Queens College of the City University of New York

Graduate School of Library and Information Science

GSLIS 709 Research and Bibliographic Methods

Professor Karen P. Smith

The Many Faces of Metadata: A Content Analysis of the Effects of Information Technology Upon Cataloging Theory and Practice Within the Library Setting

Brian Rocco

May 24, 2004

ABSTRACT

Recent technological advances in computer memory, storage and processing power, coupled by an increase in bandwidth for network connection, have revolutionized the world of communication. Institutions worldwide must readily adapt to new hardware and software configurations in order to keep abreast with competition. As access to the Information Highway can be easily obtained remotely, the library as an information authority is threatened. What is critical for the library discipline to realize is the fact that they must meet these new challenges and evolve into a "modern" information provider. The library community currently faces a period of transition where they must reevaluate their antiquated cataloging theories and practice, and adjust to contemporary strategies. This research project addresses this concern and investigates the markup languages and metadata schemas being deployed by the online infrastructure—and how the library community can greatly benefit from them.

This study examines contemporary cataloging theories and practices expressed in articles published in library science journals. An annotated bibliography was created to structure and summarize theories propounded. As a result of the research, it has become clear that the MARC record format and *AACR2* rules of description are nearing obsolescence. Recommendations for markup languages and schemes practical for the future of bibliographic control and description are examined.

TABLE OF CONTENTS

CHAPTER I

THE PROBLEM

Introduction

By the late 1960s, technological advancements integrated within the library environment resulted in the reconstruction of both theory and practice of library cataloging. Mainframe computers, Library of Congress standards, and the MARC record format all contributed to the reevaluation and reconfiguration of the library catalog. Empowered by network capability, libraries were beginning to realize that their strength and functionality were multiplying considerably—the ability to share resources and data was converting the local library community into a global "power." The card catalog was nearing obsolescence as the library catalog was beginning to go virtual.

The twentieth century gave birth to international standards of bibliographic description. In essence, the twentieth century was the era of codes (Taylor 45). For contemporary libraries, the encoding of texts and documents must obviously meet new criteria in order to be operable in today's digital environment. Centuries previous, concern for bibliographic description was handled quite differently. Nonetheless, basic principals of cataloging and the organization of information practiced long ago are relevant and still critical to understand in today's modern, technological milieu. A summary of the history1 of cataloging is important to discuss in order to establish the basis from which today's theories evolved.

---

1 The following synopsis is loosely based on Arlene Taylor, *The Organization of Information,* Chapter 3, "Development of the Organization of Recorded Information in Western Civilization."

A Sumerian tablet which contained sixty-two literary titles, estimated to be created around 2000 B.C, was found at Nippur. (Twenty-four of those titles are resources present in today's global network.) Though it is not irrefutably ascertained that this tablet was in fact a catalog, many historians feel that, among other functions, the tablet was indeed used for cataloging bibliographic resources of the community. .

The Hittites, around 1500 B.C., apparently discovered the practical functions of recording bibliographic information with the written work: bibliographic description was born. Their tablets incorporated the number of tablets in a series, its title, and frequently the name of the scribe.

Greek civilization, it has been argued, engendered the first cataloging schema, accredited to Callimachus, who some argue was the very first cataloger. Callimachus employed *pinakes*—a tablet with wax in the middle—to record bibliographic information. Entries were arranged under general subject, and then grouped either alphabetically or chronologically. This practice gave modern society the premise that the "main entry" ought to correlate to "the author" (Taylor 39).

Church and monastery libraries were prevalent during the Middle Ages. Monks copied numerous theoretical texts which resulted in the creation of a system to list the copied texts. In other words, in order to keep track of what the monastery library had successfully finished documenting, an inventory list of the plentiful copies of text was being constructed. It was rare to find any bibliographic control concerning these lists, i.e. no ordering scheme existed. However, towards the end of the thirteenth century a project that was started but never finished is considered to be a landmark in cataloging theory. This is known as the *Registration Librorum Angliae*, which was essentially a union list of

the holdings of English monastery libraries; each library was assigned a number for coding practicality—very much similar to today's implementation done by union catalogs.

