Volume : 6 | Issue : 4 | April - 2017 ISSN - 2250-1991 | IF : 5.761 | IC Value : 79.96

Original Research Paper Library

Understanding Semantic Web Technologies in Digital Libraries

Payal Prajapati (PhD scholar, Dept. of Library and Information Science, Rai University, Ahmedabad)

M. K. Prajapati 16-A, Nandanvan Township, Nr Avasar Party Plot, Modhera Road, Mehsana-384002

ABSTRACT

In recent years more and more information has been made available on the Web. High quality information is often stored in dedicated databases of digital libraries, which are on their way to become expanding islands of well-organized information. However, managing this information still poses challenges. The Semantic Web provides technologies that are about help to meet these challenges. This paper gives an overview of recent Semantic Web technologies which can be used to enhance digital libraries in semantic nature. Deploying Semantic Web technologies would lead to efficient and more precise representation of digital library content and hence better retrieval. This paper explores the Semantic Web and the technologies that support its functioning, including XML, RDF, and Ontology. Semantic Web technologies are considered from the perspective of digital Libraries and how Semantic Web technologies can enhance the functioning of digital libraries. The paper additionally explores on some existing semantic digital library systems like DSpace, SMILE, JeromeDL and BRICKS.

KEYWORDS digital libraries, semantic web, web 3.0, ontology

Introduction

The machines cannot read, understand the webpages like human beings, and perform the work without the human involvement. Thus need has arisen for the current web, which can enable the machines to find, combine and act upon the information by interpreting the data. If the information sources are semantically structured, this enables the machines to understand and respond to complex human requests based on their meaning. The Metadata tags provide a method by which computers can categorize the content on the web pages. In digital libraries semantic techniques are often deployed to reduce and catching the search. The expensive manual overhead for indexing documents, maintaining metadata. Digital Libraries provide a vast amount of digitized information ranging from collections of cultural heritage to specialized topic centered portals.

One of the essential differences between digital libraries and unstructured collections such as the Web is the focus on information quality. In contrast typical Web search engines base their indexing on text-based measures from information retrieval and structural properties of the collection, e.g. link analysis, whereas digital libraries usually use indexes (manually) crafted from document metadata. Since metadata can express concepts not explicitly occurring in the document, (or leave out concepts explicitly mentioned, but not relevant for the document) the use of a metadata index generally leads to better precision and recall in information services. However, given the exponential increase in newly published items even for focused collections, librarians face two serious problems. First it is increasingly costly and time consuming to properly index new items (leading to a delay in actually offering the item to customers); second in an ideal collection, the indexing has to foresee all possible (future) uses for a specific item. Moreover, the information overload for the user and the increasing specialization of (research) interests force indexes to be more and more specific in the choice of appropriate indexing terms. In fact, the vision of today's digital libraries is to provide personalized information spaces for each individual customer. To this end, semantic technologies have been recently proposed to bring a higher rate of automation into the indexing process and changed the way of web searching.

What is the Semantic Web?

The Semantic Web provides a common framework that allows data to be shared and reused across applications, enterprise, and community boundaries. It is a collaborative effort led by W3C with participation from a large number of researchers and industrial partners. Its objective is to convert all the unstructured documents on the web into a web data. It is based on the Resource Description Framework (RDF). –Wikipedia.

According to the W3C, "The Semantic Web provides a common framework that allows data to be shared and reused across application, enterprise, and community boundaries." The term was coined by Tim Berners-Lee for a web of data that can be processed by machines.

Semantic Web Framework

MCF: Meta Content Framework

The Meta Content Framework (MCF) is a tool to provide information about information. The primary goal is to make the Web (Internet or Intranet) more like a library and less like a messy heap of books on the floor. In order to understand MCF, there are three things that you'll need to learn. 1. Objects, Categories, and Properties - the conceptual building 2. The XML syntax in which MCF is stored. 3.0 The Directed Linked Graph mathematical model which lies behind MCF, which can be used by computer programmers to build efficient MCF implementations building blocks. Meta Content Framework (MCF) was a specification of a content format for structuring metadata about web sites and other data. MCF was developed by Ramanathan V. Guha at Apple Computer between 1995 and 1997. One application utilizing MCF was HotSauce, also developed by Guha while at Apple. It generated a 3D visualization of a web site's table of contents, based on MCF description.

RDF

RDF is a standard model for data interchange on the Web. RDF has features that facilitate data merging even if the underlying schemas differ, and it specifically supports the evolution of schemas over time without requiring all the data consumers to be changed.

RDF extends the linking structure of the Web to use URIs to name the relationship between things as well as the two ends of the link (this is usually referred to as a "triple"). Using this simple model, it allows structured and semi-structured data to be mixed, exposed, and shared across different applications.

This linking structure forms a directed, labeled graph, where the edges represent the named link between two resources,

represented by the graph nodes. This graph view is the easiest possible mental model for RDF and is often used in easy-to-understand visual explanations.

Ontology
Ontologies define the concepts and relationships used to describe and represent an area of knowledge. Ontologies are used to classify the terms used in a particular application, characterize possible relationships, and define possible constraints on using those relationships. In practice, ontologies can be very complex (with several thousands of terms) or very simple (describing one or two concepts only).

