ISSN 2319-2720
Volume 2, No.3, July - September 2013
International Journal of Computing, Communications and Networking
Available Online at http://warse.org/pdfs/2013/ijccn01232013.pdf

An Investigation on Semantic Web

R. GOGULAKRISHNAN[1], Dr. K. THIRUMALAIVASAN[2], S. NITHIYA[3]
[1,3]Dept. of Computer Science, Saradha Gangadharan College, Puducherry, India
gogulakrishnan16@yahoo.com
[2]Dept. of ECE, Achariya College of Engineering and Technology, Puducherry, India
thirumalaivasan@ieee.org

ABSTRACT

The web is meant for human consumption rather than the machine consumption. At present scenario, everything on the web is machine readable, not machine understandable. The method for data handling in the traditional web (The Syntactic web) is tedious, time consuming, and also provides many unrelated information. The Syntactic web (Current Web) provides an interface for users to render HTML documents and retrieve linked documents with simple user interface commands. On the other hand, Semantic Web making the web more understandable by machines and provide accurate results. So Semantic Web is needed to express the information in a precise, machine interpretable form. This Web aims to convert the current web, which is dominated by unstructured and semi-structured documents into 'web of data'. In addition, Semantic web integrates the information in an intelligent way and providing semantic based access to the internet. Semantic is building an appropriate infrastructure for intelligent agents to run around the web performing complex action for their users extracting information from texts. Semantic web also provides automated information access based on Machine – processable semantics of data and heuristics that use these metadata. This paper presents a comparative study of Syntactic and Semantic web technologies.

Key words: Ontology, RDF, Semantic web, Syntactic web, XML.

1. INTRODUCTION

The Internet is the large container, and the Web is a part within the container. The World Wide Web (WWW) or "Web" is large software subset of the Internet dedicated to broadcasting Hyper Text Markup Language (HTML) pages. The web is viewed by using free software called web browsers. The web is based on Hyper Text Transfer Protocol (HTTP), the language which allows you and me to "jump" (hyperlink) to any other public web page. There are over 40 billion public web pages on the Web today.

Web was invented by Tim Berners-Lee (amongst others), a physicist working at CERN (European Organization for Nuclear Research). Tim Berners-Lee's original vision of the Web was much more ambitious than the reality of the existing (syntactic) Web [1].

Rest of the paper is organized as follows: Section 2 discuss about syntactic web and their architectural model. Semantic web and their layers are presented in Section 3. Semantic web services are detailed in Section 4. Section 5 concludes the paper.

2. SYNTACTIC WEB

The architectural model of Web is client-server and it is based on three main technological components:

- Universal Resource Identifiers (URI) that provide a global addressing scheme for documents
- Hyper Text Transfer Protocol (HTTP) that is a stateless request-reply protocol to retrieve documents by their URIs
- HTML, a markup language that can be used to specify the presentation structure of a document and to include links to other documents using their URIs.

The primary software components of the Web are *Web servers* (at the server side) that host HTML documents, and *Web browsers* (at the client side) that provide an interface for users to render HTML documents and retrieve linked documents with simple user interface commands.

The main use of the Web is *surfing*: a user retrieves, evaluates, and optionally reads documents with a Web browser. An attempt to manage the varied content types has led to the specification of a general markup language framework called eXtensible Markup Language (XML). Numerous domain-specific representation languages have been created on top of it: Scalable Vector Graphics (SVG) for vector graphics, Simple Object Access Protocol (SOAP) for messages, XML User interface Language (XUL) for graphical user interfaces, and so on [2].

2.1 LIMITATIONS OF HTML

➢ Not extensible and could not able to customize

- Could not able to accommodate special needs (e.g. mathematics, chemical formulae)
- Proprietary, vendor-specific tags to extend capabilities

➢ Only codes for display, no document structure and semantics

3. SEMANTIC WEB

The word 'semantic' stands for the 'meaning of'. The Semantic Web = a Web with the meaning. 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 co-operation [3]. The Semantic Web is given initiative attempts to define controlled vocabularies or ontologies which is set of conceptual terms labeled by Uniform Resource Locators (URLs) that can be used in XML documents to give XML structures the semantics required by automatic reasoning. Semantics based searches that find results according to user needs, which differ from existing web environment by providing services automatically.

Semantic Web technology has been developed very fast in the recent past years and continues to grow as the importance of knowledge and technologies working together for human benefits becomes a necessary part in all the known domains namely information technology, communication, economic, social, health and even political. Semantic Web is a strong and bigger collaboration between researchers and business corporations for industry solutions and products which use semantic web technology to increase profits and reduce costs [4].

Emerging Semantic Web Services Standards like OWL-S, Web Service Modeling Ontology (WSMO) and Semantic Annotations for WSDL and XML Schema (SAWSDL). Enrich Web Services Standards like Web Services for Devices (WSD) and Business Process Execution Language 4 Web Services (BPEL4WS). The Figure 1 shows the building layers of semantic web. The phases occurred in layers of semantic web is discussed as follows:

