The Electronic Library, 2004, Vol. 22 Iss: 3, pp.238 – 248.
ISSN: 0264-0473
DOI: 10.1108/02640470410541633
http://www.emeraldinsight.com/index.htm
http://www.emeraldinsight.com/journals.htm?issn=0264-0473
http://www.emeraldinsight.com/journals.htm?issn=0264-0473&volume=22&issue=3&articleid=862058&show=html

Metadata and cataloging practices

Magda El-Sherbini and George Klim

Abstract

Metadata standards existing today range from very complex to very simple. Relative simplicity or complexity of metadata standards depends in large part on the resources for which they were created and the depth of description that is deemed necessary to make these resources accessible. This paper reviews the differences between metadata standards and current cataloging practices, and discusses how the various metadata standards are applied in libraries. In addressing these issues, the authors introduce definitions of key concepts of metadata and cataloging standards and provide an overview of the most common metadata schemes. The discussion of current cataloging practices includes an overview of the most commonly used cataloging practices and standards, the impact of metadata on library practice and the role of librarians related to metadata. The authors will discuss the OHIOLINK Electronic Thesis and Dissertations (ETD) as an example of how Anglo-American Cataloging Rules 2nd (AACR2) and Machine Readable Cataloging (MARC21) are used as metadata to store, describe and access this unique information resource.

Introduction

Metadata are data about data. The term came into being with the appearance of electronic resources and it initially referred to standards that assisted in identifying, describing and locating electronic resources. This concept continued to evolve as it was applied to a variety of "non-traditional" formats, resulting in proliferation of metadata schemes.

Although the term was used at first in relation to digital information, general understanding of the term has since broadened to include any kind of standardized descriptive information about resources, including non-digital formats. Some examples of this would include library catalogs, abstracting and indexing tools, and archival finding aids. Metadata standards existing today range from very complex to very simple. Relative simplicity or complexity of metadata standards depends in large part on the resources for which they were created and the depth of description that is deemed necessary to make these resources accessible.

This paper reviews the differences between metadata standards and current cataloging practices, and discusses how the various metadata standards are applied in libraries. In addressing these issues, definitions of key concepts of metadata and cataloging standards are introduced and an overview of the most common metadata schemes is provided. The discussion of current cataloging practices includes an overview of the most commonly used cataloging practices and standards, the impact of metadata on library practice and the role of librarians related to metadata. We will discuss the OHIOLINK Electronic Thesis and Dissertations (ETD) as an example of how Anglo-American Cataloging Rules 2nd (AACR2) and Machine Readable

Cataloging (MARC21) are used as metadata to store, describe and access this unique information resource.

A brief review of literature

Current research on metadata tends to concentrate on the variety of metadata standards and their application. A comprehensive survey of metadata literature is well beyond the scope of this paper, but some emerging trends in current literature are worth noting. Proliferation of metadata research is dictated by the relative flexibility of metadata and their usefulness in describing a growing variety of objects.

In "Practical applications of metadata at Oregon State University" (Banerjee, 1999), the author outlines some of the challenges encountered during a pilot project at Oregon State University, where efforts were made to use metadata to facilitate access to networked university resources through the online catalog. Based on his analysis, the author concluded that a library that is interested in using metadata for a project to access to resources should first develop a plan to address the human, technical, and organizational challenges. On the technical side, the author recommends that a library should also be aware of the existing metadata tools, such as Dublin Core (DC).

A good description of the DC metadata and the means by which they can be implemented was discussed by Miller (1996). The author begins by offering a definition of metadata and discusses existing approaches to metadata (the experts and the search engine approaches). He goes on to review in some detail the DC standard and its constituent elements. In conclusion, he discusses the potential of HTML, metadata creation and search engines.

Much professional literature is devoted to the debate whether to use the traditional cataloging standards to organize networked information or not. Some authors have indicated that the current cataloging tools are too complex and expensive to use for cataloging Internet resources. These authors make an argument in favor of using DC instead (Medeiros, 2000; Chepesiuk, 1999).

On the opposite end of the spectrum, an argument is made in favor of using traditional cataloging tools to describe electronic resources. A number of researchers recommended that the current cataloging practices are sufficient for describing Internet resources and suggest that they should be used for their effectiveness and flexibility to organize electronic information (Mandel and Wolven, 1996).

As the new metadata schemes are introduced, the issue of teaching cataloging standards to librarians and in library schools gains prominence (Hsieh-Yee, 2000; Glaviano, 2000). The authors of these two papers describe some of the methods that were used to teach basic and advanced cataloging. Both authors point out the need to devote a significant amount of time to the treatment of Internet resources. Instructors would do well to provide specific training on the application of metadata and ought to allow for sufficient hands-on practice. Additional discussions focused on how to use DC as a cataloging tool and how educators are incorporating these standards into their cataloging courses.

