BEYOND WEB 2.0: SEMANTIC WEB AND IT'S INFLUENCE ON LIBRARY SERVICES

John Paul Anbu K.[1] and Sridevi Jetty[2]

[1]Head, Periodicals, University of Swaziland Libraries, Swaziland, Southern Africa.
[2]Deputy Librarian, Central Library, Bundelkhand University, Jhansi, U.P., India.

ABSTRACT

States that Web 2.0 applications have drastically transformed the knowledge and information dissemination channels and have also changed the mindset of information seekers. While there is no doubt that the libraries struggle to cope with the changes looming the library horizon many libraries have incorporated the advancements of current innovations and provide services that skillfully make use of the current technology. Services like content enrichment service, digital reference service and the recent advancements in Online Public Access Catalogues (OPAC) prove beyond doubt that the information scientists are working overtime to deliver the advancements of the web. With the current web nearing its run time the future of the information services at the dawn of the newer web technologies seem to be in the quandary. The concept of semantic web is clearly seen as the next major advancement in the web technologies. Examines the various advancements achieved through the web 2.0 technologies especially in the libraries and information centers and goes further to analyse the challenges that lie ahead and the opportunities it has created for the information professionals at the wake of semantic web technologies.

KEYWORDS: Library Services, Information Services, Web 2.0, Semantic Web

Introduction

The transformation of libraries and information centers commenced with the advent of microprocessors into the library services. From the humble beginning of transferring piles and piles of catalogue cards into the computers, till the latest distributed content enrichment services which integrate a number of enhancements for library users with the invigorating colorful mosaic of services, the library and its services have grown a long way. The unprecedented developments in the field of ICT and the rapid advancements in the World Wide Web and telecommunication technologies have greatly redefined the concept of information, especially the way information is produced, accessed and disseminated. With the current available tools and techniques and the variety of information sources, it is possible to access a wide variety of information on myriad access technologies. The Information landscape is filled with newer and efficient tools and technologies to dig the information anywhere from the entire digital ecosystem. While information explosion has produced a deluge of information snippets, it has also redefined the mindset of users. The users are no longer the passive information seekers instead they have become "information confirmers and information consolidators (Kataria,

2009)." Current information technologies and its various Services allow users to identify, seek, confirm and consolidate the information they look for. In this process we are once again at the crossroads where the existing technologies are reaching their zenith and newer technologies have begun to warm up the information horizon.

Web 2.0 and Library Services

Web 2.0 is the term coined by Dale Dougherty of O'Reilly Media to distinguish between the old and the new generations of websites, but it had no clear definition when it was coined. Razmerita observes that "in the last few years we have witnessed a transformation of the web from a static web towards a 'living web' where the users bring content, collaborate and share knowledge (Razmerita, 2009)." This transformation of web from static to dynamic is termed as Web 2.0. Minsk observes that "in recent years, new software design patterns and business models are observed on the Web which is commonly referred to as Web 2.0 (Minsk, 2007)." While some people look at web 2.0 as a marketing ploy born out of the ashes of the dot.com collapse, some agree that it made significant enhancement to the original web technologies. Though Tim Berners-Lee, the father of internet, reasserts that there is no such thing as web 2.0 since the original web was perceived to do exactly what the current web technologies do, there is no doubt that the evolution of web and its services are very evident. The binding link between the webs of its origin to the current one is the level of interactivity. The result of this interactivity is the emergence of services like Instant Messaging, Steaming Media, Blogs, News Feeds, Social Bookmarking Tools, Tagging and AJAX, which are grouped together called as Web 2.0 services.   Using these tools "people no longer passively consume information; rather, they actively contribute, even customize tools and technologies for their use (Razmerita, 2009)." These revolutionary ideas are new to libraries yet they are very important for the consolidation of information as the users are now able to identify, seek, confirm and consolidate the kind of information which they look for and a number of library services skillfully use this advantage.

The history of Information services show that the early digital library services have used the technology only as an extension of its existing manual operations. The manual card catalogues were converted into electronic catalogues and in turn the Online Public Access Catalogues (OPACs) manifested those electronic transformations. Later consuming activities of the library operations like circulation, order processing, and information dissemination were slowly integrated in the library system which gave birth to the Integrated Library Systems (ILS). When the web based catalogues and web based digital content delivery modes were initiated it provided room and scope for many exciting new services. Many exciting Web 2.0 services started capturing the imagination of information professionals. Services such as Synchronous Communication (Instant Messaging) gave way for Instant Referencing services. Content Delivery (RSS Feeds, News Feeds) and Streaming Media made the streaming content and syndicated content available on the library catalogues on the fly. Collaborative Publishing Tools (Blogs, Wikis, Tagging etc.,) added new dimension to information resource for the libraries. Social Networks with the interactive approach provided ample scope for exciting information dissemination and content delivery to library users. All these newer services have slowly become part of library services.