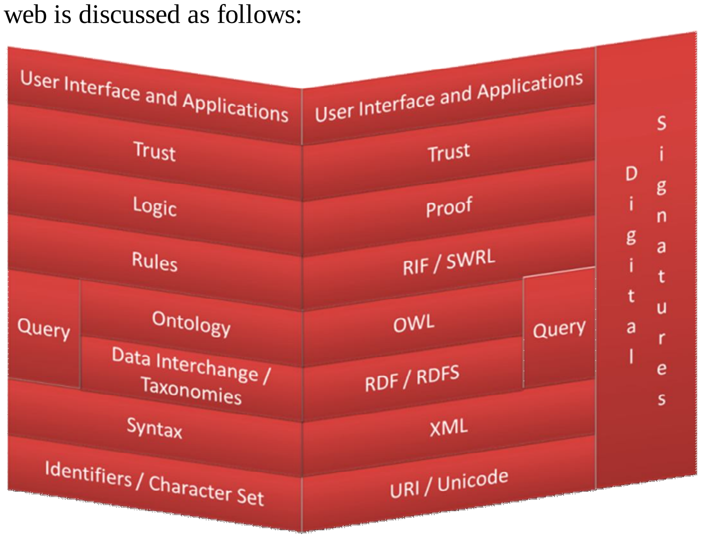

Figure 1: Building Layers of Semantic Web

3.1 UNICODE

Unicode is computing industry standard for consistent encoding, representation and handling of text expressed in most of the world's writing systems. Unicode provides a unique number for every character [5].

3.2 UNIVERSAL RESOURCE IDENTIFIERS (URIs)

URIs identifies resources; it is the central to the Semantic Web enterprise. A global naming convention (however arbitrary the syntax) provides the global network effects that drive the Web's benefits. URIs has global scope and that is interpreted consistently across contexts. Associating a URI with a resource means that anyone can link to it; refer to it; or retrieve a representation of it. URIs provides the grounding for both our objects and relations. They underpin the Semantic Web, allowing machines to process data directly. In this way, the Semantic Web shifts the emphasis from documents to data. Much of the motivation for the Semantic Web comes from the value locked in relational databases. To release this value, database objects must be exported to the Web as first-class objects and therefore must be mapped into a system of URIs [6].

3.3 EXTENSIBLE MARKUP LANGUAGE (XML)

XML is a first step in the design process. Metadata used within documents, not across documents. It is a prescriptive, not descriptive. It has no commitment on vocabulary and modeling primitives. XML more complicated but more powerful. It is a subset of Standard Generalized Markup Language (SGML) which omits some of SGML's more complex features in return for making it easier to process and specifically for making it easier to deliver on the Web. It is (like SGML) a Meta language [7].

A. The XML family of standards

- Extensible Stylesheet Language (XSL) is used to transform and render XML documents
- XML Linking Language (XLink) provides methods for creating internal and external links within XML documents, and associating metadata with those links
- XML Pointer Language (XPointer)

B. Extensible Hyper Text Markup Language (XHTML)

It is a family of XML markup language that mirror or extend versions of the widely used HTML, the language in which web pages are written. While HTML (prior to HTML5) was defined as an application of SGML, a very flexible markup language framework, XHTML is an application of XML, a more restrictive subset of SGML. Because XHTML documents need to be well-formed, they can be parsed using standard XML parsers - unlike HTML, which requires a lenient HTML-specific parser [8].

C. Some XML features

- DTD is allowed but not required; well-formed and valid documents
- Employs Unicode character set
- Markup minimization features not required
- Sophisticated stylesheet language (XSL) [7].

D. Limitations of XML

XML is not machine accessible meaning; it is accessible only to the people. For example, one can use an element as 'Author'; another can use it as 'Writer'. Here, human can make out that both are same, but how system can? This creates confusion when machines try to share data with each other.

E. Advantages of XML over HTML

- By defining own markup language; It can code documents more precisely
- Reflects structure and semantics of documents better searching and navigation
- Tagging/content separate from display
- Allow single document to be used many ways
- It places emphasis on descriptive rather than procedural markup;
- It distinguishes the concepts of syntactic correctness and of validity with respect to a document type definition;
- It is independent of hardware or software system [9].

F. Disadvantages of XML

- More difficult, demanding, and precise than HTML
- Lack of browser support/ end user application
- Still experiment/not solidified.

3.4 RESOURCE DESCRIPTION FRAMEWORK (RDF)

RDF is the foundation of the Semantic Web, which is a simple metadata representation framework that uses URIs to identify Web based resources and a graph model for describing relationships between resources or Documents. RDF is a language for expressing data models in XML syntax that provides an elemental syntax to structure the data. It provides the meaning to that structured data which is used to describe web resources. Semantic Web is built on XML language capacity to define ordinary schemes much closer to data representation. But the language that did marked the beginning of a real web of data was the RDF language with its triplets: subject, property and object (represented by URI-s) to form a direct, labeled graph which connects data [4].

Consisting of triples or sentences:

<subject, property, verb>

Ex. <Tolkien, wrote, The Lord of the Rings>

RDF essentially uses XML syntax. One needs to use RDF for integrating and exchange information in a meaningful way on the web. RDF is only specification language for expressing syntax and semantics [10].

A. RDF and RDFS

RDF and RDF Schema are XML based representational frameworks that intend to provide interoperable descriptions for web resources, defining ontologies or metadata schemas and their properties and relationships. RDFS extends RDF with standard ontology vocabulary:

Class, Property
Type, subClassOf
Domain, range
Notation: RDF(S) = RDF + RDFS

RDF offers a simple graph reference model. RDF was proposed in 1998 as a simple graph model, followed a year later by RDFS. RDF Schema (RDFS) offers a simple vocabulary and axioms for object-oriented modeling.

B. Terse RDF Triple Language (Turtle)

It is a format for expressing data in the RDF data model, similar to SPARQL. RDF, in turn, represents information using triples, each of which consists of a subject, a predicate, and an object. Each of those items is expressed as a Web URI [11].