An example for the role of ontologies or rules on the Semantic Web is to help data integration when, for example, ambiguities may exist on the terms used in the different data sets, or when a bit of extra knowledge may lead to the discovery of new relationships.

OWL
Ontology Web Language is a language which allows us to describe the semantics of classes and properties, add more vocabulary in the domains of internet. Web ontologies provide richer integration and interoperability of data; the applications developed using OWL are intelligent, work at the level of human conceptual level, and searches across diverse communities and integrate the information.

SEMANTIC SEARCH
Semantic search is when a whole lot of resources are used in order to perform a search, rather than just keywords. In other words, semantic search uses a wide variety of information to give you search results. In short, the purpose of semantic search is to go beyond the 'static' dictionary meaning of a word or phrase to understand the intent of a searcher's query within a specific context. By learning from past results and creating links between entities, a search engine might then be able to deduce the answer to a searcher's query, rather than provide ten blue links that may or may not provide the correct answer. Major web search engines like Google and Bing incorporate some elements of semantic search.

SEMANTIC DIGITAL LIBRARIES
PRE-DIGITAL LIBRARIES
Computers have made revolutionary changes in every field of life; undoubtedly, the field of library and information has been no different. Importantly, conventional libraries moved to the concept of digital libraries, which ultimately made gaining knowledge more efficient and organized. However, a notable important fact here is that the digital library should stand for more than a well-organized centralized form of information.

Furthermore, they should also embody the essence of communication, which was originally the aspect of face-to-face interaction between the people at a conventional library.

POST DIGITAL LIBRARY OR SEMENTIC DIGITAL LIBRARIES
The advent of digital libraries in our lives, another innovative step followed. This step was made in relation to making the search more meaningful and intelligent. Essentially, it growth of Web 2.0 has given way to new methods of accessing information and contributing opinions. Notably, semantic digital libraries enable the user to get the intended information concerning an object without the presence of the exact word in the search. This integrated form of information is based on different metadata which provides a more meaningful data. These libraries tend to provide a better and more convenient form of browsing interfaces.

Applications of RDF and Digital Libraries (Examples)
Friend of a Friend (FOAF) - FOAF is an application of RDF, a machine readable ontology which describes people, the links between them and other people. It provides a rich vocabulary to describe personal information.

Simple Knowledge Organization System (SKOS) - SKOS is an application of RDF which provides a data model for Knowledge Organization System (KOS) being developed by W3C. SKOS enables data to be linked or merged by semantic web applications with other controlled vocabularies or subject indexes where complex data integration is required. It is used to improve the recall, retrieval precision and provides a number of searching methods for users.

EXISTING SEMANTIC DIGITAL LIBRARY SYSTEMS
JeromeDL - Can be considered as a social semantic digital library. Based on Semantic Web as well as social networking in order to promote collaborative activities along with other common uses of semantic digital libraries.

DSpace (IR) - Semantic Search -The first semantic search operability for DSpace developed at HPCLab, University of Patras.

SIMILE3 - (Semantic Interoperability of Metadata and Information in unlike Environments) this system focuses on enhancing the integration aspect of metadata, services etc. to increase accessibility.

BRICKS - This system focuses on the basic infrastructure of a digital library network so that information can be shared amongst users in the cultural heritage domain

Conclusion
A semantically intelligent Integrated Library System will provide effective functioning in the provision of library services. The Librarians who are having full skills, talent and knowledge will become the advocates of Semantic Web and they can fulfill the vision and mission of the Semantic Web. Ultimately, the Transcendent Web (Web 3.0 - Semantic Web) will depend on a high level of artificial intelligence underlying many Web processes. Using inputs from different sources, including browsing history, user-specified preference, and contextual information such as location, these systems will profile users to better understand both the content and the context of their requests in digital library spear.

REFERENCES

1. Alotaibi, Sara Jeza: Semantic Web Technologies for Digital Libraries: From Libraries to Social Semantic Digital Libraries (SSDL), Over Semantic Digital Libraries (SDL); Southampton / United Kingdom
2. http://en.wikipedia.org/wiki/Semantic_Web
3. https://www.w3.org/RDF/
4. Kruk, Sebastian Ryszard and Decker, Stefan and Zieborak, Lech: JeromeDL - Adding Semantic Web Technologies to Digital Libraries; 2005
5. Lagoze, C., Krafft, D., B., Payette, S., Jesuroga, S., 2005, What Is a Digital Library Anymore, Anyway? D-Lib Magazine November 2005, Volume 11 Number 11, http://dlib.org/dlib/november05/lagoze/11lagoze.html
6. Borgman, C., L., 2000, Digital libraries and the continuum of scholarly communication, Journal of Documentation, 56 (4), pp. 412-430
7. Borgman, C., L., 1999, what are digital libraries? Competing visions, Information Processing & Management, pp. 227-243,
8. Celino, I., Turati, A., Della Valle, E., and Cerizza, D. (2006). Squiggle Med: Semantic search for medical digital library, Technical report, CEFRIEL.
9. https://www.crazyegg.com/blog/everything-about-semantic-search/
10. tmu.ac.in/college-of-computing-sciences-and-it/wp.../sites/.../ICAC-1604178.pdf