

THE 8th UITMCJ ACADEMIC CONFERENCE

Globalisation and Dynamics Towards a World Class University

CD PROCEEDING ACCON 08

Penyelenggara:

Rohani binti Jangga Ainol Hasanal bin Jalaluddin Dr. Carolyn Soo Kum Yoke Noridah binti Sain

Terbitan:

BPJIA UiTM Cawangan Johor

2012

Cetakan Pertama 2012

Hak Cipta Terpelihara. Tidak dibenarkan mengeluar ulang mana-mana bahagian artikel, ilustrasi, dan isi kandungan buku ini dalam apa jua bentuk dan dengan cara apa jua sama ada secara elektronik, fotokopi, mekanik, rakaman, atau cara lain sebelum mendapat izin bertulis daripada BPJIA UiTM Johor.

Perpustakaan Negara Malaysia

Data Pengkatalogan-dalam-Penerbitan

CD Proceeding for ACCON 08

8th Academic Conference, UiTM Johor

1. Proceeding

ISBN 978-967-0479-03-3

Diterbitkan oleh: Bahagian Penyelidikan Jaringan Industri dan Alumni UiTM Cawangan Johor KM 12 Jalan Muar 85000 Segamat JOHOR DT

Dicetak oleh:

Pereka grafik: Unit Korporat dan Media UiTM Johor

Percetakan Segamat

Jalan Orkid 85000 Segamat JOHOR DT

CONTENT

CLICK ON THE TITLE TO LINK TO THE FULL PAPER.

*A=Accouting; B= Information Management, Marketing & Management; C=Economy & Finance; D= FSKM; E=Other Discipline (CITU, APB, LAW)

NOS.	REF.*	TITLE	AUTHOR
1.	E103	Penglibatan Syarikat Swasta Di Malaysia, Singapura Dan Indonesia Dalam Mengembangkan Potensi Harta Wakaf: Suatu Perbandingan	Mohd Ali Muhamad Don
2.	E104	Acquisition Of Verb-Tenses Among Esl Lecturers: An MDAB Pilot Study	 Dr Carolyn Soo Kum Yoke Nor Haniza Hasan
3.	A105	Tax Knowledge And Tax Compliance Among Self-Employed In Malaysia	1. Fatimah Alwi 2. Noraizah Abu Bakar
4.	B106	Knowledge Management Initiatives For Sirim Berhad	 Siti Nuur-Ila Mat Kamal Zainab Sulaiman
5.	B108	Managing Public Transport Service Operation In Reducing Travel Fear Factor	 Rohana Sham Mashita Sham Khairunnisa Rahman Nurazlina Samsudin
6.	C109	Employees' Perspective Of Ethics	 Rabiatul Alawiyah Zainal Abidin Raziah Bi Mohamed Sadique Dr. Norhayati Alias Pn Noraizah Abu Bakar
7.	E110	Using Online Corrective Feedback In The Editing Stage Of Academic Writing Among ESL undergraduates	 Cecilia Bai Rajendran Dr. Carolyn Soo Kum Yoke Noridah Sain Puteri Nur Hidayah Kamaludin Sofwah Md Nawi Suhaili Mohd Yusof