Arguably, the fourteenth century and European Renaissance had as much an impact on scientific and artistic thought as it did on library and bibliographic theory. Simple bibliographic lists began to morph into "complex" shelflists. Breakthroughs in cataloging theory included call numbers, content description, and something paramount to library cataloging progression: analytic entries and alphabetic listings (Pass 12).

By the fifteenth century, cross-referencing was beginning to become a common practice amongst scholars, theologians, and researchers. But by the middle of the fifteenth century the invention of the printing press had without question the greatest impact on the theory and practice of bibliographic control. Now faced with identical versions of texts, the role of a bibliographer was born. By the end of the fifteenth century, Johann Tritheim, a German bibliographer and librarian, appended to his bibliography an alphabetical author index, which was indeed a revolutionary idea and concept (Pass 34).

Over the next three centuries, the concepts of bibliographic control began to evolve as a direct result of Guttenberg's revolutionary invention. Indices incorporating authors and subject terms began to prosper. Moreover, as an indirect result of the French Revolution, the French government created instructions for creating a national code; and thus the first use of card catalogs began to be implemented.

With the start of the nineteenth century, and as the production of texts proliferated, arguments on the theory of cataloging prospered. In 1876 Charles Cutter

published his seminal *Rules for a Printed Dictionary Catalogue.* Cutter professed cataloging principals that are still honored today, i.e. catalogs should be recognized as collocating devices; rules for subject headings as a path to gain subject access to catalogued material; as well as rules for instruction on description and filing.

Acceptance of an international cataloging code was heavily debated throughout the first half of the twentieth century. It was not until 1961 with the "Paris Principals" that standardization for cataloging was agreed upon. From there, international standards such as the *Anglo-American Cataloging Rules (AACR),* originally published in 1967; and the *International Standard Bibliographic Description* (ISBD), conceived of in 1974, set the stage for international cataloging interoperability. However, it wasn't until the invention of MAchine Readable Cataloging (MARC) format in the late 1960s that the interoperability of library catalogs reached its apex. As a direct result of the MARC records, OCLC was created, enabling international catalog sharing via cable and terminal communication.

Background of Problem

Unconditionally accepted worldwide for decades, the MARC record format was the initial standard encoding format for bibliographic data in libraries. For years MARC met no comparable alternative for describing content of and format for library resources. However, recent technological advancements-- namely breakthroughs in memory, storage, and processing power-- have forced librarians worldwide to reevaluate the standards of bibliographic control practiced for close to forty years. In other words, the new technological advancements have begun to highlight the inherent weaknesses of the

MARC encoding schema, as well as, to a certain extent, create new dilemmas with the format. Moreover, the World Wide Web is widely recognized as a powerful information tool; but is also admitted universally to be hindered by a lack of bibliographic control—or to put it a bit more dramatically: disturbed by bibliographic chaos. Should librarians follow the old adage: "If it ain't broke, don't fix it," and continue employing MARC for bibliographic control? Or are the new technological advancements a reasonable and beneficial replacement?

Statement of Problem

The internet has had an enormous impact on almost every facet of business, entertainment, education, and communication. Institutions and industries worldwide have readily adapted to the digital revolution by reconfiguring both hardware and software dependencies. But this revolution has by no means reached its summit—institutions must keep up-to-date with technological progress. The library community presently faces a period of transition where the reevaluation of the theory and practice of cataloging must be reconsidered, and perhaps even drastically altered. In particular, the MARC record format was conditioned to describe *physical* objects, i.e., a text, a map, a sound recording device. However, the digital revolution has created innumerable resources that have no physical properties—they just exist within the internet ether. Libraries today are faced with the daunting challenge of not only cataloging new forms of resources, but also rethinking the standards they have successfully practiced for a number of years. Moreover, due to the global scope of the Internet environment, revisions to standards must be agreed upon on a domestic and international level.

Purpose of Study

The purpose of this research project is twofold. Firstly, the researcher will investigate the new encoding schemas and metadata standards arising in conjunction with the World Wide Web. Secondly, and more importantly, this project is designed to exist as a practical tool of access for librarians and library science students to consult for an introduction and evaluation of the current theoretical writings on the importance of metadata reinterpretation and application inspired by the digital revolution.