A. Enhanced Reference Services

The first place where the Web 2.0 services impact in the Library and Information services is in the reference section. The Synchronous Communication tools like Instant Messaging

(IM) have brought the entire reference library into the fingertips of the users. Many libraries now offer the "Ask a Librarian" or "Just Answers" services which connect the user and the library through the digital interactive link. There are different kinds of digital referencing systems available in libraries. Most common in the libraries are the instant messaging communication systems which link the reference librarian to answer the queries of the users. On the other hand the commercial services make use of the current advancements to provide either subscription based reference referrals or fixed reference answers for set reference questions. The advancement of the web has ensured that these services are offered at a much higher level which allows multiple Instant Messaging clients to be integrated simultaneously using widgets which can be embedded on a library webpage to seamlessly interact with the library as well as with buddies on a same platform. Meebome (Meebo me, 2009) is one such service which is an open source Instant Messaging (IM) tool which allows anyone to use multiple IM clients in the same platform. This service also provides the Meebome widget which can be embedded into the library website for reference referrals not only between the library authorities and the client but also from the buddies.

B. Content Enrichment Services

The different content delivery modes like RSS Feeds, News Feeds, and Streaming Media have enriched the content of library catalogues to such an extent that the information services can enhance their catalogues with rich content. Most of the libraries which use the streaming media or the syndicated feeds try to take them directly into their catalogue to provide the users with up to date content and the user's choice of streaming audio or video with a single mouse click. The evolution of multimedia and dynamic content into the library catalogues have prompted commercial subscription services to enhance the information sources with "on the fly" or "on demand" enrichments.

Notable among these enrichments is the content enrichment service available through Syndetics (Syndetics Solution, 2010) which allow the users to check cover images, biography profiles, series information, awards, summaries, author notes, first chapter, excerpts, table of contents and book reviews. These module based services can be scaled and integrated according to the needs of the libraries.

C. Next Generation Library Catalogues

The next generation of library catalogues seems to be heading towards an assortment of all the web 2.0 tools and technologies along with a number of enhancements which envy a web portal. With the library catalogues providing dynamic content enrichment from the backbone, the front end of the catalogues offer a number of other utilities. While some catalogues provide a portal interface with different channels running side by side with news feeds and reference rooms adorning the front page along with the search screen, some catalogues provide a number of interactive services. Products like Aquabrowser (Aquabrowser, 2010) provide faceted search capabilities along with relationship searching and cloud tag related term supplementation for users to choose from.

The next generation catalogues are more poised towards personalized search experience where each user will be able to personalize the way they view the catalogue, organize their registries in folders and assign tags, configure the searches that they perform regularly and subscribe to RSS feeds to get any news on the topic. While the catalogues will allow interoperability and the syndication of contents it will also allow external services, for example with bibliographical reference managers, options to purchase books, etc. The evolving catalogues will be equipped with options of predefining syndicated channels to disseminate the user's contents, news, topics, and authors. While Content Enrichment and Information Architecture will dominate the traditional information portals encompassing the table of contents, the index and the summary, the displayed information could be filtered through facets and groups with

possibilities of the users able to retrieve and re-engage the same search for future reference.

Beyond Web 2.0: The Semantic Web

While there is no doubt that the current web technologies have revolutionized the power of computing, it is also true that the current web has completed its own course of time and the dawn of new web technologies are on the horizon. Sure et al. points out the problem of web 2.0 as the use of the current Mark-up language. "The major shortcoming of HTML is that it is well suited for human consumption, but not for machine-process ability (Sure, 2005)." The experts think the next wave of web, web 3.0, is going to change this element of human interaction and make internet to connect with information. In the process the next generation of web will look like a Personal Assistant who knows virtually everything about the user and can access all the information from the internet for them to answer any question they have. While the experts are divided on how the current web can recreate itself to such a massive change, most of them feel that this kind of new web will co-exist with the current web while some feel that the new web will replace the original web.

The idea of the perceived new technology is to allow the user to relax while the internet to do all the work for them. Most experts pin their hope for this new technology through the use of semantics. Warren and Alsmeyer observe that "the basic building block of Semantic technology is the use of eXtensible Markup Language (XML) which permits the use of tags to describe objects (Warren, 2005)." The semantic web technology represents a new way of formatting data. Unlike the current web where documents are linked to other documents, web 3.0, the semantic web, will have data linked to data through the use of semantics using web ontology language. Burke feels that "The semantic web converts web pages from being readable and displayable by computers to being understandable by computers. It does this by adding extra metadata to web pages and by sharing this metadata between multiple applications. It enables computers to understand a web page in the way a human does (Burke, 2009)."

