Resource Description Framework Schemas for E-Library

Jamal F. Tawfeq
PhD Computer Science
University College Madenat Al-Elem
drjamal2004@yahoo.com

Saja Majeed Mohammed
MSc. Computer Science
Baghdad University
saja.mohammed82@yahoo.com

Abstract

Libraries can play a major role in the dissemination of knowledge as a huge repository of knowledge. This research shows the description for the content of the electronic library using metadata and Resource Description Framework (RDF) to facilitate access to content at any time based on this description, thus providing a more meaningful search. In this paper provides an integration of computer science techniques with electronic libraries, and determination of semantic features of this content by proposing a new scheme representative in the Resource Description Framework (RDF) and modification of the basic elements in RDF Dublin core.

Keywords: Digital-library, Semantic Web, Metadata, RDF, RDFS.

وصف محتوى المكتبات الالكترونية باستخدام البيانات الوسطية

سجى مجيد محمد[1]، جمال فاضل توفيق[2]

1.   قسم علوم الحاسبات، جامعة بغداد
2.   قسم علوم الحاسبات، كلية مدينة العلم الجامعة.

الخلاصة

يمكن للمكتبات أن تلعب دورا كبيرا في نشر المعرفة كونها مستودعا ضخما من المعرفة. يوضح هذه البحث وصفا لمحتوى المكتبة الإلكترونية باستخدام البيانات الوسطية وإطار وصف الموارد لتسهيل الوصول للمحتوى في أي وقت باستخدام هذا التعريف، وبالتالي توفير بحث أكثر جدوى. ويناقش هذا البحث كيف يمكن دمج تقنيات علوم الحاسبات مع المكتبات الإلكترونية، وكيف يمكن تحديد السمات الدلالية لهذا المحتوى من خلال اقتراح مخطط جديد ممثل في إطار وصف الموارد و تعديل العناصر الأساسية في دبلن آر دي أف.

1. Introduction

   The World Wide Web is the biggest sources of information, including documentation and information sources and variety that can be accessed through traditional search engines. There is a number of difficulties for finding exact ones supposed to be shown to somebody. The metadata is the solution of the problem of resource discovery on the Web.

Libraries are the main source of information for the users, therefore the digital library can be defined as " a focused collection of digital objects that can include text, visual material, audio material, video material, stored as electronic media formats (as opposed to print, microform, or other media), along with means for organizing, storing, and retrieving the files and media contained in the library collection".[1]
The benefits of digital libraries are the following:

 - Easy to bring the information to the end user, unlike the traditional library.
 - Provides infinite storage with low cost.
 - No need to physical space or boundary.
 - Available at any time.
 - Provide multiple and universal accesses.
 - Simplest to maintenance.

1.2 What is The Semantic web?

   In Web the information should be more machine understandable, So The Semantic Web can 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". [2]
Also, the semantic web technology provides a data which can be reused in many applications; it is implemented by the Resource Description Framework (RDF).

1.3 Semantic Digital Libraries

   The use of semantic technology in a digital library can improve the precision of search engine, enhance the architecture, expanding in the data repository and archives, the discovery of information to make the users can easily find the information . The "Semantic search encapsulates a concept-based, semantic matching and ontology-centric framework for providing extant search results".[3]

Also, " The benefit of the semantic approach based on ontologies is that it gives us a framework for searching and browsing information objects on the web and gives more relevance and accuracy to search processes. ".[4]
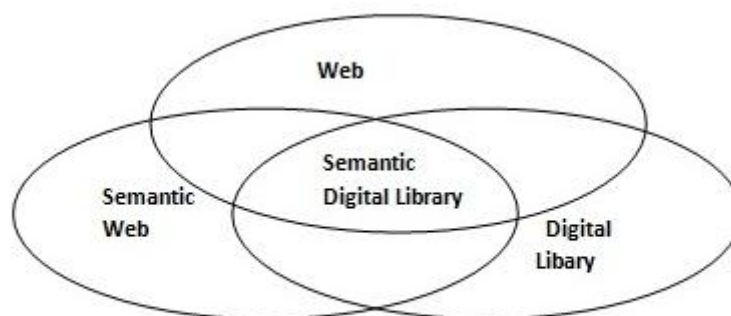
Figure 1.The process of representing a semantic digital library

2. Metadata Concept

Metadata is Structured Document about document, objects, photos, and books, when properly implemented metadata can unambiguously describe information resource, enhancing information retrieval and enabling accurate matches to be done.

