STEPPING TOWARDS THE SEMANTIC WEB

Jatinder Kumar1, Navjot Kaur 2, Rakesh Kumar 3, Rajwinder Kaur 4, Ritu Bala 5(HOD) 1,2,3,4,(Faculty), 5 (HOD)* **Department of Computer Science and Engineering, 5 (HOD) Department of Information Technology, 2 (Coordinator MCA, M.Tech) Sai Institute of Engineering and Technology, (Sai Technology Campus) VPO Manawala-143115, Amritsar, Punjab, India,**
1 jkteji@yahoo.co.in 2 er.navjotdadwal@gmail.com, 3 er_rakeshsharma@rediffmail.com
4 rajwinder1179@yahoo.co.in 5 ritu_bala2002@yahoo.com

Abstract **- The Semantic Web aims at machine agents that thrive on explicitly specified semantics of content in order to search, filter, condense, or negotiate knowledge for their human users. A core technology for making the Semantic Web happen, but also to leverage application areas like Knowledge Management and E-Business, is the field of Semantic Annotation, which turns human-understandable content into a machine understandable form.The generation of information on the web is increasing exponentially. The main need of web users is to access subject relevant and well-defined meaningful information. Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users. This study is, after analyze of Semantic Web Application and Web ontology characteristics, established inference-based web ontology that use expression of Description Logic and verified inference that established Web Ontologies. We can select the level of ontology from Meta-data, Domain to Instance Ontologies. In computer science and information science, ontology is a formal representation of 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 define the domain. Keyboard** Semantic web, RDF, Ontology**,**

1. INTRODUCTION

"The Semantic Web is an extension of the current web in which information is given well defined meaning, better enabling computers and people to work in cooperation". Tim Berners-Lee The Semantic Web is multidisciplinary and heterogeneous. Many Semantic Web researchers maintain close ties to neighboring disciplines which provide methods or application areas for their work. However, the Semantic Web has now established itself as a research field in its own rights. The Semantic Web is an effort by the W3C to enable integration and sharing of data across different organizations and applications. Though called the Semantic Web, the W3C envisions something closer to a global database than to the existing World-Wide Web. In the W3C vision, users of the Semantic Web should be able to issue structured queries over all of the data on the Internet, and receive correct and well-formed answers to those queries from a variety of different data sources that may have information relevant to the query [3]. Like other languages, the semantics of the Semantic Web depend on those of Its grammar and terms, their "context", and the applications that use them. The generation of information on the web is increasing exponentially. The main need of web users is to access subject relevant and well-defined meaningful information. Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users. Semantic web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. It is evident that, understanding level of a machine and human beings are different. The traditional web matches word by word and provides the result. But semantic web (SW) understands the meaning of each word analyzes domain and defines the relationship among different words. bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users. Semantic web is an extension of the current Web in which information is given well-defined meaning, better enabling computers and people to work in cooperation. It is evident that, understanding level of a machine and human beings are different. The traditional web

* is the main author

matches word by word and provides the result. But semantic web (SW) understands the meaning of each word analyzes domain and defines the relationship among different words.