8.	B113	Cointegration And Causality Between Overnight Policy Rate And Commercial Bank Rate	 Abdul Razak Bin Jambari Azizah Binti Daut Noranita Abdamia
9.	C115	Energy Consumption, Gross Domestic Product, Foreign Direct Investment And Co2 Emission In Malaysia	 Mohd Azim Sardan Azman Ali Tan Yan Ling Abdul Razak Jambari
10.	C118	Systematic Risk Of Malaysian Stock	 Tay Bee Hoong Tan Yan Ling Nur'asyiqin Ramdhan Zulkifli Mohamed
11.	B119	The Impact Of Semantic Web On Digital Libraries Development	 Muhammad Asyraf Bin Wahi Anuar Isma Bin Ishak
12.	E120	<u>الفِعْل) Penggunaan Kata Kerja Kini (المُضَارِع</u> المُضَارِع <u>Masa</u>	Abd Rahman Jamaan
13.	B128	Pusat Sumber Sekolah Sebagai Ejen Kecemerlangan Sijil Pelajaran Malaysia (Spm): Kajian Ke Atas Pelajar Baru Sesi Disember-April 2010, UiTM , Johor	 Nor Diana Bte Abd. Rahman Nor Ezan Bin Omar Zailani Bin Shafie
14.	E129	<u>Pengurusan Zakat Di IPTA :</u> Pengalaman Dan Cabaran di UiTM	 Nazrudin Bin Hashim Mohd Zainoddin Bin Mustaffa
15.	E130	Kajian Ke Atas Penggunaan Partikel Mod UiTM al Pelajar Johor Dalam Pembelajaran Bahasa Mandarin	Chong Peng Hwa
16.	B135	<u>A Case Study Of Digi</u> <u>Telecommunications Sdn Bhd</u> <u>Strategic Marketing Analysis</u>	 Rohanizan Md Lazan Roslina Ali Shaherah Abd Malik Rabiatul Alawiyah Zainal Abidin
17.	B138	Entrepreneurial Strategic Alliances And Partnership Of Marketing And Retailing In Small And Medium Enterprises In Agro Sapling Supply Chain	 Shaherah Abd. Malik Wan Haslin Aziah Syaidatul Zarina Mat Din Roha Mohamed Noah Norhayati Omar

18.	D139	<u>Student Timetable Asistant System</u> (<u>Stasys)</u>	 Noor Azrin Binti Zainuddin Nafisah Binti Jamingan @ Amin Mohamad Azlan Shah Bin Mohamed Azman
19.	E141	Using Readers Theatre To Improve Reading Fluency And Comprehension	 Evelyn Sharminnie Vasuthavan Dr Carolyn Soo Kum Yoke S Kunaratnam Sita Raman Nor Haniza Hasan Cecilia Bai Rajendran Noridah Sain
20.	A142	How Quantitative And Qualitative Of Data Research Can Be Embedded Together?	 Dr. Norhayati Alias Pn Noraizah Abu Bakar
21.	B143	Preliminary Study On Social Network Among Malay Smes Entrepreneurs: Gender Differences	 Oswald Timothy Edward Ehsan Fansuree Surin
22.	E144	Penambahbaikan Pembelajaran Di Sekolah Agama Kerajaan Negeri Johor: Kajian Pelaksanaan Sistem Anjal	 Prof. Madya Ahmad Nawawi Yaakob Nor Rashidah Mohamed
23.	E137	Representasi Semantik Lewat Adat- Istiadat Perkahwinan Dalam Alam Minangkabau	Ainol Hasanal Bin Jalaluddin

The Impact of Semantic Web on Digital Libraries Development

Muhammad Asyraf bin Wahi Anuar Isma bin Ishak

Lecturer, Faculty of Information Management Universiti Teknologi MARA

Abstract

Semantic web technologies enable the computer to interpret data in a machine-interpretable form on the website that can facilitate the searching process. Currently, only human can understand the information and the search engine only provide facility to do online searching but in semantic web environment, it possible for the web to understand and satisfy the requests of people and machines to use the web content. This article discusses the current state of the Semantic Web, and how it may impact on the digital libraries development over the next few years by understanding the semantic web standards, for example, metadata and taxonomies. The use of ontologies allows interoperability among applications exchanging machine-understand the user needs. With the help of authority vocabularies, ontologies, and natural language processing, the semantic web can create a higher level of search retrieval giving well defined meaning by linking data and reusing the same data to connect across many different documents on the web.

Keyword: semantic web; digital library; metadata; controlled vocabulary

Introduction

The initiation of the Semantic Web by World Wide Web Consortium (W3C) has been active for the last few years, especially in information retrieval process. Sir Tim Berners Lee, the founder of WWW, was the person responsible for introducing the concept of Semantic Web where web contents are flexible, integrated and self learning to provide a more interactive experience for users. Metadata is one of the foundations of Semantic Web in information retrieval and it has been used in libraries for centuries. Dublin Core is one of the metadata used by libraries that can provide a set of vocabulary terms which can then be used to describe resources for the purpose of discovery, either in printed or electronic format. Most websites also use the tree structure classification scheme to coordinate the information on the same or related subjects, leading to the hierarchical organization of catalogs. The approach of catalog and tree structure classification has been widely accepted and implemented in website environment, especially for online directories like Yahoo! Directory, Malaysian Manufacturers Directory or the Open Directory Project.