C. N-Triples

It is a format for storing and transmitting data. It is a line based, plain text serialization format for RDF graphs, and a subset of the Turtle (Terse RDF Triple Language) format. N-Triples should not be confused with Notation 3 which is a superset of Turtle. N-Triples was primarily developed by Dave Beckett at the University of Bristol and Art Barstow at the W3C. N-Triples was designed to be a simpler format than Notation 3 and Turtle, and therefore easier for software to parse and generate. However, because it lacks some of the shortcuts provided by other RDF serializations such as CURIEs and nested resources, which are provided by both RDF/XML and Turtle [11].

D. Notation3 (N3)

It is more commonly known, is a shorthand non-XML serialization of RDF models, designed with human-readability in mind: N3 is much more compact and readable than XML RDF notation. The format is being developed by Tim Berners-Lee and others from the Semantic Web community. N3 has several features that go beyond a serialization for RDF models, such as support for RDF based rules. Turtle is a simplified, RDF-only subset of N3 [11].

E. Embedded RDF (eRDF)

It is as syntax for writing HTML in such a way that the information in the HTML document can be extracted (with an eRDF parser or XSLT Stylesheet) into RDF. This can be of great use for searching within data. It was invented by Ian Davis in 2005, and partly inspired by microformats, a simplified approach to semantically annotate data in websites.

F. Resource Description Framework in attributes (RDFa)

It is a W3C Recommendation that adds a set of attribute-level extensions to HTML, XHTML and various XML-based document types for embedding rich metadata within Web documents. The RDF data-model mapping enables its use for embedding RDF subject-predicate-object expressions within XHTML documents; it also enables the extraction of RDF model triples by compliant user agents. XHTML+RDFa is an extended version of the XHTML markup language for supporting RDF through a collection of attributes and processing rules in the form of well-formed XML documents. This host language is one of the techniques used to develop Semantic Web content by embedding rich semantic markup

G. RDF API for PHP (RAP)

RAP is a Semantic Web toolkit for PHP developers. It offers features for parsing, manipulating, storing, querying, serving, and serializing RDF graphs. RAP was started as an open source project by the Free University of Berlin in 2002 and has been extended with code contributions from the Semantic Web community. The core of RAP is two implementations of statement storages which hold RDF graphs either in memory or in a relational database. Around these storages RAP provides rich programming interfaces for manipulating RDF graphs on different abstraction layers. Furthermore, RAP supports RDFS inference as well as some OWL entailments, allowing programmers to work with implicit (virtual) statements. Various tools complement the RAP package: an up-to-date RDF/XML parser, an integrated RDF server, and a graphical user-interface for managing database-backed RDF models as well as an implementation of the RDQL query language. SPARQL has completed the layer-cake of the SW architecture providing a query language for RDF [12].

H. Simple Knowledge Organization System (SKOS)

SKOS language is an extensible RDF language to describe concept and content of concept schemes that include semantic relationships between these concepts. SKOS core represents the core model for expressing the basic structure and content of a concept scheme. SKOS Core Vocabulary is a set of RDF properties and RDFS classes that can be used to express the content and structure of a concept scheme as an RDF graph [13].

3.5 ONTOLOGY

In philosophy, ontology studies the nature of being and existence. The term ‘ontology’ is derived from the Greek words onto, which means being, and logia, which means written or spoken discourse. Ontology formally defined as set of terms that represents concepts within a particular subject area and also defines relationship between these terms in reusable and machine readable format.

Conceptualization: The language should choose an appropriate reference model, such as Entity-relationship model and object-oriented model, and provide corresponding ontology constructs to represent factual knowledge, such as defining the entities and relations in a domain, and asserting relations among entities.

Vocabulary: Besides the semantics, the language should also cover the syntax such as symbol assignment (i.e., assigning symbols to concepts) and grammars (i.e., serializing the conceptualism into explicit representation).

Axiomatization: In order to capture the semantics for inference, rules and constraints are needed in addition to factual knowledge. For example, we can use rules to generate new facts from existing knowledge, and to validate the consistency of knowledge. Web based knowledge sharing activities demand that human and/or machine agents agree on common and explicit ontologies so as to exchange knowledge and fulfill collaboration goals.

In order to share knowledge across different communities or domains, three requirements should be considered when developing explicit ontologies:

Extensibility: In the context of the Web, ontology engineers should be able to develop ontologies in an incremental manner: reusing the existing popular concepts before creating a new concept from scratch. For example, the concept woman can be defined as a sub-class of an existing concept person in WordNet vocabulary. This requirement demands an expressive common reference model as well as distributed symbol resolution mechanisms.

Visibility: Merely publishing knowledge on the Web does not guarantee that it can be readily understood by machines or human users. In order to make knowledge visible on the Web, additional common ontological ground on syntax and semantics is required between information publishers and consumers. This requirement is especially critical to machines since they are not capable of understanding knowledge written in an unfamiliar language.

Inferenceability: An ontology not only serves the purpose of representation, i.e. enumerating factual domain knowledge, but also serves the purpose of computation, i.e., enabling logical inference on facts through axiomatization. Hence, ontologies on the Web should provide constructs for effective binding with logical inference primitives and options to support a variety of expressiveness and computational complexity requirements [14].

A. Ontology Database

The ontology database used in the local search service proposed in this study was established by transforming it into an RDF triple using an N3 (Notation 3) method of W3C with a legacy database. The N3 method made it possible to automatically transform the table of the related database system to a 'record-field-data' and triple format of 'Subject-Predicate-Object'. This method not only guarantees the automatic generation of database instances but also integrity because the data stored in existing relational database systems is only a part of the detailed information comprising ontology. The field (column), record, and data used in a conventional database are mapped into triple predicates (Property, Predicate), Subject, and Object, respectively [15].