Conceptual Assumptions

For the purpose of this research project, the following conceptual assumptions are necessary in order to foster the magnitude of the hypothesis and speculation presented by the research.

1. The Internet and World Wide Web will not be superseded by any means of information technology within the next fifty years.
2. That XML, the en vogue encoding schema, will not soon become obsolete due to a new and more efficient schema.
3. The importance and functionality of MARC—the current standard for bibliographic description—and other metadata languages is fully understood and appreciated by the library community.

Statement of Hypothesis

This research project examines the effects the digital revolution has had on the library community. In effect, the researcher speculates that information technology has reinvented library cataloging theory and practice.

Research Questions

In order to study the full effect information technology has had on library cataloging, the following four research question need to be examined:

1. How functional are the MARC record format and AACR2 rules for bibliographic description in the contemporary library environment that houses "modern" forms of documents?
2. When migrating records to new standards and schemas, what factors do librarians face in order to enable interoperability?
3. How important is it for librarians to both understand and implement eXtensible Markup Language (XML)?
4. What are the new metadata standards available and appropriate for digital and special collection library cataloging implementation?

Importance of Study

Many librarians today are astonished by the evolution of metadata standards and its many configurations engendered by the World Wide Web. Of vital importance to any library is a delineation of the "many faces" of metadata. Librarians and institutions must continue to acquaint and educate themselves with not only new metadata standards, but also the corresponding theories and applications. This research study is an introduction to the complex world of metadata and cataloging practice.

Definition of Terms

Terminology apparent to catalogers and metadata implementers will be used throughout the course of this assignment. A concise list of definitions follows. In

today's world of technology, the use of acronyms is rampant. However, acronyms are not included in this list. Whenever an acronym is referenced during this project, it will be followed by what it stands for.

1. Access point**:** Any term in a bibliographic record that may be used to locate a record.
2. Authority control: The development of maintaining consistency in the verbal form used to reference access points, i.e. names, works, subjects.
3. Bibliographic Control: The act of creating, arranging, and maintaining systems for information retrieval.
4. Bibliographic Description: The process of identify elements that identify an information package
5. Document: An information package.

Scope and Delimitations of Research

This research project will limit its study to academic libraries and special collections. The impact information technology is having on the rethinking of library cataloging is examined. In order to explore the modern theories evolving within the library community, the researcher examined theories propounded and cataloging projects discussed by library scholars in journal articles and conference papers. Due to the fact that information technology is rapidly evolving, the scope of years on the published articles is within the past five years.

Cataloging theory directly associated with both public libraries and school media libraries will not be evaluated. Nevertheless, the principals and theories expressed via the

following research will more often than not correlate to the interests of both public and school media libraries.

The main focus of this project is to examine the impact the Internet and the World Wide Web has had on the organization of documents. Obviously, the clear classification of web pages as “documents” is complicated by its proclivity to exist as multimedia constructions; this is something the research project will take into consideration during the course of the project. In other words, the project’s emphasis will be on ostensible documents, and will only peripherally address concerns of multimedia information objects as they correlate to the traditional document-- and not as individual independent information objects.

CHAPTER II

REVIEW OF RELATED LITERATURE

Overview

In order to recognize the present challenges confronting bibliographic description and control in the library community, it is important to fully investigate the role the Internet has had on information science. In particular, an examination of the history and functionality of markup languages and its impact on bibliographic description is vital. Once a framework on the development of the Internet is formulated, as well as an introduction to the corresponding language syntaxes, this chapter will investigate current trends and practices by institutions outside of the library community implementing markup schemas for bibliographic description and control. But before delving into the World Wide Web infrastructure, it is important to discuss various other mechanical impetuses that spurred the art of bibliographic documentation previous to the Internet.