A fascinating view of cataloging and its future in light of ongoing developments in digital technologies was addressed by Levy (1995). This non-librarian author "explores the nature of cataloging as it is now practiced". The author draws on library science literature, as well as recent research in other disciplines, such as anthropology and sociology. Based on his findings,

he characterizes cataloging as a form of order-making based on the way books are arranged and published. Levy goes on to illustrate how this established order is being challenged by the changes introduced by digital technologies.

Authors of "Metadata: cataloging practice and Internet subject-based information gateways" (Chapman *et al.,* 1998) make an interesting attempt at comparing traditional library cataloging concepts and practices with the template creation in Resources Organization and Discovery in Subject-Oriented Services (ROADS)-based gateways designed to describe Internet resources. In their efforts, they try to define the catalog and its elements. In the conclusion, they found substantial similarities between traditional library cataloging practice and the creation of subject gateways of ROADS template.

This small sample of research trends related to metadata and cataloging reflects the multi-level debate that is taking place between the library community and the information community at large. It suggests that the issue of access to information is no longer the domain of a single profession. The debate is joined by information producers, technology experts and members of a variety of disciplines who shied away from these issues in the past. Many of these findings tend to support the notion that there are substantial commonalities between old and new metadata schemes.

Definition of key concepts

Metadata and cataloging standards

A brief survey of literature reveals that most researches are in agreement as to the definition of metadata. Metadata are "data about data" and not data in and of themselves. They describe the content, quality, condition, and other characteristics of data. Although the term metadata was originally applied to those bibliographic description activities that were aimed at classifying electronic resources, general understanding of the term has since been broadened to include standardized descriptive information about all kinds of resources. This includes digital and non-digital resources alike.

In his "A gentle introduction to metadata", Good (2002) begins with the notion of the humble origins of metadata. He points out that even a simple citation contains basic metadata elements, but argues in favor of a more inclusive approach:

> An annotated bibliography, for example, also constitutes metadata which is very much like a list of references except that it also includes an extra level of description in addition to the basic metadata for the document.

Metadata relating to a print resource may consist of information such as author, title, year of publication, publisher, and so forth. This information may be organized in a card catalog card, or its electronic iteration. Both types of records are held in the library catalog, electronic or otherwise, which then becomes a repository of metadata about materials that are held by that particular library.

In discussing "traditional cataloging standards", reference is being made to those tools that have been developed over time for the purpose of cataloging. These include the AACR2 with its editions, MARC 21 formats and standards. Library of Congress Subject Headings (LCSH), Dewey Decimal Classification (DDC), Sears Subject Headings, and other cataloging

tools that libraries are using to describe and organize knowledge.

AACR2, for example, is an internationally accepted standard for descriptive cataloging. It contains rules for describing and providing access to all types of library materials including books, serials, computer files, maps, music, motion pictures, etc., through library catalogs. AACR2 is also a standard for structuring catalogs with headings and references to provide links between items with similar or related characteristics.

How are metadata different from the traditional cataloging standards? In looking at this issue from the point of view of purpose or intent of metadata, one arrives at the inevitable conclusion that the differences are not substantial. Both approaches attempt to provide bibliographic description. This can be extended further to include the fundamental mechanisms governing the creation and the structure of metadata. Like traditional cataloging standards, it is governed by the same principles, even when those are applied to a diversity of materials.

Milstead and Feldman (1999) argue that the term as applied to electronic resources "refers to 'data' in the broadest sense - datasets, textual information, Web pages, graphics, music, and anything else that is likely to appear electronically".

What, then, is the reason for the evolution of metadata and what distinguishes them from what came to be known as traditional cataloging standards? The word "Internet" provides a short, albeit incomplete answer. The world of information has moved beyond paper and microform as the primary carriers of information. Digital resources have become abundant and with them came the need for classification. With their proliferation came the perception that the available cataloging standards could not be satisfactorily adapted to the demands of these new formats. This development coincided with a new trend in publishing and bibliographic description. Publishers began to provide libraries that acquired their books with skeletal pre-publication descriptions of their books. As these descriptions became more accurate and complete, the libraries saw an opportunity to use them in their cataloging process. The door was opened to the idea that the library was not the only place where information about materials could be built into a record describing the materials.

As the amount of digital, Internet accessible information grew, librarians came to realize the need to apply some sort of scheme to describe them and that they themselves could not deal with the "workload". Not everything that came to be considered information could receive the full cataloging treatment used to describe print materials. Prohibitive costs of full cataloging along with perceived inflexibility of existing cataloging standards were two of the key factors leading to the revolutionary changes in information processing; the development of a simplified, flexible standard or standards of cataloging that could accommodate the diversity of electronic formats, and taking cataloging out of the library.