Ontology is a key to the Semantic Web. Some basic ontology languages are RDF and OWL (Web Ontology Language). Some of the development editors are altova, protégé, ontolingna, UNSPSC, Rosetta Net. Ontologies may be specified using RDF syntax. A model-driven architecture based approach for specifying semantic web service compositions through the use of a UML profile that extends class and activity diagrams. This profile is used in transformations that facilitate automatic construction of OWL-S specifications from UML diagrams. Conditions required by the composition, such as those on control constructs, are specified using OCL and transformed into SWRL during the construction process. OWL facilitates greater machine interpretability of Web content than that supported by XML, RDF by providing additional vocabulary along with a formal semantics [16].

B. OWL Language

OWL is based on Description Logics (DL) knowledge representation formalism. OWL (DL) benefits from many years of DL research: Well defined semantics, Formal properties well understood (complexity, decidability), Known reasoning algorithms, Implemented systems (highly optimised), Three species of OWL:

- OWL full is union of OWL syntax and RDF
- OWL DL restricted to FOL fragment
- OWL Lite is easier to implement; subset of OWL DL

C. Structure of Ontology

Ontology typically has two distinct components as shown in Figure 2. Names for important concepts and relationships in the domain

- Elephant is a concept whose members are a kind of animal
- Herbivore is a concept whose members are exactly those animals who eat only plants or parts of plants

Background knowledge/constraints on the domain

- Adult Elephants weigh at least 2,000 kg
- No individual can be both a Herbivore and a Carnivore

OWL offers additional knowledge base oriented ontology constructs and axioms.

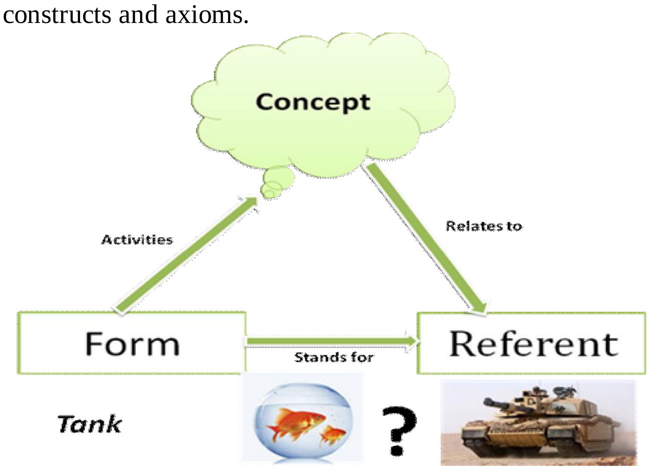

Figure 2: Ontology in linguistics

Independent contemporary efforts in DARPA Agent Markup Language (DAML) and Ontology Inference Layer (OIL) merged into DAML+OIL in 2001 and finally evolved into OWL, which was drafted in 2002 and became a W3C recommendation in 2004.

D. Advantages of Ontology

- Provide a shared understanding of domain
- Useful for the organization and navigation of web sites
- Useful for improving the accuracy of web searches
- Web searches can exploit generalization/specializati on information.

E. Friend of a Friend (FOAF)

FOAF is a machine-readable ontology describing persons, their activities and their relations to other people and objects. Anyone can use FOAF to describe him or herself. FOAF allows groups of people to describe social networks without the need for a centralized database. FOAF is a descriptive vocabulary expressed using the RDF and the OWL. The FOAF vocabulary includes classes and properties found useful to describe people online. Friendship networks connected by FOAF relationships can provide insights into features and patterns of social networks (like facebook, twitter, etc.) in the semantic web and advance the theories and models of social structures. Computers may use these FOAF profiles to find, for example, to list all people both you and a friend of yours know. This is accomplished by defining relationships between people. Each profile has a unique identifier (such as the person's e-mail addresses, a Jabber ID, or a URI of the homepage or weblog of the person), which is used when defining these relationships [23].

3.6 RULES AND RULE SYSTEMS

A rule is perhaps one of the simplest notions in computer science: it is an IF - THEN construction. If some condition (the IF part) that is checkable in some dataset holds, then the conclusion (the THEN part) is processed. Deriving somewhat from its roots in logic, rule systems use a notion of predicates that hold or not of some data object or objects. For example, the fact that two people are married might be represented with predicates as MARRIED (JENI, ASHOK). MARRIED is a predicate that can be said to hold between JENI and ASHOK. Adding the notion of variables, a rule could be something like:

IF MARRIED(?x, ?y) THEN LOVES(?x, ?y)

We would expect that for every pair of ?x and ?y (e.g. JENI and ASHOK) for which the MARRIED predicate holds, some computer system that could understand this rule would conclude that the LOVES predicate holds for that pair as well. Rules are a simple way of encoding knowledge, and are a drastic simplification of first order logic for which it is relatively easy to implement inference engines that can process the conditions and draw the right conclusions. A rule system is an implementation of a particular syntax and semantics of rules, which may extend the simple notion described above to include existential quantification, disjunction, logical conjunction, negation, functions, non monotonicity, and many other features. Rule systems have been implemented and studied since the mid-1970s and saw significant uptake in the 1980s during the height of so-called Expert Systems [17].

A. Rule Interchange Format (RIF)

RIF is part of the infrastructure for the semantic web, along with (principally) SPARQL, RDF and OWL. Although originally envisioned by many as a "rules layer" for the semantic web, in reality the design of RIF is based on the observation that there are many "rules languages" in existence, and what is needed is to exchange rules between them. RIF includes three dialects, a Core dialect which is extended into a Basic Logic Dialect (BLD) and Production Rule Dialect (PRD). The standard RIF dialects are Core, BLD and PRD. These dialects depend on an extensive list of datatypes with builtin functions and predicates on those datatypes. Relations of various RIF dialects are shown in the following Venn diagram.