The main ingredient of the future web technology will be the ability to find, interpret and relate the data to the context of the search. So, basically, a search by two different persons on the same topic should yield different results. This will be possible since browsers will start to understand more about the content of what the user is browsing and the context within which it has been searched and will start making recommendations based on the user's unique "Internet profile" based on the user's browsing history. Applications like FOAF (Friend of a Friend) which are currently used to describe people online within the social networking circles without the need of the centralized database and newer browser applications which can seamlessly integrate the browsing history and the user profile can analyze each search according to the requirements of the user. Semantic search engines will further increase the accuracy of the results as they will attempt to give smarter results by first searching for concepts and making the results more meaningful and interactive for people who want to further their search.

Current web technologies make use of Application Programming Interfaces (APIs) to allow developers to create applications which can manipulate the resources. This manipulation of data will be the fundamental of the next generation of web where the data will play an integral part using web ontology. For the next generation of web is to be very

effective a comprehensive and detailed ontology is to be created. The concept of mesh up which is prevalent in web 2.0 will be effectively used to combine two or more applications to have an effective search sequence.

In short the next generation of web, the semantic Web, will be able to understand information on the web. The semantic web will be fundamentally about using new technology which can help to remix, reuse and repurpose data on the web in new ways. It is first and foremost about the emergence of a data web where the entire web will be a giant interconnected database which will have simulated intelligence through which it can understand and respond.

Semantic Web and Libraries

Against this backdrop, it is imperative to think what lies for libraries and Information scientists in terms of the anticipated technology and its services. Macgregor points out that "for many digital libraries or cultural institutions, the semantic web offers an opportunity to better expose the valuable digital resources (Macgregor, 2008)." This will be achieved through the anticipated better individualized search capabilities of future browsers which will use the unique internet profiles and FOAF information and at the same time use better ontology descriptions used by these valuable resources. He further observes that the "Semantic Web technologies are capable of enhancing digital libraries or repositories by facilitating improved navigation and retrieval within heterogeneous document environments, user profiling, personalization and contextualization, improved user interface and human computer interaction (Macgregor, 2008)." While pondering on the future, it is equally prudent to brainstorm the initiatives in the Information Science horizon to see how far the capabilities of the evolving technologies can be stretched.

A. Knowledge Organization

The information professionals and semantic web practitioners have much in common as both of them delve on the organization of knowledge. The semantic web implies on seamlessly combining data using the ontologies. Currently information professionals make use of classification schemes, taxonomies, thesauri or subject headings to strengthen the metadata. Through the use of Web Ontology Language (OWL) or through the usage of Resource Description Framework (RDF) these metadata can be further strengthened since they are based on the concept that every data item should have a unique web identifier called a URI (Uniform Resource Identifier) and that every data item can be connected to every other item. Organizing the universe of knowledge using the ontology based language or resource based language will certainly make information much more reliable and valuable at the time of retrieval. The "Simple Knowledge Organization Scheme" (Simple Knowledge Organisation Systems (SKOS)., 2011) model is a valuable initiation and exploration into the Knowledge Organization using web ontology. SKOS provides specifications and standards to support the use of knowledge organization systems such as thesauri, classification schemes, subject heading systems and taxonomies within the framework of the semantic web. It provides a standard way to represent knowledge organization system using Resource Description framework (RDF).

B. Collaborative Referencing

The power of Semantic Web can be utilized much better with initiatives of collaborative referencing for subjects using powerful networks or consortia on special subjects and interests. Projects like Knowledge Exchange and Capacity Building Division (KCE) (Knowledge Exchanges and Capacity Building Division (KCE)., 2011) of FAO provides

comprehensive standards and norms and a collaborative technical support for Information Management. The AGROVOC Concept server Workbench (ACSW) again initiated by FAO is a web service java tool for collaborative building and structuring multilingual ontology and terminology systems in the area of agriculture which uses powerful referencing platform which makes use of collaborative terms, subject headings, definitions and relationships in their documents (Yongyuth, 2008). These initiatives will not only enhance the knowledge organization but it will also enable the distributed and collaborative management of information which in turn will facilitate the reuse of existing information and extension of the same information into different facets.