Old and new metadata are based on common practices. Cataloging standards were created and developed as a way to organize information, in order to facilitate the retrieval and access to this information. Standards are the foundation on which all the cataloging and metadata rules are developed. Without these cataloging standards, a single item would be cataloged many times over and each cataloging record is likely to contain different information. Without the existence of cataloging standards it would be difficult to imagine how scholars could access information, how libraries could share resources, and how patrons could benefit from the library collection. Cataloging standards help to organize knowledge and have served scholars and research very well in accessing information relatively quickly and efficiently.

Cataloging and metadata standards provide consistency and also exhibit tremendous flexibility. As new publishing formats appear; microforms, sound recording, and computer files

could serve as examples; new standards are developed to their description. The most recent developments in AACR2-2002 revision were made to reflect the need to catalog electronic resources.

Types of metadata

Metadata types range from simple formats to rich and complex formats:

- Unstructured data, which are automatically extracted from resources and indexed for use by robot-based Web services such as AltaVista.

- Structured formats, which are simple enough to be created by non-specialist users. Usually manually created, but some data may be extracted automatically. Examples include ROADS templates and DC.

- Structured formats, which are rich and complex to organize complex relations between objects or collections of objects and are often based on implementations of SGML. Examples include MARC21, TEI headers and CIMI formats.

There are several ways to associate metadata with resources (Caslon Analytics, 2003). Metadata can be embedded within the resource by the author at the time of creating the resources itself.

Associated metadata are maintained in files attached to the resources they describe or tightly coupled using protocols or identifiers.

Third-party metadata are maintained in a separate repository (databases) by an organization. A library, for example, could create another version of metadata either from scratch or derived from a previous record. A library may or may not have direct control over access to the content of resources.

Metadata can be embedded in the document or created as a separate record. There are two main approaches to metadata creation. One is to use elaborate and specialized schemes like the MARC21-record. This requires time, money and qualified staff and will probably only be used on a small, choice selection of valuable and stable resources.

The other approach is to create a simple scheme, such as DC, that could be used by the author of a document to create a bibliographic record (Eriksson, 1997). The controversy created by having the author of the resources create the metadata instead of a trained cataloger is still unresolved. Massive numbers of electronic resources that need metadata lead to the conclusion that some metadata can be created by authors, while trained librarians can evaluate and enhance some of these records. Trained catalogers could provide an invaluable service when specific analysis of a collection is required.

Why metadata are important

Metadata describing electronic resources are essential in the retrieval process. Great proliferation of electronic content that is accessible through the World Wide Web makes it necessary to create and apply metadata standards that will make navigation more effective. Existing search engines reach a small portion of Web-based resources, making effective metadata application a necessity.

Recent studies document that although some metadata are applied to many of the Web-

based resources, standard metadata schemes are applied to a very small percentage of these resources:

> A discouraging aspect of metadata usage trends on the public Web over the last five years is the seeming reluctance of content creators to adopt formal metadata schemes with which to describe their documents (O'Neill *et al.*, 2003).

As scholars, researchers and other users continue to increasingly rely on information that is available via the Internet, standardized metadata schemes become an essential element in the information retrieval process.

Who creates the catalog record and who creates metadata?

The job of organizing information for the purpose of providing access to it is the domain of librarians. This profession developed the first classification tools and indexes, which enabled the user to locate print materials.

The need to organize this information came with the creation of the first collections and libraries. As collections grew beyond a single shelf or a single bookcase, locating needed items became an issue. Organizers of these collections had to develop tools that would locate the needed information in the space allocated for print materials.

As the business of producing and distributing information grew in complexity and sophistication, so did the challenges faced by those who collected these materials and who were charged with providing access to them. Demand to create more sophisticated search tools grew with the size of libraries and the user's need to access information in many different ways. Library professionals and practitioners are responsible for the creation of indexes, subject guides, and thesauri as well as the standards that governed the creation of such tools.

The Internet era ushered in a revolution in information creation and distribution. The very concepts of publisher, publication, title, author that formed the foundation of the cataloging record had to be analyzed and redefined. This task fell to professional catalogers, whose expertise and experience continued to be indispensable.

Not all the resources, however, could be cataloged by professional librarians using the MARC21 format. Metadata schemes, such as DC, were introduced to address the needs of the new information formats. The very simplicity of the new standards paved the way for the trend that would take cataloging out of the library. Creation of a DC record does not require the detailed knowledge of cataloging practices that is needed for traditional cataloging.

Development and evolution of metadata standards follow the familiar patterns established in the library community. Metadata came into being to meet the need to classify and to organize the vast amounts of information and data that are produced and distributed digitally. Most, if not all metadata standards are a development, a refinement, or, in some cases, simplification of existing classification tools and standards. The library online catalog serves as an example of the relationship between traditional cataloging and metadata. Records in the online library catalog consist of metadata about books, including information such as physical dimensions, publication dates, and author names. Metadata has existed under various names in the bibliographic description professions for decades, providing enough information to manage and retrieve resources.