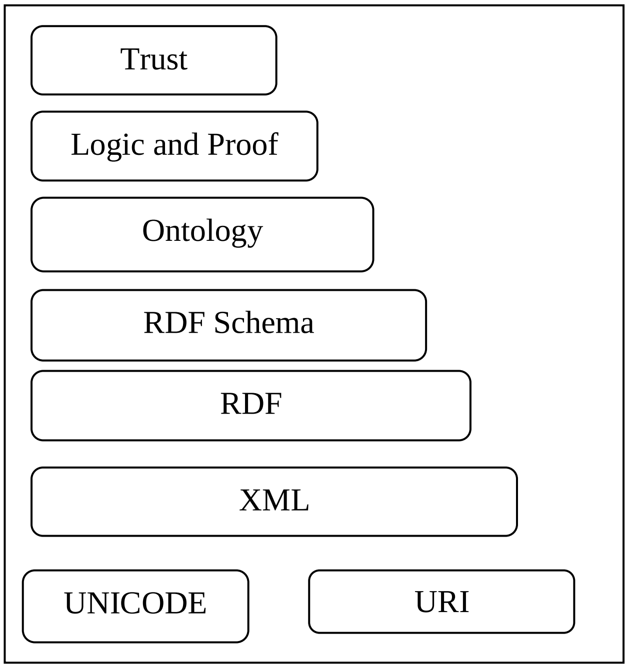

Fig.:1 Layered Architecture of Semantic web

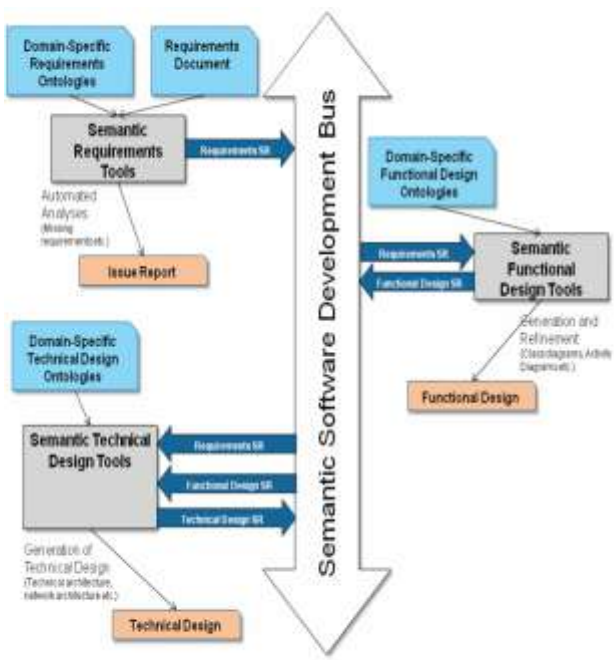

Fig.:2 Semantic Software Development Bus [4]

2. Historical Evolution of Semantic Web

There is not any specific research about the web generations from the web advent in formal. outlined three qualities of the web based on an analytical distinction. Web 1.0 was introduced as a tool for thought, web 2.0 as a medium for communication between humans and web 3.0 as networked digital technology to support co-operation of humans. San Murugesan [6] described web 2.0 as the second phase in the web's evolution in which technologies, services, development approaches and tools of web 2.0 were introduced in detail.

Table.1 Comparison of Web 1.0 and Web 2.0

Web 1.0	Web 2.0
Reading	Reading/Writing
Companies	Communities
Client-Server	Peer to Peer
HTML, Portals	XML, RSS
Taxonomy	Tags
Owning	Sharing
IPOs	Trade sales
Netscape	Google
Web Forms	Web Applications
Screen scraping	API,s
Dialup	Broadband
Hardware costs	Bandwidth costs
Lectures	Conversation
Advertising	Word of mouth
Services sold over the web	Web services
Information portals	Platforms

Table.2 Comparison of Web 2.0 and Web 3.0

Web 2.0	Web 3.0
Read /Write Web	Portable Personal Web
Communities	Individuals
Sharing Contents	Consolidating dynamic contents
Blogs	lifestreams
Ajax	RDF
Wikipedia, google	Dbpedia, igooge
Tagging	User engagement

3. Relationship to the hypertext web Limitations of HTML

Many files on a typical computer can be loosely divided into documents and data. Documents like mail messages, reports, and brochures are read by humans. Data, like calendars, address books, play lists, and spreadsheets are presented using an application program which lets them be viewed, searched and combined in many ways. Currently, the World Wide Web is based mainly on documents written in Hypertext Markup Language (HTML), a markup convention that is used for coding a body of text interspersed with multimedia objects such as images and interactive forms. Metadata tags,

for example <meta name="keywords" content="computing, computer studies, computer">
<meta name="description" content="Cheap widgets for sale">
<meta name="author" content="John Doe">