Datatypes and Built-Ins (DTB) specifies a list of datatypes, built-in functions and built-in predicates expected to be supported by RIF dialects. Some of the datatypes are adapted from XML Schema Datatypes, XPath functions and RDF: PlainLiteral functions. The Core dialect comprises a common subset of most rule dialect. RIF-Core is a subset of both RIF-BLD and RIF-PRD. Framework for Logic Dialects (FLD) describes mechanisms for specifying the syntax and semantics of logic RIF dialects, including the RIF-BLD and RIF-Core, but not RIF-PRD which is not a logic-based RIF dialect. The Basic Logic Dialect (BLD) adds features to the Core dialect that are not directly available such as: logic functions, equality in the then-part and named arguments. RIF BLD corresponds to positive datalogs, that is, logic programs without functions or negations. RIF-BLD has a model-theoretic semantics.

The Production Rules Dialect (PRD) can be used to model production rules. Features that are notably in PRD but not BLD include negation and retraction of facts (thus, PRD is not monotonic). PRD rules are order dependent, hence conflict resolution strategies are needed when multiple rules can be fired. The PRD specification defines one such resolution strategy based on forward chaining reasoning. RIF-PRD has a operational semantics, whereas the condition formulas also have a model-theoretic semantics [17].

B. Semantic Web Rule Language (SWRL)

It is a proposal for a Semantic Web rules language, combining sublanguages of the OWL (OWL DL and Lite) with those of the Rule Markup Language (Unary/Binary Data log). SWRL has the full power of OWL DL, but at the price of decidability and practical implementations. Rules are of the form of an implication between an antecedent (body) and consequent (head). The intended meaning can be read as: whenever the conditions specified in the antecedent hold, then the conditions specified in the consequent must also hold [18].

3.7 DIGITAL SIGNATURE

In the field of semantics, the information is divided into triples, which can be described as phrases in which someone (the subject) says something (the predicate) about another resource (the object). The Semantic Web places great importance on concepts and less on form, for which the classic definition of the electronic signature, in which what matters is the integrity of data rather than its meaning, does not fit in with this philosophy. For this reason, Safe layer is investigating the definition of a new signature format, specifically, a semantic digital signature that better adapts to the new standards of the Future Internet.

The traditional digital signature verifies the integrity of a document, treating the information as mere strings of bytes.

For example, changing the order of two lines of a signed XML file results in the negative verification of a signature, even though there may have been no change at all in the meaning of the information this file represents. We are looking to go one step further and develop mechanisms that support verifying the integrity of the meaning of the concepts. In, for example, a semantically signed RDF file, changing the order of the triples would not affect the result of the digital signature verification. The semantic uses a less syntax. For example, the XML document is formatted for adding spaces or line breaks. So, the semantic digital signature can protect the concepts as well as the documents.

Safe layer has developed a semantic digital signature application prototype that demonstrates the validity of this concept for the use case of the FOAF ontology. FOAF supports defining concepts of personal identity in semantic format. It can also be integrated with any other vocabulary, making it possible to express information on an identity in a complete manner. For example, it is possible to define a résumé as a FOAF profile. If each of the fragments of information described in it (such as personal data, qualifications and work experience) were signed by trusted entities that guarantee them (public administrations, educational institutions and companies, respectively), the trust in the content of the document increases notable [19].

4. SEMANTIC WEB SERVICES (SWS)

The Figure 3 shows about the web service enabled with semantic web. Current Web Services are SOAP for message transport, Web Services Description Language (WSDL) for service description, and UDDI for service advertisement and discovery. These are Web Service discovery and description, No semantic (formal) description and don't support automatic web service discovery, mediation, composition into complex services and negotiation.

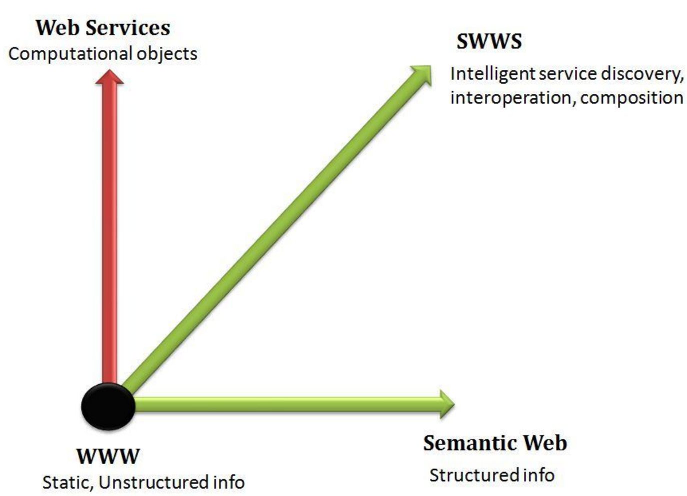

Figure 3: Semantic Web-enabled Web Services

4.1 OWL-S: SEMANTIC MARKUP FOR WEB SERVICES

The OWL-S approach proposes ontology of services motivated by the need to provide three essential types of knowledge about a Web service. OWL-S attempts to combine the representational technologies of the Semantic Web (RDF and OWL) with the dominant Web services standards, such as WSDL.

- Profile – is used to advertise the service. The service profile elements include preconditions, inputs, outputs, results and service category.
- Process model – includes inputs, outputs, preconditions, effects and the behavior of the service (data and control flow).
- Grounding – provides the needed details about transport protocols [20].