The major distinction is that the object of this new classification standard, a digital entity, is not physically located in a library or a collection. This important distinction has contributed to the development of the notion that librarians are no longer needed to develop standards and tools

to access electronic information. The assumption was that rudimentary metadata tags accompanying each electronic object will enable Internet search engines to locate all the information that is needed. This simple metadata could be created by almost anyone, including the publisher of electronic resource, the author or the recourse creator, experts, metadata creation agencies, the Webmasters, or the institution.

What simplifies the process of creating metadata for electronic resources are the metadata tools that have been developed to assist the non-professional in this process. There are a number of metadata creation tools that are readily available. These include "DC-dot"[l] which will retrieve a Web page and automatically generate Dublin Core metadata for embedding in the < META > section of the HTML pages; "Nordic Metadata Project Metadata Templates" [2]; and "Reggie Metadata Editor" [3].

Discussion of metadata standards

Overview of the most common projects using metadata schemes

A review of the most commonly used metadata tools is intended to highlight their key characteristics and commonalities. Whether one looks at new metadata tools, such as DC, or the traditional cataloging tools, such as AACR2 or MARC21, one is immediately aware of the need for standards that are required for the development of effective research tools. The limited number of initiatives that were chosen for discussion here does not exhaust all possibilities but is intended to illustrate some of the key metadata concepts.

National Library of Medicine (NLM) metadata schema

The National Library of Medicine (NLM) has created a standardized set of metadata elements to be used with electronic resources published by the library. This standard is based on the DC metadata element set. In addition, the NLM incorporated to DC Metadata schema some additional elements and qualifiers identified as requirements by NLM[4].

Categories for the Description of Works of Art (CDWA)

Categories for the Description of Works of Art (CDWA) is a metadata standard for art objects, architecture, groups of objects, and visual and textual surrogates. It is created by the Art Information Task Force (AITF) sponsored by the College of Art Association and the Getty Trust. This standard is based on several metadata standards for describing works of art and material culture, such as the Foundation for Documents of Architecture/Architectural Drawing Advisory Group (FDA/ADAG) Guide to the Description of Architectural Drawings data categories and the Art Museum Image Consortium (AMICO) data dictionary.

The CDWA categories are divided into two broad groups:

(1) Object, architecture or group.
(2) Authorities/vocabulary control.

These groups are then divided into categories and subcategories. Some of these categories and subcategories are marked by the task force as core. Those represent the minimum information necessary to uniquely and unambiguously identify a particular work of art or museum object. The categories map to the MARC21 format and the DC Metadata Element Set. The CDWA Web site provides a number of examples to illustrate the use of these categories [5].

The Independents Media Arts Preservation (IMAP)

The Independents Media Arts Preservation (IMAP) group provides a Web site that allows independent producers and art and culture organizations to catalog their collections of digital resources. The IMAP was organized in 1999 with focusing on the preservation of independent electronic media for cultural and educational use by future generations. These works are found throughout the country in museums, arts centers, artists' spaces, dance and theater companies, libraries, university departments, non-profit distributors, public television stations, and with individual artists or producers. The site provides a standardized template for use by these organizations. The template is based on MARC21 and AACR2 cataloging rules. There are several record examples on the IMAP Web site [6].

Recordkeeping Metadata Standard for Commonwealth Agencies

The National Archives of Australia have made available on their Web site a metadata standard that the National Archives of Australia recommends should be captured in the record-keeping systems used by Commonwealth government agencies. This standard is based on DC. Part one of the standard addresses the issue of standardized record-keeping metadata and details the scope, intended application, and features of the standard. Part two of the standard provides details on the basic set of 20 metadata elements and 65 sub-elements and defines them in relation to their purpose and rationale [7].

Online Information Exchange (ONIX)

Online Information Exchange (ONIX) is the international standard for publishers to distribute electronic information about their books to wholesale, e-tail and retail booksellers, other publishers, and anyone else involved in the sale of books [8]. ONIX was developed and is maintained by EDItEUR (an international group coordinating development of the standards infrastructure for electronic commerce in the book and serials industries) [9] jointly with Book Industry Communication and the Book Industry Study Group, and with user groups in France, Germany and Latin America. One of the reasons for creating ONIX is book promotion and sale.

The standard allows for a publisher to use either of two standards - level 1 or level 2. Level 1 contains all the information in level 2. Standard data elements in level 1 are targeted to publishers who have not established an in-house database of product information. Level 2 is targeted for those publishers who feel that level 1 data elements are not adequate.

The ONIX standard defines both a list of data fields about a book and how to send those data in an "ONIX message". ONIX specifies over 200 data elements, each of which has a standard definition, so that everyone can be sure they are referring to the same thing. Some of these data elements, such as ISBN, author name, and title, are required; others, such as book reviews and cover image, remain optional. While most data elements consist of text (e.g. contributor biography), many are multimedia files, such as images and audio files.