C. Recommender System

Currently a number of Integrated Library Systems (ILS) use recommender systems within the periscope of the knowledge facets. It is a simple system using the traditional knowledge classification. But the future web technologies, especially the personalized browser environment and the Friend of Friend (FOAF) technologies, promise a much more robust recommender lists while giving the results. Vascoda's (Vascoda, 2009) "search term recommender system" is another step towards creating a system which can even suggest the search term before the user engages in the search process. These kinds of recommender systems can not only supplement the user's needs, they can also bridge the gap between the levels of user searches. The Telplus project funded by eContentplus (Isaac, 2010) program is a good beginning towards building such recommender system where it uses the automatic vocabulary alignment method for semantic searches. This system can suggest the suitable search terms using intelligent agents. Telplus harps on how mapping between different vocabularies can improve multilingual access to heterogeneous collections. They have experimented with two different alignment schemes; Lexical alignment and Extensional alignment and have evaluated them for further research and development.

D. Personalized Library Catalogues

With more and more users desiring to obtain personalized services in search results, the future of library catalogues is pointed towards providing such personalized catalogues. Ferran asserts that "two elements determine the functionalities of the desired personalization system: first the user's profile, including navigational history and user preferences; and second, the information collected from the navigational behavior of the digital library users (Ferran, 2005)." The future catalogues point towards creating personalized library catalogues using the above mentioned concepts. Ferran presents a model which uses sub-ontologies where they can be effectively used to create personalized library catalogue where an ontology for describing the way users browse can be kept as a parallel criteria along with the ontology for describing the content in the library. Such catalogues will greatly complement the recommender systems which can be an integral part of the personalized library catalogues.

Conclusion

With the current web reaching its zenith and the enormous amount of information churned out of the web nearing its pinnacle, the information community needs to realign their reservoirs to develop strategies in line with the future of the libraries and information centres. While there is no doubt that the enormous potential of web 2.0 is rightly channeled to the various services of the library labyrinth, it remains to be seen how the library community will react to the enormous challenges lying ahead especially with the

introduction of newer technologies. Semantic web is still in its infancy and making this web to work the way it is envisaged remains the prime focus of information specialists. Library and Information professionals should be reoriented to these developments so that the power of the web can be fully harnessed.

REFERENCES

Aquabrowser. (2011). Retrieved on December 10, 2010 from http://www.serialsolutions.com/aquabrowser/

Burke, M. (2009). The semantic web and the digital library. *Aslib Proceedings: New Information Perspectives, 61*(3), 316-322.

Ferran, N. (2005). Towards personalization in digital libraries through ontologies. *Library Management, 26*(4/5), 206-217.

Isaac, A. (2010). TELplus automatic subject alignment experiments. Retrieved on February 7, 2011 from http://www.theeuropeanlibrary.org/portal/organisation/ cooperation/telplus/documents/TELplus_D3.5_04012010.pdf

Kataria, S. (2009). Applications of Web 2.0 in the enhancement of services and resources in academic libraries: An experiment @JIIT University Noida, India. In Globalizing Academic Libraries: Vision 2020 (pp. 583-589). New Delhi: Mittal.

Knowledge Exchanges and Capacity Building Division (KCE). (2011). Retrieved May 10, 2011 from http://www.fao.org/kce/about_kce/en/

Macgregor, G. (2008). Introduction to a special issue on digital libraries and the semantic web: Context, applications and research. *Library Review, 57*(3), 173-177.

Meebo me. (2009). Retrieved on January 28, 2010 from http://www.meebome.com/

Minsk, C. (2007). Web 2.0.: Concepts and technologies for dynamic B2B integration. Emerging Technologies & Factory Automation (ETFA). IEEE Conference, 315-321.

Razmerita, L. (2009). Personal knowledge management: The role of web 2.0 tools for managing knowledge at individual and organizational levels. *Online Information Review 33*(6), 1021-1039.

Simple Knowledge Organisation Systems (SKOS). (2011). Retrieved on November 12, 2010 from http://www.w3.org/2004/02/skos

Sure, Y. & Studer, R. (2005). Semantic web technologies for digital libraries. *Library Management, 26*(4/5), 190-195.

Syndetics Solution. (2010). Retrieved on November 18, 2010 from http://www.bowker.com/syndetics/

Vascoda. (2009). Retrieved on November 11, 2010 from http://www.vascoda.de/

Warren, P. & Alsmeyer, D. (2005). Applying semantic technology to a digital library: A case study. *Library Management, 26*(4/5), 196-205.

Yongyuth, P. (2008). The AGROVOC concept server workbench: A collaborative tool for managing multilingual knowledge. Retrieved on December 10, 2010 from ftp://ftp.fao.org/docrep/fao/010/ ai168e/ai168e00.pdf