4.2 WSMO OR WEB SERVICE MODELING ONTOLOGY

WSMO is a conceptual model for four top level elements as the main concepts which have to be described in order to describe Semantic Web services. It provides an ontology based framework, which supports the deployment and interoperability of Semantic Web Services. The WSMO has four main components:

- Goals – The client's objectives when consulting a Web Service.
- Ontologies – A formal Semantic description of the information used by all other components.
- Mediators – Connectors between components with mediation facilities. Provides interoperability between different ontologies.
- Web Services – Semantic description of Web Services. May include functional (Capability) and usage (Interface) descriptions.

Descriptions of a WSMO service comprise non functional properties, a provided interface and a provided capability. Descriptions comprise of a WSMO goal comprise non functional properties, a requested interface and a requested capability. A WSMO interface describes messages sent to/by a WSMO service and the visible behavior of that service. A WSMO capability includes: non functional properties, preconditions, assumptions, post-conditions, and effects [2], [20].

4.3 SEMANTIC ANNOTATIONS FOR WSDL AND XML (SAWSDL)

SAWSDL is a set of extensions for WSDL, which provides a standard description format for Web services. WSDL uses XML as a common flexible data-exchange format and applies. XML Schema for data typing. SAWSDL extends

WSDL with pointers to semantics that are crucial for achieving automation. Adding semantics to Web services mainly aims to automate certain tasks that must be performed with services before or during invocation. Based on various efforts in SWS and service-oriented computing communities (such as OWL-S and WSMO), the generally accepted tasks are discovery, negotiation, filtering, selection, and invocation, complemented by composition and interspersed with mediation.

SAWSDL is the first step toward standardizing SWS. It forms the basis for interoperation between the various SWS efforts that previously couldn't seem to find any common ground. SAWSDL itself isn't a complete technology for allowing automation; indeed we must provide service ontology and the appropriate domain ontologies to describe Web services. The major SWS frameworks (WSMO and OWL-S) have already started to embrace SAWSDL for grounding (connecting the semantic framework to the WSDL descriptions of Web services) [21].

4.4 SIMPLE SEMANTIC WEB ARCHITECTURE AND PROTOCOL

(SSWAP pronounced "swap") SSWAP is architecture, a protocol and a platform to semantically discover and integrate heterogeneous disparate resources on the web. Unfortunately, this approach heavily relies on the provided metadata, which is usually very poor. Other approaches focus on the development of interfaces to assist in the location of web resources; for example, presents a client engine for the automatic and dynamic development of service interfaces built on top of the BioMoby standard [22].

SSWAP utilizes OWL ontologies to describe the features and capabilities of web services and standard HTTP methods to execute these web services. SSWAP aims to combine web services and semantic web technologies to enable high-throughput discovery, assessment, and integration of data and services between distributed parties. Semantic Web ontologies encoded in OWL are used to describe information about a web service such as the service category, types of input the service consumes, and the types of output the service produces. SSWAP differs from other Semantic Web Services architectures by not adopting the XML-based Web Service technologies such as WSDL and SOAP. SSWAP does not specify rules for authentication or security; but it is designed to work on protocols such as SSL and HTTPS that already address these issues.

SSWAP originated from the Semantic MOBY project which was a branch of BioMOBY project. Under the umbrella of BioMOBY, Semantic MOBY developed the fundamental model for a semantic web approach, while MOBY Services developed the web services approach commonly referred to as "BioMOBY". Semantic MOBY project was followed by The Virtual Plant Information Network (VPIN) that eventually turned into SSWAP. Evolution of syntactic and semantic web for more than the two decade is tabulated in Table 1. Different languages used for web and their respective description; merits and demerits are listed in the table since 1986.

Table 1: Evolution of Syntactic and Semantic Web

Year	Languages	Description	Merits	Demerits
1986	SGML (Standard Generalized Markup Language)	International standard (ISO 8879) for describing the structure of a document	ISO standard, platform independent	Tools are expensive
1990	HTML	HTML as an application of SGML; Is the first language used in www	Every browser supported, easy learn & use	Static, plain pages and Security features are not good
1992	HTML+	Richer version of the original HTML; A superset of HTML	This allows a gradual rollover from the previous format(HTML); Tables, Fill-out forms	Many incompatibilities and not international standard
1994	HTML 2.0	Defined by the Internet Engineering Task Force (IETF). It included ideas from the HTML and HTML+	Easy to learn, interactive forms	Standardization and deployment of the whole proposal in browsers of the time proved unwieldy
1995	HTML 3 .0	HTML 3.0 standard was proposed to the IETF	Many additional potentialities over HTML 2.0 such as tables, text flow around figures and display of complex math elements	Static and not international standard
	CSS (Cascading Style	Describing the presentation semantics (the look and formatting)	Separation of content from presentation;	Browser incompatibility