Most online search engines retrieve and arrange information based on keywords and the popularity of the websites but this does not guarantee that the information retrieved is valid and up to date. The effort of the Semantic Web initiative to continue to learn and explore the metadata provides additional method of managing online resources. Semantically annotated contents of the digital library's database can be retrieved by new properties. These properties interlink different resources and provide new meaningful information to the existing full text and bibliographic descriptions. This paper will discuss how metadata usage in Semantic Web will impact the development of the digital library.

Semantic web definition

According to Heidari (2009), the Semantic Web is an evolving extension of the World Wide Web in which the semantics of information and services on the web is defined, making it possible for the web to understand and satisfy the requests of the people and machines that use the web contents. It is derived from the World Wide Web Consortium director, Sir Tim Berners-Lee's, vision of the Web as a universal medium for data, information, and knowledge exchange. Semantic Web is able to describe things in a way that computers can understand. Decker et. al. (2000) claimed that Semantic Web will enable intelligent services such as information brokers, search agents, and information filters, which offers greater functionality and interoperability than current stand-alone services.

Digital library

The advent of technology has not only revolutionized the teaching process but also redefined the students' methods of learning and retrieving information. As for the information management students, they are well trained to be the information professional as they are proficient to do selection on reliable information. There are many definitions of "digital library." Terms such as "electronic library" and "virtual library" are often used synonymously with it. The elements that have been identified as being common for digital library term, as identified by The Association of Research Libraries (1995), are:

- the digital library is not a single entity;
- the digital library requires technology to link resources;
- the linkages between many digital libraries and information services are transparent to the end users;
- universal access to digital libraries and information services is the primary goal; and
- digital library collections are not limited to document surrogates

The revolution of technology has led the world to rely on online resources instead of printed material as they are easy to retrieve and are accessible from anywhere. Libraries also prefer digital collections for many reasons, including, but not limited to, the following: digital journals can be linked from and to indexing and abstracting databases; access can be from the user's home, office, or dormitory whether or not the physical library is open; the library can get usage statistics that are not available for print collections; and digital collections save space and are relatively easy to maintain (Tenopir, 2003).

Digital libraries are large, organized collections of information objects. Well-designed digital library software has the potential to enable non-specialist people to conceive, assemble, build, and disseminate new information collections. This has great social impact because it democratizes the dissemination of information. In particular, it will revolutionize the way in which education is conducted and educational materials are prepared (Witten, Bainbridge & Nichols, 2006). Digital libraries containing representations of original works provide opportunities for students and scholars to conduct research from personal computers and workstations located in their homes, offices and laboratories. Facilitating remote research from private and public locations offers the learners opportunities that are unconstrained either by the limited resources found in particular physical libraries or by their schedules of operation (Downs & Friedman, 1999). Digital libraries can be viewed as infrastructures for supporting the creation of

information sources, facilitating the movement of information across global networks, and allowing the effective and efficient interaction among knowledge producers, librarians, and information and knowledge seekers (Adam, Atluri, & Adiwijaya, 2000).

Digital library interaction shares experiences from two distinct communities, the Human Computer Interaction (HCI) and the information science community. The HCI community is carrying the expertise on the improvement of user interaction with a new information management medium, while the information science community adds the scent of domain knowledge in the sense of information behavior. This conjunction imposes the investigation of the iterative exchange of dialogue elements between the user and a digital library system, which are translated through an interface and aims to fulfill the user's informational needs (Tsakonas, Kapidakis, & Papatheodorou, 2007). Campbell and Fast (2004), consider the potential of Semantic Web for future catalogues in academic research libraries. They recognize that interoperable transfer of metadata in metadata harvesting programs is important, for example, the Open Archives Initiative Protocol for Metadata Harvesting.

