

CATALOGING WEBSITES FOR A LIBRARY ONLINE CATALOGUE

Wendy Tan

*Head of Cataloging
Hunter College Library
The City University of New York
New York, U.S.A.*

Abstract

Cataloging Websites for a library online catalogue is a difficult task for catalogers, but this activity is getting to be a trend in order to expand library services. However, with the changing nature of Web materials and the complexity of cataloging them, library administrators have to know the upsides and downsides of adding Websites to OPAC and do some experiments before making the decision to go with the flow. This article lists a few advantages and disadvantages of this effect for discussions, and also talks about metadata and some tools of cataloging Internet resources.

Keywords :

Internet resources cataloging; Metadata; MACHine Readable Cataloging; Dublin core; CORC

Introduction

The focus of this article is to explore the management of freely available Websites on an OPAC environment from a cataloger's perspective. While cataloging Websites has been a growing activity, thus far there are very few articles devoted to this topic.

An OCLC study shows that there were only 100 Websites in 1993 when the World Wide Web started to gain ground in the information world, but this number went up to 7.4 million in year 2000. Undoubtedly there is a huge amount of valuable resources in the Internet world.

but unfortunately they are hidden in a chaotic jungle! How can they be organized for accessibility and utilization? As Nancy Olson pointed out in *Cataloging Internet Resources : A manual and practical guide* :

Using existing library technique and procedures and creating records for retrieval through existing online catalogs is the most efficient method of accessing these resource.

Though, as Ms Olson perceived, librarians do possess skills and experiences required to take on this challenge, library administrators have to think over whether it is beneficial to patrons and libraries to use OPAC as a gateway for access to Web resources very carefully before making a commitment to take this approach.

Before I have this issue further examined, I think a short introduction of metadata will be a useful icebreaker because understanding metadata seems a prerequisite to doing cataloging of Internet resources.

Metadata

The word “metadata” has become increasingly noticeable in many journal articles and documents related to data storage and retrieval ever since Internet resources have made rapid influx to the information highway and subsequently revolutionized the method of information delivery. However, the meaning of this buzzword may still not be clear to most information seekers.

In brief, metadata is structured data about data. They are pieces of reference that describe an information resource such as the author and title of a particular book; the responsibility and date of publication; and what this book is about, etc. As a matter of fact, the existence of metadata predated the advent of the Web. Therefore, it is very true that, without knowing it, people have been using metadata long before it was stored in an electronic format.

Types of Metadata

To name a few, library catalog cards; MARC21 (MARC); Dublin Core (DC); Text Encoding Initiative (TEI); Encoded Archival Description (EAD); Government Information Locator Service (GILS) etc. all fit in the realm of metadata. Among them, MARC and DC, besides catalog cards which have rendered their glorious service and are bidding farewell to the library world, are attracting more attention in the information community, so I will explain ;—to the best of my knowledge ;—the basics needed to understand how these two types of metadata interweave with information resources.

MARC

The acronym MARC stands for Machine-Readable Cataloging. This means a computer can read the information shown on a traditional library catalog card. The availability of MARC standard made it possible for shared cataloging at an international level in the past and also set the tone for subsequent automation of library catalogue.

MARC format was originally developed by the Library of Congress as US MARC in 1960s, and through years there were many revisions by MARC Committees. In 1987, the Library of Congress issued the 1st edition of MARC21 to replace US MARC. Basically standards of cataloging remain unchanged.

A MARC record is divided into fixed field and variable field. The fixed field is subdivided by textual names, and 3-digit tags represent the variable field. By reading tag numbers, the computer can interpret the type of data such as : author, title, publisher, subject headings, etc. and make proper markings on the online catalogue. The frequently used tags are as follows :

- 010 Library of Congress number
- 020 International Standard Book Number
- 100 Main entry (personal name)
- 110 Main entry (corporate name)
- 245 Title statement
- 250 Edition statement

260 Publication, distribution, etc.

300 Physical description

650 Topical subject headings

7xx Added entries

8xx Series added entries

MARC standard has been widely followed in the library community since its inception, so millions of bibliographical records have been created in this format. Nevertheless, even among librarians, especially for those who are non-catalogers, MARC is **not** known for its simplicity, and it does take practice to understand and apply it correctly.

For more information : [http : //lcweb.loc.gov/marc](http://lcweb.loc.gov/marc)

Dublin Core

With the explosion of Internet information, it is not practical to apply MARC standard, due to its complexity, to cyberspace for data description. There was a crying need to create a simplified method. DC was considered as a solution to this problem.

An OCLC research scientist brought up the idea in 1994, but the development of this 15-element metadata was the joint efforts of Dublin Core workshops. While it can be applied as a traditional cataloging method, it was intended to facilitate the management of Internet resources by non-catalogers. The fifteen elements are :

- | | | | |
|--------------|----------------|------------|-----------------|
| 1. Title | 2. Creator | 3. Subject | 4. Descriptions |
| 5. Publisher | 6. Contributor | 7. Date | 8. Type |
| 9. Format | 10. Identifier | 11. Source | 12. Language |
| 13. Relation | 14. Coverage | 15. Rights | |

For more information : [http : //purl.org/metadata/dublin_core](http://purl.org/metadata/dublin_core)

Pros and Cons of Cataloging Websites

A. Advantages of adding Web sites to OPAC

1. There is so much new and valuable information through Internet, so this is a way to enrich the collection and enhance the value of the library catalogue, and patrons can retrieve library materials and Internet

resources with the same search technique.

2. Libraries will serve their patrons better by providing them with information that is available exclusively on the Internet.

3. The integration of library catalog and Web resources will save library patrons the searching time and eliminate the hassle of juggling between OPAC and Web browser or wading through thousands of results by a search engine.