	Sheets)	of a document written in a markup language; Style web pages written in HTML and XHTML	Saving bandwidth & Page reformatting	
1996	XML 1.0 (eXtensible Markup Language)	A Meta-language is used to design other languages	Breaking the Tag Monopoly and Supports Unicode	Encourage non-relational data structure (de-normalized)
	SHOE (Simple HTML Ontology Extensions)	World-Wide Web authors to annotate their pages with ontology-based knowledge about page contents	Search is useful	Complex queries are constructed automatically
1997	HTML 3.2	This dropped the majority of the new features in HTML 3.0; instead adopted many browser-specific element types and attributes	Included tables, applets, text flow around images, subscripts and superscripts.	Not international standard
	RDF (Resource Description Framework)	RDF provided a simple but powerful triple-based representation language for Universal Resource Identifiers (URIs)	The RDF triple storage provides a standard way to share(import, export) data between different components	Some data not easy be represented in RDF; Low efficient to query data in the RDF triples, compared against RDBMS; Immature tools
	HTML 4.0	New elements are introduced, changes to attributes. Authors may provide long descriptions of tables, images, and frames	Strict, Transitional, Frameset	Need plug-ins and not international standard
1998	XSLT 1.0 (eXtensible Style Language Transformations)	Transformations to an XML document and the output can be HTML, XML, or any other structured document	Easy to merge XML data into presentation	It is difficult to implement complicate business rules in XSLT
	CSS 2	This includes a number of new capabilities like absolute, relative, and fixed positioning of elements and z-index, the concept of media types, support for aural style sheets and bidirectional text, and new font properties such as shadows	Saves time, Pages load faster, Easy maintenance, Superior styles to HTML	Lack of variables and Collapsing margins
	RDF Schema	Standard mechanism for declaring classes and (global) properties as well as defining relationships between classes and properties using RDF syntax	Defines a set of modeling primitives for structured vocabularies for machine-processable semantics of information	There are some difficulties with the semantics of RDF
1999	HTML 4.01	It offers the same three flavors as HTML 4.0	It helps to became an international standard (ISO/IEC 15445:2000)	---

	Web 2.0	Web 2.0 applications are the latest and newest trend in website designing; Millions of Internet companies and users are turning to web 2.0 for its added benefits; This version has brought about a revolution in how sites are built and applied in practical usage, thus increasing its overall functionality.	Flexibility, as far as the possibility of choosing technologies is concerned; Easier and faster access to information, When and where it is needed;	Viruses can be found; Information cannot be made private
	XPath1.0	XPath is a language for selecting parts of an XML document.	Locating nodes in a tree and performing operations over data	Function call is not possible, need XSLT 1.0 together with XPath 1.0 to achieve that result
2000	XHTML	XHTML is a separate language that began as a reformulation of HTML 4.01 using XML 1.0.	Overcomes the disadvantages of HTML; Ability to separate markup from content, strict guidelines on form and structure	Requires all elements to be closed properly, work in legacy browsers
	URI	An important principle of Web architecture is that all important resources be identifiable by URI	Linking, Bookmarking, Caching, etc	Does not provide people with a way to specify Web resources using their own alphabets
	XML Schema 1.0	It is also known as XSD (XML Schema Definition); Description of a type of XML document	Integration and accessibility; Data typing and namespaces	XML Schema is complex and hard to learn; Potential security problem
2001	Semantic Web	Meaning of Web; It makes interaction between user and machine	Machine understandable; Efficient retrieval of Information	Browser incompatibility; Semantic Web fairly unknown & only few people know how to write a Web page in RDF
	XHTML1.1	It is based on XHTML 1.0 Strict; This includes minor changes, can be customized, and is reformulated using modules from Modularization of XHTML	XHTML 1.1 represents a departure from both HTML 4 and XHTML 1.0	It is pure XML, and only intended to be XML. It cannot reliably be sent to legacy Browsers
	XQuery 1.0	XQuery 1.0 is based on XPath 2.0, allowing XQuery 1.0 to take advantage of all new XPath 2.0 features; XQuery 1.0 builds on XPath 2.0 to provide full XML Query capability	XQuery module support allows queries to be broken up into reusable fragments	---
	XSLT 2.0	XSLT 2.0 is based on XPath 2.0; User-defined functions can be defined in the XSLT language and are callable using XPath 2.0	Can write to multiple result documents in a single stylesheet execution; Supports regular expressions to analyze and separate strings	New features of XPath 2.0 with XSLT 1.0 is generally not supported
	XPath 2.0	XPath 2.0 has been improved to support the XPath 2.0 and XQuery 1.0 Data Model (XDM); It is a superset of XPath 1.0 and subset of XQuery 1.0	More expressive power than XPath 1.0 New operators, data model and functions	---
	XML 1.1	Second edition not new version of XML	Using character references to the control characters; XML 1.1 defines a set of	---

			constraints called “full normalization”	
2003	OWL (Web Ontology Language)	The Web Ontology Language (OWL) is a family of knowledge representation languages for authoring ontologies; The languages are characterized by formal semantics and RDF/XML-based serializations for the Semantic Web;OWL is written in XML	OWL is a part of the "Semantic Web Vision"; OWL comes with a larger vocabulary and stronger syntax than RDF	OWL was not designed for being read by people; Barriers to Compatibility
2004	CSS2.1	Fixes errors in CSS 2, removes poorly supported or not fully interoperable features	Better than css2; No errors	Submitted as a single document with all the Cascading Style Sheets information within it
	XML Schema 1.1	Superset of XML Schema1.0	Assertions, Open content, Conditional types and Schema-wide attributes	---
	SPARQL	Is an RDF query language, able to retrieve and manipulate data stored in RDF	SPARQL allows users to write unambiguous queries	Browsers Incompatibility
	SWRL	Rule language for Semantic Web	Provides a formally sound way of inferring information in OWL ontologies	Computational complexity
2005	eRDF	Simplified approach to semantically annotate data in HTML websites	Provides a natural way of embedding RDF in existing HTML documents	Only supports a fairly small subset of RDF
2008	HTML 5	A major attempt to standardize HTML as a Web application platform is HTML Version 5	Mutuality, Improved Semantics, Elegant forms, offline application cache, client-side database, geo-location support, consistency	HTML is not a perfect tool for designing graphic-intensive sites
	XQuery 1.1	Extended version of XQuery 1.0	Introduces new switch expression	---
	RDFa 1.0	Bridging the Human and Data Webs	Rich structured data markup for web documents	This is embedded in XHTML it generates a greater overhead than XML
2009	XHTML 2.0	XHTML 2.0 is incompatible with XHTML 1.x; More accurate to characterize as an XHTML-inspired new language than an update to XHTML 1.x	A richer, more reusable structure and easier addition of metadata; More device independence, accessibility, and semantics	Not be backwards compatible
	OWL 2	This is an ontology language for the Semantic Web with formally defined meaning; Provide classes, properties, individuals, and data values and are stored as Semantic Web documents	OWL's functional syntax closer to an RDF graph, also offering a formal equivalence to UML; The changes also allow for mapping from an RDF graph back to OWL, which was not possible before	---
	SPARQL 1.1	Update operations performed on collection of graphs in a Graphs store	Query results in XML format; Time – permitting features	Browsers Incompatibility