The Metadata and RDF

The RDF (Klyne and Carroll, 2004) is a simple XML-based language to define computerunderstandable vocabularies that people and programs can use to describe things of interest such as web sites, newspaper articles, e-mail messages, people, books, events, or web services. RDF mimics human languages in that it allows one to introduce new terms (individuals, classes and properties) that are defined (partially, at least) in context with existing terms. RDF Schema (Hayes, 2004) extends RDF by providing vocabulary to build logical object-oriented schema, including a simple typing system, subclasses, sup-properties, inheritance, etc. RDF, a core part of the Semantic Web, uses uniform resource identifiers (URIs) to uniquely identify resources and provides a framework to describe a resource in terms of its properties and its relationship with other resources. RDF uses a graph of nodes and arcs representing the resources, their properties and values (Yan Han, 2006).

Currently the metadata of library materials are contributed by the librarian responsible for creating the catalog record for all library collection, based on the bibliographic descriptor such as ISBD, AACR, MARC21 or Dublin Core in a machine readable format. In order to ensure all information has a standard description and makes it easy for retrieval, subject heading is used for standard vocabulary control like LCSH and SLSH. By publishing controlled vocabularies in one place, which can then be accessed by all users across the Web, library catalogs can use the same Web-accessible vocabularies for cataloguing, marking up items with the most relevant terms for the domain of interest and search engines can use the same vocabularies in their search to ensure that the most relevant items of information are returned. The Semantic Web offers open standard Resource Description Framework (RDF) with useful flexibility for describing information and resources. The implementation of RDF has led to the emergence of web technology to become more intelligent, whereby the computer are able to understand the relationship between things. For example, computers can understand that the keyboard is a part of computer and computer is related to technology and internet or any other things that is related and relevant. This shows that the Semantic Web use the metadata to understand the user's needs instead of only providing the result based on the user's keywords.

The Semantic Web will bring structure to the meaningful content of Web pages, creating an environment where software agents roaming from page to page can readily carry out sophisticated tasks for users (Berners-Lee, Hendler, and Lassila, 2001). The introduction and implementation of RDF in Semantic Web environment enables the computer to learn more in understanding the content and identifying the most relevant object to be published in the search result. The web already acts as the information source and the issue of information overload has led to the questionable result retrieval as the result usually will be based on the hits or ranking of the website. RDF has some very important roles in controlling and providing the most relevant result by establishing a relationship between resources that are interrelated.

Impact areas

Semantic Web and digital library are different terms but will it produce a semantic digital library? The ability of RDF in Semantic Web to aggregate internet resources has led to the new concept of information sharing and retrieval process as it can provide interoperability with other systems either at metadata or communication level or both and not limited to the digital library system only. The integrated information based on the metadata like resources, user profiles, bookmarks, social network and others, provide highly meaningful information to be published and used.

The Semantic Web enhances the capabilities of search engines which forms familiar parts of the current Web so that they can become useful information management tools. The Web is already an information source of choice for many learners and researchers. A more structured and directed approach to managing this information space, both within institutions and across the whole community, can make this information more useful, with less wasted effort, and more capacity to measure the quality of information. Both Semantic Web and digital library gain advantages in their own development as for the digital library, they started to emerge into a huge online entity as the Semantic Web with the RDF technology enables the computer to understand better about the user's needs and provides internet resources in a collective manner. RDF allows the digital library to use the online search engine facility as an added value to the OPAC.

	Digital Library	Semantic Digital Library
Storage	Database & archive	Database & archive
Metadata	Digital bibliographic description	Semantic bibliographic description
	(MARC, ISBD, AACR)	(interconnected metadata) + provided by users
Interface	Full text search (OPAC)	Search and browsing on ontologies (Web)
Pros	content accessible online	• User contribute to the classification process
	Faster searching process	• User enhance digital content using blog,
		wiki etc.
		• User understand community driven
		annotation
Cons	• Need to know the right keyword	• How to convince the legacy metadata
	Lonely user	provider to the value of social annotation

•	Problem interconnecting different	
	libraries	

Table 1.0 Comparison of digital library and the semantic digital library environment.