An ONIX message is a set of data elements defined by "tags" that is written in the computer language Extensible Markup Language (XML) and that conforms to a specific template, or set of rules, also known as the ONIX document type definition (DTD). The DTD defines, among other things, how to order the data elements, and how the elements are interrelated.

Metadata Object Description Schema (MODS)

Metadata Object Description Schema (MODS) is a bibliographic element set developed at the Library of Congress Network Development and MARC21 Standards Office. It is "intended to be able to carry selected data from existing MARC21 records as well as to enable the creation of original resource description records. It includes a subset of MARC21 fields and uses language-based tags rather than numeric ones, in some cases regrouping elements from the MARC21 bibliographic format" [10].

The Moving to Distributed Environments for Library Services (MODELS)

The Moving to Distributed Environments for Library Services (MODELS) project is funded by the Joint Information Systems Committee of the United Kingdom Higher Education Funding Council under its electronic libraries (eLIB) program. "MODELS is providing a forum within which the UK library and information communities can explore shared concerns, address design and implementation issues, initiate concerted actions, and work towards a shared view of preferred systems and architectural solutions" -the Web site. MODELS was created to manage the heterogeneous information resources and services being offered to libraries and their users [11].

The Resources Organization and Discovery in Subject-Oriented Services (ROADS)

The Resources Organization and Discovery in Subject-Oriented services (ROADS) is a project from the access to networked resources section of the electronic libraries programme (eLib) and is funded by the Joint Information Systems Committee (JISC). The main objective of the ROADS project is to design and implement a user-oriented resource discovery system. The project provides software tools and support for the creation of Internet subject services or information gateways. The ROADS project has some purposes:

- to produce a software package which can be used to set up subject-specific gateways;
- to investigate methods of cross-searching and interoperability within and between gateways; and
- to participate in the development of standards for the indexing, cataloging and searching of subject-specific resources.

The software allows you to set up a subject gateway. This is a Web-accessed database of records of resources, all of which are relevant to one particular subject area, such as medicine, chemistry, football etc. Each resource record consists of a number of fields, e.g. the resource URL, title, a description, and keywords. The gateway can be either searched or browsed by end-users. ROADS used simple templates adapted from anonymous FTP archive templates [12].

OAIster Project

There are many scholarly collections on the Internet that are not indexed in commonly used and popular search engines such as Google and AltaVista. The reason why these resources are not retrievable is because the resources are generated by CGI scripts or exist within repositories that search engines cannot delve into and index.

These information resources that are held in databases or in sites that cannot be found without previous knowledge of their Web address are commonly referred to as the "hidden" or "invisible" Web. Examples of material in these databases hidden from most search engines

include:

- audio;
- e-books;
- images;
- movies; and
- online journals.

OAIster, a Mellon-funded project of the University of Michigan Digital Library Production Services is a search interface project, which allows users to freely access and search a wide variety of digital resources from various institutions. The database currently contains 1,149,563 records from 160 institutions. The project is using Dublin Core as a standard for the description of resources, with each record leading to an actual digital collection hosted at an institution. Users can view each collection separately or search the database by keyword, author, title, or subject.

OAIster uses Open Archives Initiative (OAI)-based metadata harvesting services developed by the Library of the University of Illinois at Urbana-Champaign, to enable the discovery and retrieval of scholarly works, hidden on the Web (in databases, finding aids, and XML documents) beyond the reach of search engines [13].

Other examples of unique metadata projects were discussed in "Metadata and the future of cataloging". In this paper, there are detailed descriptions of the following projects:

- BeOnlin, Alexandria Digital Library Project (ADL);
- Art, Design, Architecture, and Media (ADAM);
- Nordic Literature and Art;
- Australia New Zealand Land Information Council (ANZLIC);
- Federal Geographic Data Committee (FGDC);
- Government Information Locator Service (GILS);
- DC;
- Colorado Digitization Project (CDP); and
- MARC21 (El-Sherbini, 2001).

Discussion of current cataloging practice and metadata

Metadata created by the application of traditional cataloging tools produce the most successful results in resource discovery of print and digital materials alike. National and international libraries alike have followed the basic standards that allow for sharing of bibliographic records and information in a simple way.

Many major libraries are using the AACR2 as well as the Library of Congress's practice in subject heading forms, as represented in the LCSH. They also follow the MARC21 format as a means of encoding metadata defined in cataloging rules. In addition, many libraries are using either OCLC or RLIN databases to share these records.

These traditional standards provide an effective way of describing information and allow for sharing of metadata among various systems. Their flexibility and applicability to many existing cataloging practices, as well as adherence to strict standards makes them invaluable. On the other hand, these methods and practices are labor intensive and may be considered too costly

to make them a viable cataloging tool for the Internet-based resources.