Metadata are used to describe data, information and any resource. it can describe just about any thing you find on a computer. Metadata is structured information that describes, explains, locates or otherwise makes it easier to retrieve, use or manage information resource[2]. In this paper the information resource is Books.

T. B. Lee give a well-formed definition to metadata as [3][4]:

2.1. Definition

Metadata is machine understandable information about web resource or other things.

Books metadata is structured information about book. Such as titel, auther , edition, publishing, and so on.

Throughout the search in this paper, it was founding that, there are three principles for book metadata:

1. Metadata is essential to identify and track the book.
2. Ownership metadata must never be removed.
3. Metadata must be written in formats that are understood by all.

If all application writes metadata using open standards that are fully consistent and interoperable, then information will be able to move seamlessly between different system and environments."Metadata is structured information that describes, explains, locates, or otherwise makes it easier to retrieve, use, or manage an information resource. Metadata is often called data about data or information about information". [5]

2.2. Dublin Core Metadata Element Set

" The Dublin Core is meant to be sufficiently simple to be understood and used by the wide range of authors and casual publishers who contribute information to the internet. Dublin Core elements have become widely used in documenting Internet resources". [6]

*   Title
*   Author or Creator
*   Subject and Keywords
*   Description
*   Publisher

- Other Contributor
- Date
- Resource Type
- Format
- Resource Identifier
- Source
- Language
- Relation
- Coverage
- RightsManagement

3. RDF and RDF Schemas

RDF (Resource Description Framework) is the primary technology for the semantic web implementation which is "developed by W3C are standard web ontology languages for implementing the Semantic Web". [7]

The RDF is represent and exchange the information on the web, RDF is composed from triple structure (resource , property , value) which represent a statements about resource.

Assume this example of statements:

"http://www.example.org/index.html has a creator whose value is John Smith

could be represented by an RDF statement having:

- resource is a (subject) http://www.example.org/index.html.
- property is a (predicate) creator.
- and a value is an (object) is John Smith.

RDF statementscan be represented as nodes and arcs in a graph:

- A nodes for the subject and the object.

    An arc for the predicate, directed from the subject node to the object node".[6]

Figure 2. represent the RDF statement above in a graph

The RDF/XML to this graph is:

```
<rdf:RDF>
    <rdf:Description    rdf:about= http://www.example.org/index.html >
        <creator>John Smith</creator>
    </rdf:Description>
</rdf:RDF>
```

The RDF has vagueness . To clear this limitation, the RDF Schema is required.The RDF schema is a Worldwide language developed by W3C (World Wide Web Consortium) for making ontology in RDF and for describing the the metadata structure to illustrate the WWW resources by using personal vocabularies .

4. RDF Library Schemas

These schemas are:

a. Modify Dublin Core Matadata elements

The Dublin Core Metadata element set for the digital library bibliographic search information is modified as below:

- **Title**: Title of the book.
- **Description**: Summary of the content of the book.
- **Author**: An entity associated with the creation of the book.
- **Publisher**: An entity responsible for making the book available.
- **Date:** A point or period of time associated with the publishe of the book.
- **Language**: The native language of the book.
- **Rights:** Copyright of the book.
- **Format:** The data format(e.g., pdf , word) of the resource.
- **Subject:** The topic of the content of the book. It could be one keyword or more to describe the book.
- **Type:** Field of this resource.
- **Keywords:** a keyword means any word that you wish to find in the book.