The integration of RDF and library metadata will produce a comprehensive result and enhance information sharing. The success of social networks has already changed the internet users' attitude in information sharing. As we can see, the information sharing culture started to develop when people realized that the social network is one of the popular methods of sharing something because there will always people who read about it instead of updating it into their own websites or blogs. Semantic digital library also will lead to the same situation as it can capture any related information associated with any library material and make the collection more comprehensive. Users also could share any of the library material through a website since the integration of library metadata and RDF provides full capabilities in information sharing and distribution. Once the sharing process happened, the information will get new community annotations that can be used to improve the internet resources retrieval.

Benefits of semantic digital library

Technically, semantic digital library involve 3 major components which are content, user and the system. In addition, each of these components is interrelated in term of performance, usability and usefulness. In the context of content development, the semantic digital library offers manageable and highly organized information. The rapid growth of information may lead to information overload and if there is no proper information management, it might cause a waste of resources because the information or data is not connected to each other. The semantic digital library is able to manage complex information storage and searching activity by having a flexible resource management like dynamic views on database instead of static view in a single system database.

The collaboration between the library materials with blogs, wiki and forum will bring the semantic digital library to its full potential as the users will contribute in the content annotation process. On the other hand, the users also have the opportunity to share their knowledge within a community and have a proper platform to exchange opinion and ideas. However, information sharing could be harmful if there is no proper monitoring and evaluation process especially when two different levels of users, such as a typical end user and an expert user, contribute to the annotation. Annotation plays important roles as it determines the relationship between the information and makes the semantic technology able to retrieve it. Moreover, the digital library also gains advantages from the semantic because it will be able to identify and select information from complex resources.

Content is the prime reason why people interact with a digital library. The more advance the technology involved with the internet, the more people will contribute since technology made sharing easier and this situation will lead to the huge number of information produced every day and it will become more complex to be retrieved. The concept of semantic digital library is to makes the users be involved in the content annotation process, and allow the users or readers to share their knowledge within a community. Currently the concept of a library is concentrated on content or information delivering but not on knowledge sharing within a community of users. At the same time, the semantic digital library provides better communication between users in and across communities. Once the collaboration happen,

it means that the information will be disseminated by itself and there will be more contributors who will share their information and knowledge related to that particular topic.



Fig 1.0 Architecture of semantic digital libraries based on Information Triptych Model

The user characteristics are complex and constantly evolving time by time as there is no specific mechanism to control users' need and to satisfy them. In terms of the digital library, it is impossible to provide detail overview on the library material such as books, journal and others. The librarians need time to give feedback to the users, since they need time to seek information and review. In the semantic digital library concept, it can provide opportunities for the users to retrieve any content from different sources and not only depend on the digital library collection. This is because the users can join the community, share the same interest and have a discussion related to that particular topic or issues. This is not only limited to the semantic digital library community discussion facility. Users can retrieve information from external services like social networks or other related websites that share the same interest. This shows that the content and users are now interrelated to each other as there is no limitation on content and user discussion. Moreover, it will encourage the knowledge sharing attitude to so that it becomes a culture. Furthermore, this type of knowledge sharing will give benefits to everyone as it has no limitation and everyone will be able to contribute.

The integration between other websites and systems provide a huge opportunity in information sharing and at the same time, encourage the community contribution to the digital information environment. The most important connection to the semantic digital library is the social network since the social network has now become a new trend in web usage. Traditionally, the main communication via the internet is e-mail, instant messaging and forum but the social network has become a very popular medium in communication and community integration. This trend significantly changes the way humans communicate and how internet users are connected to each other. Today, people use social networks not only to virtually develop the connection to each other, but also share the latest information through it as the medium is very popular and has become a trend. The semantic digital library adds to the advantage by integrating the library service with social networking like the open access and page sharing via the RSS and tags. By having these functions, the library can develop a knowledge society and improve the function of the library in daily application.