Metadata and traditional cataloging tools and practices have as their objective description and organization of information. Both can be applied to either print or electronic formats.

AACR2 may serve as an example of a cataloging standard that was developed for traditional cataloging, but was sufficiently flexible to adapt to new formats. AACR2 was developed and expanded to accommodate new formats. When microforms became the accepted format for preserving paper formats, AACR2 developed a chapter on cataloging of these materials. With the appearance of Internet resources, AACR2 added a chapter on handling this format. Changes were also made in MARC 21 to accommodate this new format.

There is a great deal of flexibility within traditional cataloging tools that allow librarians to effectively catalog and classify library collections. These tools can be adopted to meet specific needs of an individual library or collection or even particular user groups. Even within the chosen cataloging tool, catalogers and librarians have the option of choosing what level of detail is required to provide an adequate description to an item at hand. The choice depends on the users and the collection. Bibliographic formats and standards, for example, offers two levels of inputting standards for bibliographic records prepared for the OCLC database:

- Level I represents full or complete cataloging.
- Level K represents minimal-level cataloging.

These two levels of cataloging provide the needed flexibility in determining those data elements that are essential. They "are not preventing a cataloging agency from using any valid MARC21 bibliographic data element. They are provided to facilitate the standardization of the content of MARC21 bibliographic records"[14]. In addition to these two levels, OCLC adapted the core-level standard that was defined in 1994 by the Program for Cooperative Cataloging [15].

Development of the Connexion cataloging platform by OCLC provides an environment in which traditional cataloging tools and metadata tools coexist side by side. Connexion will assist libraries in cataloging the Internet resources using either MARC21 or DC standards. This product will enable libraries to apply DC standards to a Web site or an e-book or any other Internet resources and obtain brief bibliographic records based on standards. This bibliographic record could then be exported to the library online system using MARC21 formats[16].

Role of librarians related to metadata

Librarians have always worked with cataloging and classification concepts that form the basis of metadata. Practical experience in library work led to the creation of the existing cataloging standards. The same experience allows today's librarians to adapt the new metadata schemes and, in some instances, to participate in their creation.

Librarians continued to monitor the changes taking place in the information marketplace and they responded effectively to these changes, making new information resources available to researchers regardless of format or mode of distribution. In order to make those adjustments and meet those challenges, librarians had to devise flexible methods and procedures.

Librarians were able to easily adapt the new concept of metadata, because the concept is not new to them, as it remains part of the main current of cataloging and classification tradition. Librarians of today play an important role in the area of networked resources as they continue to

contribute their experience gained in organizing information.

There are seven roles that have been identified for librarians related to metadata (Kung, 1997):

(1) Selection, evaluation and description of networked information.
(2) Coding metadata records.
(3) Standards setting.
(4) Training.
(5) Advocacy.
(6) Document repositories.
(7) Registries (PURLs; metadadat, naming).

Other basic roles of librarians in the area of Internet organization are:

- intermediary;
- guide;
- facilitator;
- educator; and
- Web site builder (Lynch, 1999).

Librarians can continue to play an important role in modify cataloging practices for effective organization and access to Internet resources. In order to use the current cataloging practice and tools, catalogers have to recognize the dynamic nature of Internet resources and modify these current standards to create suitable sets of elements that will ensure quality and at the same time provide faster cataloging as well as access to Internet resources. This modifying standard "aims to fulfill the role of the catalog as a finding tool, an evaluating tool, a collocating tool, and a locating tool by including only elements essential for the identification and subject collocation of Internet resources" (Hsieh-Yee, n.d.).

Librarians have an excellent opportunity to work with computer scientists and professionals from multiple disciplines, to complement each other's skills in information management. With the librarians' traditional skills and the knowledge of the computer specialists, there is a great potential of both sides to cooperate and work side by side to organize the Internet resources.

During the last decade, great emphasis was placed on what librarians can offer in the era of Internet resources. Today, it is obvious that librarians accepted this challenge and have assumed an important role in organizing Internet resources. Nearly every academic library and many public libraries are engaged in projects that involve cataloging of electronic resources and many librarians attempt to integrate the new metadata standards into their work. The challenge that librarians are facing today is in the area of continuing education and training - the profession looks to library schools to assume a leadership role in providing continuing education to traditional librarians.

Can the current cataloging tools be used to locate digital information?

In this section, the authors will discuss OHIOLINK Electronic Thesis and Dissertations Center (ETD) as an example of applying AACR2 and MARC21 as metadata to store, describe and access this unique information. The focus will be on cataloging ETD at the Ohio State

University Libraries.