Below is a modify Dublin Core Schema, and we used an Arabic Vocabulary.

```
<rdf:RDF
Xmlns="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#"
Xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#">
<rdf:Property rdf:ID="title">
<label  xml:lang="en">title</label>
<SubPropertyOf
        rdf:resource="http://pur1.org/dc/elements/1.1/title" />
</rdf:Property>

<rdf:Property rdf:ID="creator">
<label  xml:lang="en">Author</label>
<SubPropertyOf
        rdf:resource="http://pur1.org/dc/elements/1.1/creator" />
</rdf:Property>

<rdf:Property rdf:ID="subject">
<label  xml:lang="en">Subject</label>
<SubPropertyOf
        rdf:resource="http://pur1.org/dc/elements/1.1/ subject " />
</rdf:Property>

<rdf:Property rdf:ID="description">
<label  xml:lang="en">Description</label>
<SubProbertyOf
        rdf:resource="http://pur1.org/dc/elements/1.1/ description " />
</rdf:Property>

<rdf:Property rdf:ID="publisher">
<label  xml:lang="en">Publisher</label>
```

```
<subPropertyOf
        rdf:resource="http://pur1.org/dc/elements/1.1/ publisher " />
</rdf:Property>

<rdf:Property rdf:ID="date">
<label  xml:lang="en">Date</label>
<subPropertyOf
        rdf:resource="http://purl.org/dc/elements/1.1/date" />
</rdf:Property>

<rdf:Property rdf:ID="type">
<label  xml:lang="en"> type</label>
<subPropertyOf
        rdf:resource="http://purl.org/dc/elements/1.1/type" />
</rdf:Property>

<rdf:Property rdf:ID="format">
<label  xml:lang="en">Format</label>
<subPropertyOf
        rdf:resource="http://purl.org/dc/elements/1.1/format" />
</rdf:Property>

<rdf:Property rdf:ID="rights">
<label  xml:lang="en">Rights</label>
<subPropertyOf
        rdf:resource="http://purl.org/dc/elements/1.1/rights" />
</rdf:Property>

<rdf:Property rdf:ID="language">
<label  xml:lang="en">Language</label>
<subPropertyOf
        rdf:resource="http://purl.org/dc/elements/1.1/language" />
</rdf:Property>

<rdf:Property rdf:ID="keywords">
<label  xml:lang="en">Keywords</label>
<subPropertyOf
        rdf:resource="http://purl.org/dc/elements/1.1/keywords" />
</rdf:Property>

</rdf:RDF>
```

b. Book Schema

   The book schema contains the keyword use in the "Subject" property of Dublin Core schema. That property should contain as many of the following keywords as are applicable. It makes a specific statement about a piece of media content.

The key words have the following meaning:

- **Edition**: The edition number of the book.
- **Place**: The geographic Place of publication.
- **Published-Date:** A date associated with the availability of the book.
- **Notes**: Additional information about the book.
- **Institution**: Institution that contributed the book.
- **Abstract:** a brief statement that presents the main points in a concise form.
- **ISBN:** International Standard Book Number.
- **ISSN:** International Standard Serial Number.
- **URL:** If the resource exists online at holding institution, provide the URL.

```
<rdf:RDF
xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns"
xmlns="http://www.w3.org/2000/01/rdf-schema#"
xmlns:book="">
<Class rdf:ID="keywords">
<comment xml:lang="en">An enumeration of keywords to describe the subject of
book.</comment>
</Class>
<content:keywords rdf:ID=" Edition">
    <label xml:lang="en"> edition </label>
</content:keywords>
<content:keywords rdf:ID=" Place ">
   <label xml:lang="en">place</label>
</content:keywords>
<content:keywords rdf:ID=" Published-Date ">
   <label xml:lang="en"> published-date </label>
</content:keywords>
<content:keywords rdf:ID="Note">
   <label xml:lang="en">note</label>
</content:keywords>
<content:keywords rdf:ID=" Institution ">
   <label xml:lang="en">institution </label>
</content:keywords>
<content:keywords rdf:ID=" Abstract ">
   <label xml:lang="en"> abstract </label>
</content:keywords>
<content:keywords rdf:ID=" ISBN ">
   <label xml:lang="en"> ISBN </label>
</content:keywords>
<content:keywords rdf:ID=" ISSN ">
   <label xml:lang="en"> ISSN </label>
</content:keywords>
<content:keywords rdf:ID=" URL ">
   <label xml:lang="en"> URL </label>
</content:keywords>
```