Conclusion

The Semantic Web provides exciting potential for the extension of library web services but with a high overhead of time and resources for the creation and maintenance of metadata and taxonomies. The way semantic digital library capture the annotation from the users is very useful since the user will contribute more than just a librarian but on the other hand, the issues of controlling and the reliability of the annotation can be questionable if there is no proper citation or references. Another major challenge is about the metadata itself. Currently each library has its own database to store the collection description and not share with others. If there is no proper sharing mechanism to make it possible for the library to open the database for other search engine to do the searching and indexing on their database, the semantic digital library initiative will be difficult to be fully implemented and utilized. This is because the concept of the Semantic Web is to share and capture every resource based on the description contributed by the internet user, and if the library database is not able to provide sharing ability, it will limit the retrieving process

However, the librarians are still important and relevant but they need to be very skillful especially in information seeking process. The task of the librarian is not only to provide and assist users to retrieve information but in the era of semantic digital library, they need to be competent in the latest technology and be an expert in manipulating the searching process. Moreover, a librarian needs to understand the complexity of information resources and should be very skilful in developing and implementing metadata standards and controlled vocabulary. Another challenge is about the privacy of data, where in semantic environment everything is shared and everyone contributes. The privacy might be the main concern as most of the content is copyright protected and some security mechanism should be applied in order to protect the information from misuse. The semantic digital library still needs further improvement in terms of its user functionality, usefulness and the performance of the information system in capturing the right information with the proper technique.

References

Association of Research Libraries (1995). Definition and Purpose of a Digital Library , *Association of Research Libraries*, Washington, DC available at: http://www.arl.org/resources/pubs/mmproceedings/126mmappen2.shtml (accessed March 17, 2012).

Adam, Nabil R., Atluri, V., & Adiwijaya, I. (2000). SI in digital libraries. *Communications of the ACM*, 43(6), 64-72.

Berners-Lee, T., Hendler, J. and Lassila, O. (2001). The Semantic Web. *Scientific American*, May 2001, p.

29-37.

Campbell, D.G. and Fast, K.V. (2004). Academic libraries and the semantic web: what the future may hold for research-supporting library catalogues. *Journal of Academic Librarianship*, 3(5), pp. 382-90.

Decker, S. et. al. (2000). The semantic web: the roles of XML and RDF. IEEE Internet Computing, 15, 63-

74. Available at http://www.cs.vu.nl/~frankh/postscript/IEEE-IC00.pdf (accessed March 17, 2012).

- Downs, Robert R., & Friedman, Edward A. (1999). Digital library support for scholarly research. *Information Processing and Management*, 35, 281-291.
- Hayes, P. (2004). Rdf semantics. *w3c recommendation*, available at: www.w3.org/TR/2004/REC-rdf-mt-20040210/ (accessed: June 10, 2012).
- Heidari, K. (2009). The impact of semantic web on e-commerce. *World Academy of Science*, 51, 303-306. Available at http://www.waset.org/journals/waset/v51/v51-53.pdf (accessed March 17, 2012).

Tenopir, C., Hitchcock, B. & Pillow, A. (2003). Use and Users of Electronic Library Resources: An Overview

and Analysis of Recent Research Studies. *Council on Library and Information Resources*, Washington, DC available at: www.clir.org/pubs/reports/pub120/pub120.pdf (accessed March 17, 2012).

Klyne, G. and Carroll, J.J. (2004). Resource description framework (RDF): concepts and abstract syntax. Available at: http://www.w3.org/TR/2004/REC-rdf-concepts-20040210 (accessed June 10, 2012).

Tsakonas, G., Kapidakis, S. & Papatheodorou, C. (2007). Evaluation of user interaction in digital libraries.

International Journal on Digital Libraries, 8(1), 21-38. Available at http://www.springerlink.com/index/w2w0j4h272k26812.pdf

Witten, Ian H, Bainbridge, D. & Nichols, David M. (2006). Digital libraries education: specialized trainning

course. *UNESCO*, Moscow available at: http://iite.unesco.org/pics/publications/en/files/3214563.pdf (accessed March 17, 2012).

Yan Han, (2006). A RDF-based digital library system. *Library Hi Tech*, 24(2), 234 - 240. Available at: http://dx.doi.org/10.1108/07378830610669600 (accessed: June 10, 2012).