4. Librarians have expertise in acquiring materials for library users, so their selections on Websites could pose certain professional criteria to the library patrons for the access to the uncensored Internet resources.

5. The abundance of Internet information can serve as alternatives, if appropriate, of print materials, so in some cases adding Web resources to the library catalog is a cost effective way to ease the tightness of library book budget.

B. Obstacles or concerns for cataloging Websites

1. Cataloging Websites is a difficult task, because the volatile and growing nature of Internet resources compounds the description of the information and subject analysis of the piece.

2. It goes without saying that cataloging Websites requires some special training and it holds true even for experienced catalogers. Meanwhile, more staff at the workforce is a must. Is the cataloging operation budgeted for this additional cost along with other expenses such as equipment and software etc.? For this purpose, cost analysis has to be made and tested before an effort of cataloging gets started.

3. In the library, who will be responsible for doing selections? Are the librarians qualified to select the suitable Internet sites for **all** subjects? Should quality of materials weigh more than quantity? Or vice versa. Besides, it has to be taken into consideration that limited access may deprive library patrons of the freedom to knowledge, which plays an important role in information discovery.

4. Due to the variety of subjects, cataloging of some sites could be very complicated and time consuming. How can a planning of workload factor this situation in properly? Where would Internet cataloging stand

in terms of setting priority of workflow?

5. MARC standard is **not** recommended to be observed in cataloging Websites, but it has been used for other types of materials in the cataloging world. As a result, the records will have different appearances on OPAC. Will this cause confusion to patrons? Besides, will there be a mechanism installed to constantly monitor the changing of URL and keep the sites updated since apparently catalogers cannot afford the time to do the maintenance after the record creations?

6. Will this approach really function better than simply book marking Web sites and making book marked pages, independent from OPAC, available to library patrons? Is the justification worth the extra efforts of doing the cataloging?

Tool for Cataloging Websites

When creating records for Web sites, catalogers may choose to do it manually as they do for other materials. On the other hand, if preferred, tools for creating Web-based metadata are evolving. The following are the two types, which have earned some credits to this effect.

A. MARCIt

This commercial software, developed by MARCIt, Inc., is a descriptive cataloging software program. It automatically generates comprehensive MARC records in a library automation system from the site currently displayed in the browser. MARCIt conforms to the most recent MARC record specifications, and can be used in conjunction with Book Where, which is a Windows program that can search online catalogs and databases on the Internet using the Z39.50 protocol. For more information of this software : [http : //www.marcit.com](http://www.marcit.com)

B. CORC (Cooperative Online Resource Catalog)

It was started as an OCLC research project, but made available for the public in January 1999. Its goal is to improve integration of Web and local resources.

CORC is sets of Web-accessible cataloging tools and databases combined. As MARCIt does, it can harvest information from a Web site



and create a record that describes the targeted electronic resource with an option of MARC, DC, or RDF formats for a library automation system. However, the availability of CORC is different from the way MARCIt is made accessible. Since CORC is one of OCLC services, the acquisition is on a membership basis. Meanwhile, in addition to the subscription, charges are also incurred with the utilization of their existing records.

For more information : [http : //www.oclc.org/oclc/corc/](http://www.oclc.org/oclc/corc/)

Conclusion

I will conclude by revisiting one of my favorite quotes of wisdom :

In real situation, there is no panacea for everything! A life saving medicine to one patient could be a fatal poison to the other.

The philosophy of this old saying should be pondered over when making a decision whether to catalog Websites for a library online catalog.

References

- Chepsiuk, Ron. "Organizing the Internet : The 'core' of the challenge," *American Libraries*, 30 : 1 (Jan. 1999) : 60-63.
- Hillmann, Diane. "Using Dublin Core," available at <[http : //dublincore.org/documents/2000/07/16/usageguide/](http://dublincore.org/documents/2000/07/16/usageguide/)>
- Jul, Erik. "Cataloging Internet Resources : An assessment and prospectus," *The Serials Librarian*, 34 : 1/2 (1998) : 92.
- Miller, Paul. "Metadata for the Masses," available at <[http : //www.ariadne.ac.uk/issue5/metadata-masses/](http://www.ariadne.ac.uk/issue5/metadata-masses/)>
- Olson, Nancy. ed. *Cataloging Internet Resources : A manual and practical guide*. Dublin, Ohio : OCLC, 1997.
- Richard, Trina. "Managing Internet Resources : Cataloging," available at <[http : //www.ryerson.ca/~trichard/mir.html](http://www.ryerson.ca/~trichard/mir.html)>
- Robertson, Wendy. "Cataloging Web Resources : CORC and its alternatives," available at <[http : //staffweb.lib.uiowa.edu/wrobertson/ila-talk.htm](http://staffweb.lib.uiowa.edu/wrobertson/ila-talk.htm)>
- Schneider, Karen G. "Cataloging Internet Resources : Concerns and caveats," *American Libraries*, (March 1997) : 77
- Suttle, George. "Library Websites : Cataloging and control," available at <[http : //](http://)



georgesuttle.com/presentation/JMU/IIIA.shtml>

Vellucci, Sherry L. "Options for Organizing Electronic Resources : The coexistence of metadata,"

Bulletin of the American Society for Information Science, 24 : 1 (Oct. 1997) : 14-17.

Weibel, Stuart. "Metadata : The foundations of resource description," available at <[http :](http://www.dlib.org/dlib/july95/07weibel.html)

[//www.dlib.org/dlib/july95/07weibel.html](http://www.dlib.org/dlib/july95/07weibel.html)>

Younger, Jennifer A. "Resources Description in the Digital Age," *Library Trends*, 45 : 3 (1997) :

462-487.