The Concept of Semantic Web in Library Services
Saikat Goswami*, Payel Biswas**

INTRODUCTION
The World Wide Web contains huge amount of information which can be easily accessed by specifying URI (Uniform Resource Identifier) addresses, search engines and following links to find other related resources. This simplicity of usage has made the web so popular but such simplicity comes with a price. The enormous amount of data has made it increasingly difficult to find, access, present and maintain the information required by a wide variety of users. It is very easy to get lost or discover irrelevant and unrelated information because information content is presented in natural language. Search engines cannot promise precision since the indexing is based on quantity with no clue to the context in which the terms occur. For this problem, a support is essential for bringing the web to its full potential. Tim Berners-Lee, the inventor of the World Wide Web, put forward the concept of meaningful Web or semantic web. He referred to the future of the current web as the semantic web an extended web of machine-readable information and automated services that extend far beyond current capabilities.

SEMANTIC WEB CONCEPTS
The word semantic’ stands for the meaning of ’or, word web defines it as of or relating to meaning or the study of meaning. In the term semantic web also indicates that the meaning of data on the web can be discovered not just by people also by computers. According to Tim Berners-Lee (1998) the word semantic means ’machine-possible’. Tim Berners-Lee et al. (2001) describe the semantic web as: “an extension of the current web in which information is given well-defined meaning, better enabling computers and people to work in cooperation.” The key enabler of the semantic web is the need of many communities to put machine - understandable data on the web which can be shared and processed by automated tools as well as by people. Tim Berners-Lee [2] has clearly stated the main goal of the semantic web in his statement, if html and the web made all the online documents look like one huge book,
RDF; schema and inference language will make all the data in the world look like one huge database. Semantic web aims to develop such technologies that make the information more meaningful for the machines to process which in turn makes search and retrieval of information more effective. In the semantic web data itself becomes a part of the web and is able to be processed independently of application, platform, or domain. The semantic web on the other hand is about having data as well as documents on the web so that machines can process, transform, assemble and even act on the data in more useful and meaningful ways.

**IMPORTANT FEATURES OF SEMANTIC WEB**

- The semantic web comprises the standards and tools of XML (Extensible Markup Language), XML Schema, RDF (Resource Description Framework), RDF Schema and OWL (Web Ontology Language).
- Two important technologies for developing the semantic web are already in place: Extensible Markup Language (XML) and the Resource Description Framework (RDF).
- XML lets everyone create their own tags hidden labels such as `<zip code>` or `<alma mater>` that annotates web pages or sections of text on a page.
- Meaning is expressed by RDF which encodes it in sets of triples each being rather like the subject verb and object of an elementary sentence. These triples can be written using XML tags.
- In RDF, a document makes assertions that particular things (people, web pages or whatever) have properties (such as "is a sister of," "is the author of") with certain values (another person, another web page). This structure turns out to be a natural way to describe the vast majority of the data processed by machines.
- Subject and object are each identified by a Universal Resource Identifier (URI), just as used in a link on a web page. (URIs, Uniform Resource Locators are the most common type of URI.) The verbs are also identified by URIs which enables anyone to define a new concept, a new verb, just by defining a URI for it somewhere on the web.
- Subject and object are each identified by a Universal Resource Identifier (URI), just as used in a link on a web page. (URIs, Uniform Resource Locators are the most common type of URI.) The verbs are also identified by URIs which enables anyone to define a new concept, a new verb, just by defining a URI for it somewhere on the web.

**XML (Extensible Markup Language)**

The basic building block of semantic technology is the use of Extensible Markup Language (XML) which permits the use of tags to describe objects. It is a markup language much like html and was designed to carry data not to display data. XML tags are not predefined so one can define or create his/her own tags hidden labels such as or that explain web pages of sections of text on a page, but it says nothing about what the structures mean. XML is particularly suitable for sending documents across the web.

**RDF (Resource Description Framework)**

RDF is a framework for describing resources on the web. It describes a resource the resources properties and the values of those properties, often referred to as triples that consist of a subject, predicate, and object. It is particularly intended for representing metadata about the resources, such as title, author and modification date of a web page, copyright and licensing information about a web document (or information resource published on the web), or the availability schedule for some shared resources. RDF presents the syntactic structure whereby we can express information about the resource in form of XML statements. For example we want to say that the title of the webpage http://www.caluniv.ac.in is University of Calcutta Homepage. This will be expressed in RDF as: http://www.caluniv.ac.in has title whose value is University of Calcutta Homepage.

The RDF statement comprises of three parts:
- Subject
- Predicate
- Object

**Subject**

By subject we mean the web resource that has described. In the above example the subject of the RDF statement is http://www.caluniv.ac.in. In RDF the subject has to be always represented in the form of URI (Uniform Resource Identifier) which is the identifier by which the web resource can be identified in the web. URIs is of two types: URL (Uniform Resource Locator) and URN (Uniform Resource Number). URLs can locate and retrieve the resource on the web, whereas URN works to identify the web resources which may not be retrievable. For example: the ISBN of the book is 0-223-59871-7. This will be represented in URN as urn: ISBN: 0-223-59871-7. Hence the subject of the RDF statement is always presented as an URI.