Historical Background of Documentation

The Documentation Movement in Europe in the 1890s signaled the beginning of the development of the mechanization of bibliography (Taylor 50). By the nineteenth century the mass output of journals and text produced by the printing press was overwhelming the science and humanities culture. The concept of bibliographic control was “literally” drowning beneath piles and piles of bibliographic material. Paul Otlet and Henri LaFontaine, the progenitors of the Documentation Movement, recognized the problem and directed a movement to create bibliographic information beyond just the text itself, towards points of access to parts of books, articles within journals, to archives,

photographs, and newspapers. In short, the main goal of the Documentation Movement was to "capture, record, and provide access to all information in all formats for the improvement of science" (Taylor 50).

Otlet and LaFontaine then went on to organize a conference in Brussels in 1892 to foster their concept Universal Bibliographic Control (UBC). Indeed, the conference resulted in various revolutionary concepts of bibliographic control, leading to the realization that a new technology was needed in order to fulfill these preconceived ideas. It was not until 1928 that a new technology capable of meeting the UBC's standards was created when Eastman Kodak invented microphotography. The new medium drastically increased the speed of comprising, collecting, storing, and accessing large volumes of documents. In 1937, the concept of Documentation was taken over by the Americans as the American Documentation Institute was founded. Within a year's time, the institute formulated the Universal Decimal Classification, a program dedicated to the complex analysis and organization of documents.

The importance and "power" of bibliographic control grew considerably due to America's involvement in World War II. Pressured by the desire to develop the atomic bomb before the Soviets, a huge output of the documentation of scientific information occurred. The importance of secrecy and the ability to rapidly disseminate documents from lab to lab pressured scientists. Another impact WW II had on the development of bibliographic control occurred when the Allies confiscated enormous amounts of scientific literature form the Germans. Furthermore, in the 1950's, the launching of Sputnik and the establishment of the Soviet's All Union Institute for Scientific and Technical Information (VINITI) also inspired the art of documentation (Garfield). Thus,

both the Second World War and the cold war garnered unparalleled intellectual attention and economic support for bibliographic control.

In 1945, one journal article extended all theoretical boundaries harnessing documentation and information retrieval. The seed of the theory behind today's information infrastructure began to germinate when Vannevar Bush published his article "As We May Think." Bush envisioned the precursor to computer technology with his idea of a device he called "memex." Bush fully grasped the complication of the lack of bibliographic control affronting researchers:

> There is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered by the findings and conclusions of thousand of other workers—conclusions which he cannot find time to grasp, much less to remember, as they appear (Bush, Screen 2).

Theoretically speaking, the memex allowed one to store enormous amounts of information which, by mechanization, could be retrieved rapidly. In his seminal article Bush, amazingly enough, goes on to delineate the hypertext world of today. Powered by associative indexing, Bush's memex envisioned the linking of items that would enable access to related documents. Bush's meditations would come to fruition decades later and drastically restructure the role of information and communication in modern society.

The Rise of the Internet

J.C.R. Licklider of MIT in 1962 began theorizing about an interconnected set of computers that would allow users to quickly access data from any site. Licklider termed his proposed concept the "Galactic Network." As head of the computer research program at DARPA, Licklider professed theories on networking were continued by successors

Ivan Sutherland, Bob Taylor, and MIT researcher Lawrence G. Roberts. Another fellow MIT colleague, Leonard Kleinrock, published the first paper on packet switching theory, which was a groundbreaking step toward the reality of computer networking. By 1966, Roberts, through DARPA, created his concept of computer networking which he called "ARPANET." The initial line speed of the project was a paltry 2.4 kbps; it was upgraded to 56 kbps (the line speed of today's dial-up connection) when MIT, RAND, and NPL collaborated on the project. The ARPANET eventually grew into the Internet as we know it today.

A major impetus to the success and rapid growth of the Internet has been the free and open access to documents. The research community behind ARPANET endorsed the academic tradition of the open publication of theories and conclusions. Nevertheless, "the normal cycle of traditional academic publication was too formal and too slow for the dynamic exchange of ideas essential to creating networks for network communication" (Leiner *et al*.)

In 1969 an important step was taken by S. Crocker for the development of internet documentation. Crocker established the Request For Comments (RFC) series of notes. The intent of these memos was to be shared quickly and informally; but through postal mail it lacked the former. But with the advent of File Transfer Protocol (FTP), RFCs could now be prepared as online files and quickly accessed through FTP. Today, RFCs are recognized as the "documents of record" in the Internet engineering and standards community. The potential of the power and speed of the Internet was exploding.