With HTML and a tool to render it (perhaps web browser software, perhaps another user agent), one can create and present a page that lists items for sale. The HTML of this catalog page can make simple, document-level assertions such as "this document's title is 'Widget Superstore'", but there is no capability within the HTML itself to assert unambiguously that, for example, item number A142398756 is an Artificial Intelligence with a retail price of Rs. 350, or that it is a consumer product. Rather, HTML can only say that the span of text " A142398756" is something that should be positioned near " Artificial Intelligence " and " Rs. 350", etc. There is no way to say "this is a catalog" or even to establish that " Artificial Intelligence " is a kind of title or that " Rs. 350" is a price. There is also no way to express that these pieces of information are bound together in describing a discrete item, distinct from other items perhaps listed on the page.

4. Relationship to object oriented programming

A number of authors highlight the similarities which the Semantic Web shares with object-oriented programming (OOP).Both the semantic web and object-oriented programming have classes with attributes and the concept of instances or objects. Linked Data uses Dereferenceable Uniform Resource Identifiers in a manner similar to the common programming concept of pointers or "object identifiers" in OOP. Dereferenceable URIs can thus be used to access "data by reference". The Unified Modeling Language is designed to communicate about object-oriented systems, and can thus be used for both object-oriented programming and semantic web development.

5. Ontology Development

An ontology defines a common vocabulary for researchers who need to share information in a domain. It includes machine-interpretable definitions of basic concepts in the domain and relations among them[2]. Why would someone want to develop ontology? Some of the reasons are:

• To share common understanding of the structure of information among people or software agents
• To enable reuse of domain knowledge
• To make domain assumptions explicit
• To separate domain knowledge from the operational knowledge
• To analyze domain knowledge

Ontologies are used in artificial intelligence, the Semantic Web, systems engineering, software engineering, biomedical informatics, library science, enterprise book marking, and information architecture as a form of knowledge representation about the world or some part of it. The creation of domain ontologies is also fundamental to the definition and use of an enterprise architecture framework[1].

For the development of Ontologies of profile and community, we used Protégé 3.3.1 , a free, open source platform, a popular tool of Stanford University for developing Domain Ontology. Ontology Development was followed step-by-step, as defined by Noy and McGuinness.

6. Steps of Ontology Development

Following are the steps

1. Determine the domain and scope of the ontology
2. Consider reusing existing ontologies
3. Enumerate important terms in the ontology
4. Define the classes and the class hierarchy
5. Define the properties of classes—slots
6. Define the facets of the slots
7. Create instances

7. Characteristics of Web Ontology

First of all, the analysis results regarding Web Ontology characteristics to be suitable for a Semantic Web application are as follows.

1. Web ontology shall express formally and explicit to knowledge so that ontology engineer and a machine can understand.
2. Using syntax interaction, increase reuse of parsing data, or shall improve semantic interaction etc. than web environments of now.
3. An expression of information and inference shall be added to the current Web with bases. A knowledge process, share, reuse of web-based shall be possible between applications.
4. A construction language of Web ontology uses a Web ontology language like DAML+OIL, RDF, OWL

8. OWL and XML Schema Standards

One question that comes up when describing another XML/Web standard is why there is the need to use OWL language instead of the well known XML and XML Schema standards. As the W3C states at the OWL Web Ontology Language Guide (Smith, et al., 2004), there are two answers to this question.

· An ontology differs from an XML schema in that it is a knowledge representation, not a message format.

· One advantage of OWL ontologies will be the availability of tools that can reason about them.

Tools will provide generic support that is not specific to the particular subject domain, which would be the case if one were to build a system to reason about a specific industry-standard XML schema.

9. The Species of OWL

The OWL language provides three increasingly expressive sublanguages designed for use by specific communities of implementers and users.