**Predicate**

This denotes the specific property of the resource. For example, title is the property of the webpage. Predicate can be any other property of the resource such as creator, language, date of creation, etc.

**Object**

This is the actual value of the property. For example, in the
above case University of Calcutta Homepage is the value of the predicate (property) of the resource.\[12],[11]

**Figure 1 : An Example of an RDF Directed Graph**

**ONTOlOGY WEB LANGUAGES (OWL)**

The term ontology has its origin in philosophy, and has been applied in many different ways. Ontology is a formal representation of knowledge as a set of concepts within a domain and the relationships between those concepts. It is used to reason about the entities within that domain and may be used to describe the domain ontology as the structural frameworks for organizing information and are used in artificial intelligence, the semantic web, systems engineering, software engineering, 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 anthologies is also basic to the definition and use of an enterprise architecture framework.\[19],[10]

**Libraries and Information Science (LIS) and Computer Science (CS) Join Forces How Librarians and Computer Scientists Complement and Contrast:**

Some of the librarians work with knowledge that computer scientists are looking for when trying to envision the semantic web the two disciplines does have unique differences.

- While computer science is concerned with how software and associated machines interact with ontology, LIS is more concerned with how their users retrieve information and as a way to facilitate certain types of information-seeking behavior with the aid of taxonomies.
- As computer science professionals perceive hierarchies as logical structures that help machines make decisions, while LIS professionals view these information structures in terms of mapping out a topic for the benefit of patrons.
- Computer Science is concerned with how software and associated machines interact with ontology; in contrast, librarians are concerned with how patrons retrieve information with the aid of taxonomies. Despite these apparent differences, they are essentially different sides of the same coin.\[15],[4]

Nonetheless, there is collaboration to be made between LIS and Computer science, particularly when mapping concepts, skills and jargon between computer scientists and librarians encourages collaboration. Computer Science and LIS are working to solve problems of information retrieval and the exchange of knowledge between user groups. Whether they are doing it together or doing it separately, ontology and taxonomies will be extremely important to a number of computer scientists by facilitating the sharing and reuse of digital information.\[4],[23]

**Ontology and Taxonomies in the Semantic Web: How LIS and Computer Science need to join forces to make this happen**

LIS and cataloguing professionals are not only familiar with these concepts, as they often form the core of their work and part of the educational curricula. The traditional skills of librarianship - thesaurus construction, metadata design and information organization - are deeply important in the creation of this next stage of web development. Ontology Web Languages (OWL) facilitates greater machine interoperability of web content than that supported by XML, RDF and RDFS by providing additional vocabulary along with a formal semantics. It can be used to explicitly represent the meaning of terms in vocabularies and the relationships between those terms. OWL adds more vocabulary for describing properties and classes: among others, relations between classes (e.g., disjointness), cardinality (e.g., exactly one), equality, richer typing of properties and characteristics of properties (e.g. symmetry), and enumerated classes. Thus semantic web ontology consists of taxonomy and a set of inference rules from which machines can make logical conclusions.\[16],[21]

**SIMILARITIES BETWEEN THE LIBRARY AND THE SEMANTIC WEB**

The “Internet [Web] has been described as a library with all the books tossed on the floor”\[21] or “the Web is like a virtual library” the latter statement marshals little support when considering the full scope and anarchy of the web. The semantic web part of the larger web is however quite similar to the library for the following reasons:

- The library and the semantic web have mission statements grounded in service, information access, and knowledge discovery.
- The library and the semantic web have advanced as result of international and national standards.

**Missions grounded in service, information access and knowledge discovery**

The library's definitive goal is to support knowledge discovery for advancement of citizens and society. The semantic web's homepage provides a succinct definition of the semantic web that is characteristic of a mission statement. The semantic web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries.\[19] The statement highlights such components as a common framework, shared data (information) and collaboration and it parallels the library's standardization and sharing of bibliographic data, resource circulation and collaborative activities. The semantic web's overriding goal to imbue computer and human agents with intelligence which is very similar to the library's goal of advancing knowledge.
TO THE SEMANTIC WEB
APPLICABILITY OF LIBRARY FUNCTIONS

This section discusses the goals and objectives of the four primary functions underlying the modern library. The discussion also explores the applicability of each function to the semantic web based on the above analysis of library and semantic web similarities.

Collection development in the library

The collection development is it to build and maintain a various collection that services a designated constant patron population. The activities of collection development policy that viewed as a contract between the library and users. Collection development policies document the library's intent to grow the collection, identify collection strengths and limitations, and guide library staff, particularly bibliographers, in their collection development work. Guidelines also include selection criteria about preferred subjects and formats. Collection development policies are not permanent, rather they need to be reviewed and revised, as user populations' change and present new demands. Finally, collection development can help libraries with administrative activities by including procedures for acquisitions, gifts, weeding, replacing lost items and collection evaluation.

