources that form the Cloud of interlinked data. The power of the links provides consistency across data sources and exposes
relationships between data that can lead to greater knowledge creation and user experiences.
In our scenario, we may identify "Garden St" with a unique identifier. This means that any other
application or service may also "link" to Garden St and these applications know explicitly what
this entity is about and, more importantly, that it is the same entity. For example, the Sydney City Council plans to resurface "Garden Street" with new bitumen and even though they call it slightly differently, they use a unique identifier so that other authorities know exactly which entity is being referred to.

B. Vocabularies

Vocabularies define the concepts and relationships used to describe and represent an application
domain. Vocabularies are used to classify the concepts, represent possible relationships, and
define constraints on using those entities. Typically, vocabularies can range from very simple (eg a dozen entities) to be very large and complex (eg tens of thousands of entities). There is no clear
division between the what is referred to as "vocabularies" and "ontologies". The current trend is to use the word "ontology" for large, complex, formal collections of entities, and "vocabulary" for smaller cases.
In our scenario, we may classify "Garden St" and "Central Ave" as instances of types (classes) of "Roads" and that "Street" and "Avenue" are all subtypes of Road. Additionally we may define a "Corner" where at least two roads intersect.

C. Query

The Semantic Web needs programatic mechanisms to retrieve all the data, just like relational databases or XML need specific query languages. The Semantic Web, typically represented using the RDF model, needs its own query language and services. In our scenario, applications may wish to find out all current traffic incidents in Eveleigh or how many accidents have occurred on Garden St. The Semantic Web defines are way we can express these queries and the return the results to the client application.

D. Inference

Inference on the Semantic Web can be characterised by the automatic discovery of new relationships. Since the Semantic Web data is formally modelled with Ontologies via a set of named relationships, inferences can be extracted or explicitly added or returned from a query.
The Semantic Web also supports inference measures from rule sets to discover and generate new relationships based on existing ones.
In our scenario, even though we did not explicitly state this in our original data, we can infer that since Garden St is a type of "Street", which is a subtype of "Road", then Garden St is also a "Road". This is classic object-oriented inferencing. More complex inference mechanisms are extremely application specific, for example, if I know that a "party" event will be held on "Garden St", then I could infer that this is a "street party".

E. Vertical Applications

Vertical applications areas are used to explore how Semantic Web technologies can help improve operations, efficiencies, and provide better user experiences. The two most prominent areas for Semantic Web adoption are Health Care/Life Sciences and e-Government sectors as these sectors provide valuable use cases, feedback, and deployment scenarios used to improve the range of Semantic Web technologies. Of course, the Semantic Web is a generic set of technologies and can be applied to any sector/ community.
In our scenario, there would a number of different government agencies involved in responding to the traffic accident. State-based Emergency Services (eg Police, Ambulance) and Local Council services (eg to repair the damaged traffic lights) may all respond. The same information messages need to be transmitted and understood by these different agencies. In some cases, the message maybe mapped to local vocabularies or entities translated to alternatives (eg Eveleigh maybe translated to a postcode, or the geospatial coordinates of the road intersection maybe preferred for some agencies).

Semantic web architectures

An Enterprise Architecture is a rigorous and complete description of an organization and how it is decomposed into sub-systems, how the sub-systems are related and dependent, the terminology used, the external environment, and the principles and goals of the organisation. An Enterprise Architecture presents an organisation with the unique circumstances to identify opportunities for improvement and to better meet its overall objectives.
The most common layers of an Enterprise Architecture include:

- Business - strategic documents, business processes, capabilities.
- Information - metadata, information flow, data models.
- Applications - software, functional systems, application interfaces.
- Technology - middleware platforms, network infrastructure, programming languages

To date, most Enterprise Architectures have not utilised nor deployed, in any systemic way, Semantic Web technologies.
An example of an Enterprise Architecture is from the Australian Government Information Management Office (AGIMO) in which they define five layers of the Australian Government Architecture Reference Model as shown in Figure.

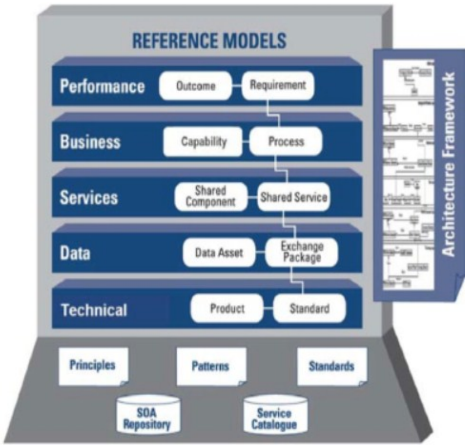


Fig. 3- Architecture of Semantic Web

Each of the layers represent and is further expanded into a new reference model. For example, the Data Reference Model, shown in Fig 3, shows it's three sub-systems:

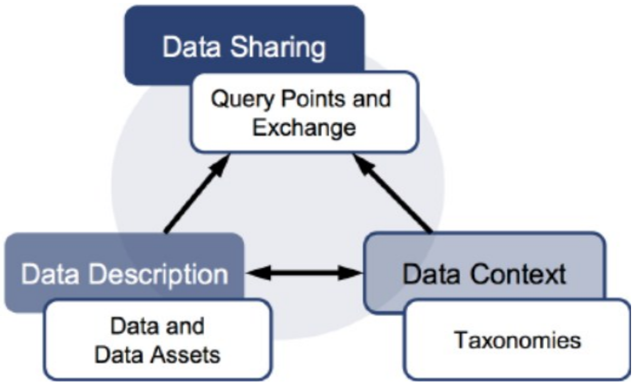
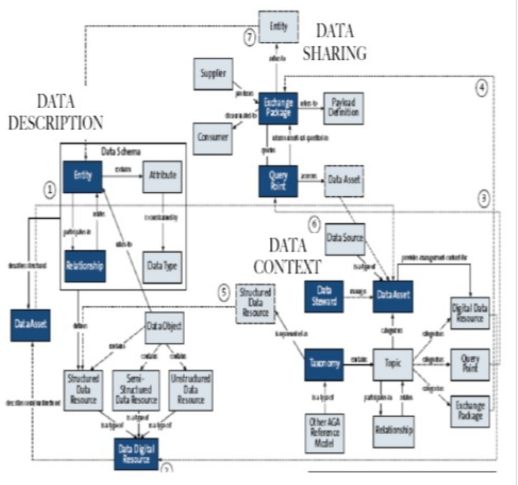


Fig. 4- Data Reference Model

- Data Description - a means to uniformly describe data.
- Data Context - the categorization of data according to taxonomies.
- Data Sharing - access, query and exchange of data between parties.



The Data Reference Model further describes and decomposes the three subsystems into an Abstract Model, and importantly, how they interact and are related, as shown in Fig 4.

The Data Description area of the abstract model identifies the various data types and their interrelationships. The focus of this area is the identification of entities and the designation of the information describing them. The Data Context area of the abstract model identifies the structures used for resources in the form of a set of terms collectively called 'categorisation' or 'classification' taxonomies. The reference model, at this point, then refers to the Semantic Web with "implementation of Taxonomies could take the form of eXtensible Markup Language
(XML) Topic Maps, Web Ontology Language (OWL) hierarchies or ISO11179 Classification schemes". Finally, the Data Sharing area of the abstract model conveys an architectural pattern for the sharing and exchange of data with examples to support their business needs.

Even with a cursory analysis, after looking at the entities in Fig. 5, we can immediately see where some of the Semantic Web technologies could form part of the core features of the Data
Reference Abstract Model:

- Data Description: RDF
- Data Context: OWL
- Data Sharing: SPARQL

In this case, there is a promising opportunity to overlay Semantic Web technologies onto the Australian Government's enterprise architecture.

However, in general, there has been little comprehensive, and large-scale, attempts to create Enterprise Architectures utilising Semantic Web technologies as the foundation information model and technology platform. There are some small pockets of activity in this space, ranging from individuals (eg the "Layered Semantic Enterprise Architecture") to niche companies promoting the Semantic Web as part of their architecture platform (eg "Towards Executable Enterprise Models" from Top Quadrant.

Conclusion

The Semantic Web has great potential, and with direct application to the HE and FE sector. However, it has been a long time in development and does require an investment of time, expertise and resources. Nevertheless, the time does seem right to start to think how best to use the simpler applications of the technology.

In the future the Semantic Web may not even be noticeable. The tools of the Semantic Web will be integrated into Virtual Learning Environments and Virtual Research Environments on our desktops, as well as in browsers and search engines. What we will have is a richer experience of IT that is better able to deliver the right information at the right time in the right way, so we can get on with the serious business of research and teaching.

The Semantic Web offers plenty of opportunities for Enterprise Architectures. Given one of the
core principles of Enterprise Architectures is a shared vision of an organisation, such a vision
needs to be expressed in clear, common, and precise semantics. RDF and OWL provide the
formal ability to provide this objective.

Future

The Semantic Web is no different and has both been criticised for being too complex and "academic" and showered with praise for its new "exciting" technologies that drive web-based user experiences.

- The potential future path of the Semantic Web may include:
- Redefined RDF language (now being discussed).
- Merge the RDF and OWL languages into one comprehensive language with various levels of semantic support.
- Support a range of simple (eg "microdata") and complex syntaxes seamlessly.
- Support greater Vocabulary/Ontology management.

References

[1] W3C Semantic Web homepage. *http://www.w3.org/standards/semanticweb/*.
[2] Tim Berners-Lee, James Hendler and Ora Lassila (2001) *Scientific American Magazine*.
[3] W3C RDF homepage *http://www.w3.org/standards/techs/rdf*.
[4] Semantic Web Tools *http://esw.w3.org/SemanticWebTools*.
[5] Koivunen MR., Miller E. (2001) *Proceedings of the Semantic Web Kick-off Seminar in Finland*.
[6] Hayes P. (2004) *http://www.w3.org/TR/rdf-mt/*.
[7] Educational Modelling Language (EML) *http://eml.ou.nl/eml-ou-nl.htm*.
[8] Johan Hjelm (2001) *Creating the semantic web with RDF*, John Wiley (UK).
[9] Dieter Fensel, (2001) *Ontologies: silver bullet for knowledge management and electronic commerce*, Berlin (DE).
[10] Tim Berners-Lee, James Hendler and Ora Lassila (2001) *Scientific American*, 284(5):35-43, *http://www.scientificamerican.com/2001/0501issue/0501berners-lee.html*.