Until recently, theses and dissertations were published only in paper copy and also preserved on microforms. Graduate student theses or dissertations were submitted to the graduate school and a bound copy of each thesis and dissertation was placed in the university library where it was available for general circulation. A copy of each Master's thesis was sent to Assured Micro-Services, where a microfiche copy was produced. This copy was also stored in the library. One copy of each doctoral dissertation was forwarded by the graduate school to University Microfilms International (UMI) for microfilming and publishing copies on demand.

Cataloging and processing of theses and dissertations is handled differently at various institutions. At the Ohio State University Libraries (OSUL), paper copies of theses and dissertations receive minimum level cataloging with no subject headings and no abstract.

This method of processing and cataloging of theses and dissertations makes it difficult to gain quick access to these materials. The processing cycle encompasses numerous steps, including duplication, binding and handling of multiple formats. Cataloging is often delayed since the microfiche and the hard copy are not received at the same time. The minimal level of cataloging does not provide sufficient access points to make this resource fully available to users.

Technological developments of recent years allowed libraries to improve on this process. Beginning in the fall of 2002, all Master's theses and, optionally, doctoral dissertations at the Ohio State University can be submitted electronically. This change in university procedures has a significant impact on processing and cataloging of theses and dissertations.

Initially, during the transition period, when both the paper and the electronic copies of a theses were still in use, a cataloging policy was implemented whereby the existing catalog record for the paper edition would be used to catalog both formats, with fields 856 (URL) and 530 (system requirements note) being added to this record.

After the requirement to submit a paper copy was eliminated by the graduate school, and all theses could be submitted electronically, OSUL decided to use the existing policy for cataloging theses and dissertations, with some adjustments to fields 530, 520, 650s, and 856. This decision was the most logical one and did not require a lot of additional training for those catalogers who were already familiar with computer file cataloging.

This change was facilitated by the cooperation with the Ohio Library and Information Network (OhioLINK), a consortium of Ohio's college and university libraries and the State Library of Ohio. OhioLINK's Electronic Thesis and Dissertations Center (ETD Center) is archiving theses and dissertations for several academic libraries, including OSU [17].

Before embarking on this project, OSU librarians were faced with the decision of whether to use the existing cataloging tools or explore the possibility of using some metadata standard that was developed specifically for electronic resources. Since OSU Libraries wanted to incorporate the theses and dissertation records into its online catalog, it was deemed appropriate to use the current theses and dissertations guidelines. OSUL chose to use the current cataloging standards (AACR2 and MARC21 format) to catalog these materials.

In addition, catalogers took advantage of the information provided by OhioLINK EDT Center to enhance the bibliographic records by adding field 520 (abstract), and field 650 (subject headings). The addition of field 856 allows those searching the catalog to link directly to these scholarly materials.

Introduction of the OCLC Connexion interface is likely to streamline the process and offer other options. Connexion will allow the cataloger to create a basic bibliographic record using either DC or MARC21. These records can then be exported to the local online system.

Conclusion

Bibliographic standards were developed to address the need for organizing information. Their specifics and sophistication were dictated largely by the form of the object they were intended to describe. These standards developed over time to address the needs of the newly developing carriers of information. Whether we call it metadata or cataloging standards, bibliographic standards will continue to be of value. The requirement to describe physical attributes of an item is likely to remain with us, as cataloging and its standards continue to be refined, renamed or reinvented.

Having said this, one cannot help but notice a certain duality of attitudes in the library community, when the topic of metadata and cataloging standards is raised. On the one hand, there is strict commitment and adherence to the standards that are universally recognized as necessary for the processing of information. On the other hand, there is the realization that the library community cannot hope to cope with the daunting task of "cataloging the Internet" using the costly, labor-intensive traditional methods.

The example of the Electronic Thesis and Dissertations Center project initiated by OhioLINK and the OSU Libraries serves as an example of an approach, limited as it is, that could serve as a model of merging the old with the new, of applying traditional standards for handling of electronic formats.

Traditional standards of cataloging and classifying of information continue to be a viable resource. They exist alongside the new metadata standards that are being developed by the information industry, often outside the traditional library world. In an attempt to meet the challenge of the digital age, libraries are finding innovative ways of making use of both approaches.

Web sites

1. DC dot: Dublin Core metadata editor, available at: www.ukoln.ac.uk/metadata/dcdot/

2. Nordic Metadata Project. Dublin Core Metadata Template, available at: www.lub.lu.se/cgi-bin/nmdc.pl

3. Reggie - The Metadata Editor, available at: http:// metadata.net/dstc/

4. National Library of Medicine (NLM), available at: www.nlm.nih.gov/tsd/cataloging/metafile.html

5. Categories for the Description of Works of Art (CDWA), available at: www.getty.edu/research/institute/standards /cdwa/index.html

6. Independents Media Arts Preservation (IMAP), available at: www.imappreserve.org/index.html

7. Recordkeeping Metadata Standard for Commonwealth Agencies, available at: www.naa.gov.au/recordkeeping/ control/rkms/introduction.html