Semantic Web selection

The semantic web initiative, as a whole, does not identify a specific type of user, although semantic web selection policies will require review and revision for the following key reasons:

- the development of new and related projects—some of which may be competitors;
- the identification of new user agents (computer and human); and
- the development of new technologies and machine capabilities.

Based on knowledge about the library community's experience developing library collection development policies, it is likely that semantic web selection policy development will require time and patience, particularly given the absence of examples specific to the semantic web. The wide availability of library collection development guidelines and resources, such as Guidelines for Writing Collection Development Policies provide a useful framework for developing Semantic Web selection policies.

Library cataloging

The purpose of cataloging is to make library collection materials findable and discoverable so they can be used. Charles A. Cutter's (1904) objectives for a library catalog, printed in the 4th edition of his Rules for a Dictionary Catalog, are among the most influential statements impacting cataloging. Cutter's objectives state that a library should:

1. Enable a person to find a book when the author, title, or subject is known;
2. Show what the library has by author, subject, and literature genre; and

Written a century before the development of the web, Cutter's objectives are still applicable to library operations today, and thus influence current cataloging activities.

Jumping a century beyond Cutter to today, digital resource cataloging (metadata creation) is being guided by principles and objectives documented in a variety of metadata schemes. Under development is the Rules for Description and Access (RDA), which includes a draft statement of objectives (RDA, 2005). For example, descriptive data (metadata) created using RDA should enable a user to “identify the resource described” and select appropriate resources “with respect to content, format, etc.” Additional objectives address access points, representation of entities identified in Functional Requirements for Bibliographic Records (FRBR) (1998), and cataloging quality criteria (e.g. data flexibility, sufficiency, and accuracy).

Semantic Web “semantic” representation

Similarities between library cataloging and producing metadata for the semantic web, both are deal with representation. In fact, the boundary between the employ of representation standards in these two environments (libraries and the semantic web) is artificial. Rather the representation activity takes place along a continuum, with simple bibliographic representation for search and retrieval on one end, and the implementation of formal ontology and machine supported deductive reasoning on the other. Similar to the library's community extensive MARC documentation the semantic web provides comprehensive documentation for working with enabling technologies, such as XML, RDF, and OWL. However, the semantic web community falls short, currently, in providing documentation to guide the use of
metadata standards and ontology. Plans, guidelines and policies are needed stating principles and objectives for semantic web representation to ensure good quality “semantics” (e.g., coherent, consistent, accurate semantic representation). A semantic representation policy would help secure a robust framework for effective semantic web operations.\[5][10]

**Reference and outreach**

The goal of reference is to provide the library community with effective information services. Reference services include personal interaction and dissemination of information. The library has a compulsion “to provide information service to support the educational, recreational, personal and economic endeavors of the members of their respective communities” (RUSA Access to Information Committee, 2000).\[20]

An extension of reference service is outreach. Libraries plan services that are of value to their users. Whereas conducting outreach to highlight collection resources that help with finding a first professional job. Outreach extends to community outreach, generally in public through the offering of classes and other services (e.g., English as a second language classes, story time for youngsters, reader advisory services, even cooking and art classes). These items extend beyond reference but deal with overall access and use of the library facility and often promote collection use.

**Semantic Web service**

The attachment is with “reference service”— the central pillar of semantic web. The semantic web depends on standardized structured metadata and semantic web algorithms capable of reading and manipulating such data but the overriding goal is to provide service to free humans from routine tasks that computers can perform—and can perform effectively. Current semantic web services facilitate knowledge and service discovery and more sophisticated forecasted activities include automatic purchasing of an airline ticket—even an airline ticket from your preferred.\[67]

**A GAP BETWEEN SEMANTIC WEB AND LIBRARY**

The semantic web and library communities are far from being healthfully integrated. On one side of this gap, the members of the semantic web community are not fully aware of the skills, talent, and knowledge that librarians (primarily catalogers) have, and which can help advance the semantic web. This is evident by the absence of a metadata representation working group within the World Wide Web Consortium (W3C), and the severely limited participation of professional librarians on various W3C working groups. Granted, the W3C’s semantic web activity has focused more on the development of enabling technologies, rather than processes or activities. On the library side of the gap, librarians have been slow to embrace the semantic web and work with semantic web enabling technologies and standards in comparison to the way in which computer scientists, engineers, and oncologists. Now a day’s real time and instant processes like blogs, email, websites and instant massaging are placed to disseminating of information via web in library.\[9][10]

**CONCLUSION**

The library and the semantic web are cultures devoted to increasing information access and knowledge discovery, it makes sense to explore the foundations of the library and consider what primary functions may help advance the semantic web initiative. The library has been society’s chief information custodian for the last several hundred years; and if the semantic web is to evolve into a chief and trusted information network, affording services and performing tasks for both humans and machines and it need to examine the applicability of the library’s primary functions to the semantic web.

This paper explored the applicability primary library functions to the semantic web. The inquiry was a discussion based on rudimentary deduction and was supported by an analysis of various library guidelines and policies. An exploration of similarities between the library institution and the semantic web served as a base. All four of the primary library functions proved applicable to the semantic web. Each library function translates to a semantic web function.

**REFERENCES**