5. Example

This example of the metadata in RDF format that is explains the book:

```
<? xml version='1.0' encoding='ISO-8859-1'?>
```

```
<rdf:RDF xmlns:RDF="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
xmlns:rdfs="http://www.w3.org/TR/1999/PR-rdf-schema-19990303#"
xmlns:pd0= "http://www.w3.org/2000/LibraryRDF/dc-1-0#"
xmlns:pd1= http://www.w3.org/2000/LibraryRDF/technical-1-0#>
<rdf:Description rdf:about="">
   <pd0:Title> Programming Abstractions in C++</pd0:Title>
   <pd0:Creator> Eric S. Roberts</pd0:Creator>
   <pd0:Subject> Programming Abstractions in C++</pd0:Subject>
   <pd0:Description> Bestselling Programming Tutorial and Reference Completely Rewritten for
         the New C++11 Standard.</pd0:Description>
   <pd0:Publisher>Stanford University </pd0:Publisher>
   <pd1:Editor>Stanford University </pd1:Editor>
   <pd0:Publication-Date> 2012</pd0:Publication-Date>
   <pd1:Place > Stanford</pd1:Place>
   <pd1:URL>http://StanfordUnivercity.org </pd1:URL>
   <pd0:Copyrighte> Stanford Univercity </pd0:Copyrighte>
   <pd1:ISBN>001245689 </pd1:ISBN>
   <pd0:Keywords>C++ </pd0:Keywords>
   <pd0:Keywords>Programming</pd0:Keywords>
   <pd0:Keywords>object orented programming </pd0:Keywords>
   <pd0:Keywords>oop </pd0:Keywords>
   <pd0:Format>PDF</pd0:Format>
   <pd0:Language>en</pd0:Language>
 </RDF:Description>
</RDF:RDF>
```

7. Conclusion

In our days people can be refer to the semantic web as the next generation of the WWW. To make an improve to the data aggregation and in order to get precise information from the available information to us. A libraries can play a big role in publishing the knowledge which it make as a huge repository of knowledge,So to provide a more meaningful search in digital libraries we modify and extend the Dublin Core standared Metadata about the bibliographic digital library search.

8. References

1- **Witten**, Ian H., Bainbridge, David Nichols. (2014) **"How to Build a Digital Library "**, ISBN 978-0123748577.

2- **Carrie** Moran, **"The Use of Semantic Web Technologies in Digital Libraries"**, December 5.

3- **Preedip** Balaji B1, Amit Kumar Sarangi1, Devika P. Madalli1. (2012) Knowledge Organization Systems for Semantic Digital Libraries, Documentation Research an Training Centre, Indian Statistical Institute, Bangalore, 10-13.

4- **Bendik** Bygstad, Gheorghita Ghinea, Geir-Tore Klaebo. **Organizational Challenges of The Semantic Web in Digital Libraries**.

5- **National** Information Standards Organization. (2004) Understanding Metadata**.** Niso Press, URL www.niso.org.

6- **Frank** Manola andEric Miller. (2004) RDF Primer, (W3C) Semantic Web Activity Lead, 10-2-2004. URL: http://www.W3C.org/TR/2004/REC-rdf-primer-20040210/.

7- **Jaehoon** Kim, Kangsoo Jung, Seog Park. (2008) An RDF Triple based Acces Authorization Specification Considering Subsumption Relationships, Department of Computer Science, Sogang University, ICDEW 2008. IEEE 24th International Conference.