8. The ONIX for Books Product, available at: www.editeur.org/onix.html

9. EDItEUR, available at: www.editeur.org/

10. Metadata Object Description Schema (MODS), available at:
    www.loc.gov/standards/mods/
    http://xml.coverpages.org/MODS-Trialllse200206.html
    www.sla.org/chapter/cdc/presentations/20030204.html
    www.freelists.org/ archives/consalacl/07-2002/msg00000.html

11. Moving to Distributed Environments for Library Services (MODELS), available at:
    www.ukoln.ac.uk/dlis/models/

12. Resources Organization and Discovery in Subject-Oriented Services (ROADS), available at: www.ilrt.bris.ac.uk/roads/

13. OAlster, available at: http://oaister.umdl.umich.edU/o/ oaister/

14. MARC21 Format for Bibliographic Data: National Level Record - Bibliographic Full Level & Minimal Level, available at: http://icweb.loc.gov/marc/bibliographic/nlr/

15. Program for Cooperative Cataloging: BIBCO Core Record Standards, available at:
    http://lcweb.loc.gov/catdir/pcc/ bibco/coreintro.html

16. OCLC Connexion, available at: www.oclc.org/connexion/

17. OhioLINK EDT Center, available at: www.ohiolink.edu/etd/

References

Banerjee, K. (1999), "Practical applications of metadata atOregon State University", Oregon Library Association Annual Conference, Seaside, OR, 2 April, available at: http://home.earthlink.net/ -banerjek/papers/olal 999.html

Caslon Analytics (2003), "Caslon Analytics Profile: metadata, engines, directories", available at: www.caslon.com.au/metadataprofile.htm

Chapman, A., Day, M. and Hiom, D. (1998), "Metadata: cataloguing practice and Internet subject-based information gateways", *Ariadne*, No. 18, available at: www.ariadne.ac.uk/issue18/metadata

Chepesiuk, R. (1999), "Organizing the Internet: the 'core' of the challenge", American *Libraries*, Vol. 30 No. 1, pp. 60-3.

El-Sherbini, M. (2001), "Metadata and the future of cataloging", *Library Review*, Vol. 50 No. 1/2, pp. 16-27.

Eriksson, J. (1997), "Metadata - the description of electronic objects, documents and services", available at: www.lub.lu.se/biblutb/vt97/metadata.html

Glaviano, C. (2000), "Teaching an information organization course with Nordic DC metadata creator", *OCLC Systems & Services*, Vol. 16 No. 1, pp. 33-40.

Good, J. (2002), "A gentle introduction to metadata", available at: www.language-archives.org/documents/gentle-intro.html

Hsieh-Yee, I. (2000), "Organizing Internet resources: teaching cataloging standards and beyond", *OCLC Systems & Services*, Vol. 16 No. 3, pp. 130-43. Hsieh-Yee, I. (n.d.), "Modifying cataloging practice and OCLC infrastructure for effective organization of Internet resources", available at: www.oclc.org/oclc/man/colloq/hsieh.htm

Kung, T. (1997), "Metadata, libraries and librarianship: speech notes", 63rd IFLA Council and General Conference. UDT Core Program Workshop, 4 September, available at: http://216.239.37.100/search?q =cache:aFnNoZVQjmwC: www.ifla.org/IV/ifla63/63kuny2.pdf+terry+kuny&hl= en&ie = UTF-8

Levy, D.M. (1995), "Cataloging in the digital order", available at: www.csdl.tamu.edu/DL95/papers/levy/levy.html

Lynch, C. (1999), "Beyond the basics: role of librarian in the Internet environment", available at: www2.vuw.ac.nz/staff/alas tair_smith/beyond/2role.htm

Mandel, C.A. and Wolven, R. (1996), "Intellectual access to digital documents: joining proven principles with new technologies", *Cataloging and Classification Quarterly*, Vol. 22 No. 3/4, pp. 25-42.

Medeiros, N. (2000), "On the Dublin Core front", *OCLC Systems & Services*, Vol. 16 No. 1, pp. 41-3.

Miller, P. (1996), "Metadata for the masses", available at: www.ariadne.ac.uk/issue5/metadata-masses/

Milstead, J. and Feldman, S. (1999), "Metadata: cataloging by any other name", *Online*, Vol. 23 No. 1, pp. 24-6, available at: www.infotoday.com/online/OL1999/milstead1 .html

O'Neill, E.T., Lavoie, B.F. and Bennett, R. (2003), "Trends in the evolution of the public Web 1998-2002", *D-Lib Magazine*, Vol. 9 No. 4.

Further reading

(The) Montague Institute (1998), "Why is metadata a hot topic?", highlights of a panel discussion on metadata presented by the Boston chapter of the Special Librarians Association (SLA), April, available at: www.montague.com/review/meta.html