· OWL Lite supports those users primarily needing a classification hierarchy and simple constraints. Owl Lite has a lower formal complexity than other type of OWL.

· OWL DL supports those users who want the maximum expressiveness while retaining computational completeness and decidability. OWL DL is so named due to its correspondence with Description Logics, a field of research that has studied the logics that form the formal foundation of OWL.

· OWL Full is meant for users who want maximum expressiveness and the syntactic freedom of RDF with no computational guarantees. OWL Full allows an ontology to augment the meaning of the pre-defined (RDF or OWL) vocabulary.

The choice between OWL Lite and OWL DL depends on the extent to which users require the more expressive restriction constructs provided by OWL DL. The choice between OWL DL and OWL Full mainly depends on the extent to which users require the meta-modeling facilities of RDF Schema (i.e. defining classes of classes). When using OWL Full as compared to OWL DL, reasoning support is less predictable. In the present methodology OWL Full was chose over OWL DL for making use of its special characteristics.

10. Challenges

Some of the challenges for the Semantic Web include vastness, vagueness, uncertainty, inconsistency and deceit. Automated reasoning systems will have to deal with all of these issues in order to deliver on the promise of the Semantic Web.

• Vastness: The amount of data published on the Web of Data has witnessed a tremendous growth in recent years. The Linked Open Data (LOD) project1 has contributed significantly to this growth. People are in fact massively following Tim Berners- Lee's advice to publish data on the Web following the "linked data" principles. The World Wide Web contains billions of pages.

• Vagueness: These are imprecise concepts like "young" or "tall". This arises from the vagueness of user queries, of concepts represented by content providers, of matching query terms to provider terms and of trying to combine different knowledge bases with overlapping but subtly different concepts. Fuzzy logic is the most common technique for dealing with vagueness.

• Uncertainty: These are precise concepts with uncertain values. For example, a patient might present a set of symptoms which correspond to a number of different distinct diagnoses each with a different probability. Probabilistic reasoning techniques are generally employed to address uncertainty.

• Inconsistency: These are logical contradictions which will inevitably arise during the development of large ontologies, and when ontologies from separate sources are combined. Deductive reasoning fails catastrophically when faced with inconsistency, because "anything follows from a contradiction". Defeasible reasoning and paraconsistent reasoning are two techniques which can be employed to deal with inconsistency.

• Deceit: This is when the producer of the information is intentionally misleading the consumer of the information. Cryptography techniques are currently utilized to alleviate this threat.

•Data Dependency: Knowledge engineering is not yet fully accustomed to the ubiquity ofinstance data. An example is current work on ontology and vocabulary alignment. The Ontology Alignment Evaluation Initiative (OAEI) annually specifies a set of ontologies for benchmarking alignment systems[5].

Refrences:

1) Gianluca Tonti1,2, Jeffrey M. Bradshaw1, Renia Jeffers1, Rebecca Montanari2, Niranjan Suri1, and Andrzej Uszok1, **"**Semantic Web Languages for Policy Representation and Reasoning: A Comparison of KAoS, Rei, and Ponder"2004.
2) Su-Kyoung, Kim, **"**Implementation of Web Ontology for Semantic Web Application**"**IEEE, 2007**.**
3) Muhammad Shahab Siddiqui, Dr. Zubair A. Shaikh, Dr. Abdul Rahman Memon, "Towards the Development of Human Community Ontology" IEEE, 2009.
4) Philip Cimin and Stefan Kopp, "Accessing the web of Data through embodied virtual characters" Semantic Web1, 83-88,2010.
5) Rinke and Hoekstra, "The knowledge reengineering bottleneck" Semantic Web1,111-115, 2010
6) San, Murugesan (2007), "Understanding Web 2.0", Journal IT Professional.
7) Eleni Tomai Maria Spanaki, " From ontology design to ontology plementation: A web tool for building geographic ontologies"2011

.