	CSS 3	CSS 3 is divided into several separate documents called "modules"; Each module adds new capabilities or extends features defined in CSS 2, over preserving backward compatibility	Selectors, Text Effects, Layout and Multi-Column Layout; Paged Media and Generated Content	The biggest problem is compatibility with old browsers and even new ones
2010	XQuery 3.0	It is an update version of XQuery and extension of XPath Version 3.0	It provides many new features like groupby, tumbling window & sliding window, allowing empty, try/catch	---
	XPath 3.0	XPath 3.0 is a superset of (XML Path Language (XPath) Version 1.0)	A backwards compatibility mode is provided to ensure that nearly all XPath 1.0 expressions continue to deliver the same result with XPath 3.0	---
2011	XQuery and XPath Full-Text	Powerful queries of character strings, numbers, dates and nodes are familiar to users of relational database systems	This offers a rich set of features; More comprehensive than the query language of most existing full-text systems	Syntax is a bit verbose and redundant, not very elegant;
2012	SciSPARQL	This targeted mainly at scientific computing and laboratory data management	This system automatically recognizes collections in RDF Turtle statements that represent numerical multi-dimensional arrays in order to represent them with a special native data type	Not International Standard
	RDFa 1.1	Syntax and processing rules for embedding RDF through attributes	Ability to add structured data to HTML pages directly; This is best compared to micro data and microformats	---

5. CONCLUSIONS

Thus the Semantic Web to express information in a precise, machine interpretable form. The Semantic Web made the more Understandable by Machines. Integrated information in an intelligent way and Provided Semantic based access to the Internet. It builds an appropriate infrastructure for intelligent agents to run around the web performing complex action for their users. Semantic web extracted information from texts. Also importance and need for semantic web is clearly explained and evolution of the same is listed in the table.

REFERENCES

1. John Davies. **Semantic Web**, *Next Generation Web Research*, BT, 2003.
2. Seppo Torma, Jukka Villstedt, ville Lehtinen, lan Oliver and Vesa Luukkala. **Semantic Web Services – A Survey**, 2008.
3. www.w3.org
4. Simona Elena Varlan. **Advantages of Semantic Web technologies in the knowledge based society**, 2010.
5. www.princeton.edu
6. Nigel Shadbolt, Wendy Hall and Tim Berners Lee. **The Semantic Web Revisited**, *IEEE Computer Society*, 2006.
7. www.gla.ac.uk
8. www.quora.com
9. www.ohsu.edu
10. Bhavani thuraisingam. **Security Issues for the Semantic Web**, *27th Annual International Computer Software & Applications Conference (COMPSAC'03).*
11. www.pearltrees.com
12. Fayed Ghaleb, Sameh Daoud, Ahmad Hasna, Jihad M.Alja'am Samir A. EL-soud, and Hosam El-sofany. **E-learning Model Based on Semantic Web Technology**, *International Journal of Computing & Information Sciences*, Pages 63-71.
13. Simona Elena Varlan and Cosmin Tomozei. **Simple Knowledge Organisation System**, *Proceedings of CNMI 2007*, pp.299-308, Sept.2007.

14. Li Ding, Pranam Kolari, Zhongli Ding, Sasikanth Avancha, Tim finin and Anupam Joshi. **Using ontologies in the semantic web: a survey**, 2005.
15. Joo-Seong Jeon and Gi-Jeong Lee**. Development of a Semantic Web Based Mobile Local Search System**, 2007.
16. Joh T.E.Timm and Gerald C. Gannod. **Specifying semantic web service compositions using UML and OCL**, 2007.
17. Semantic Web – Indian Student Association, UNO, URL: http://isa.unomaha.edu
18. Vipul Kashyap, Christoph Bussler and Matthew Moran. **The Semantic Web: Semantics for Data and Services on the Web**, Springer, 2008.
19. labs.safelayer.com
20. Yassin chabeb and Sami Tata. **Yet another semantic annotation for WSDL**, *IADIS International conference www/internet*, 2008.
21. Jacek Kopecky, Tomas Vitvar, Carine Bourney and Joel Farrell. **SAWSDL: Semantic Annotations for WSDL and XML schema**, *IEEE*, November 2007.
22. Maria Perez-Catalan, Rafael Berlanga, Ismael Sanz and Maria Jose Aramburu. **A semantic approach for the requirement-driven discovery of web resources in the Life Sciences**, *Springer*, April 2012.
23. Tim Finin, Li Ding, Lina Zhou and Anupam Joshi. **Social networking on the semantic web**, *The Learning Organization*, pp. 418-435, 2005.