The Advent of Metalanguage

Standard General Markup Language

With the rise of the Internet and production of electronic documents, a call for a standardization of document form was needed to facilitate and promote dissemination through the new medium. In order for documents to be interchangeable, a common markup language to foster standardization was essential. In 1986, ISO (the International Standards Organization) published SGML (Standard General Markup Language) as Standard IS08879. Officially endorsed by ISO, SGML was quickly accepted by numerous national and international organizations and software developers—easily, for all were almost impatiently waiting for a standardization that could foster productivity. As a metalanguage, SGML has three important characteristics that prove its usability and importance.

First and foremost, it is important to understand that SGML emphasizes descriptive rather than procedural markup (which has both limitations and benefits, something that will be discussed later in this chapter); secondly, its *document type* concept; and lastly, its independence of both hardware and software platforms representing the script in which the text was composed. It is important to further investigate each of these characteristics in order to understand the importance and functionality of SGML.

A descriptive markup system employs markup codes in order to provide names to categorize individual parts of a document. For example, the markup code <para> identifies a portion of a document and states that it should be treated as the starting point of a paragraph. Dissimilarly, a procedural markup system defines what processing will

occur at specific points in a document; for example: move the left margin 4 quads left, move the right margin 6 quads right, and skip down three lines. This formatting of the document is normally collected outside of a SGML document in separate programs. So, with descriptive markup the same document can be processed by various pieces of software, where many different processing instructions can apply to the relevant part of the document.

The next key to comprehending the nuances of SGML is that it initiates the concept of a document type, and thus the creation of the document type definition (DTD). Just like other objects processed by computers, documents are considered to possess distinct types. The definition of a document is defined by its elemental parts and structure. For example, the definition of a newspaper article might be that it contains a title and a byline. So a document lacking a byline could not officially be considered a newspaper article DTD.

Why "define" documents? For one, a variety of documents of the same type can be uniformly processed. Programs can then be constructed to manipulate the information controlled in a structured format and thus have it behave in a more intelligent and productive fashion.

Lastly, and perhaps most importantly, the dissemination of documents encoded properly with SGML would be both software and hardware independent for both front and end user. In other words, SGML documents are transportable form one environment to another without loss of information.

Understanding the textual structure of a text is important in both the analog and digital form. The dynamics of an analog text today are inherently understood by the

reader—the physical and logical structure is understood (at least the latter is when not fixed by arcane or postmodern structuring). SGML illustrates a dependable mechanism for the markup or identification of structural textual units. Moreover, it delineates the rules for how combinations of units can expressively appear in any text. Each structural component of an SGML document-- a textual unit-- is known as an element. Obviously, different types of elements are given different types of names, but other than its relationship to other element types, SGML offers no way of expressing the meaning of an individual element. What should be stressed and understood is that SGML is "entirely unconcerned with the semantics of textual elements" (*Text Encoding Initiative*).

Within a SGML document, each element must be justifiably marked or tagged to properly function. SGML offers a selection of options, the most common being an insert tag at the beginning of the element and another at the end: the start-tag and end-tag. Similar to conventional punctuation, the start- and end-tag pair separates different elements and their functions. For instance, the quotation in the previous paragraph would be encoded as such:

> …that SGML is <quote> entirely unconcerned with the semantics of textual elements </quote>

The forward slash, or solidus, in the end-tag differentiates it from the start-tag. Elements, too, can be "empty," with no data between the tags. Frequently elements will be embedded within other elements, that is to say contained entirely within a corresponding element. The following is an example of a text encoded using SGML syntax, though for sake of example, it appears disassociated, double spaced and in its simplest form in order to represent a clear and easy example of elements embedded:

```
<anthology>
<poem><title>Dying and Birth</title>
<stanza>
<line>Dying and giving birth go on</line>
<line>inside the one consciousness</line>
<line>but most people understand</line>
</stanza>
<stanza>
<line>the pure play of creative energy></line>
<line>how inside that, those</line>
<line>are one event</line>
</stanza>
</poem>
<!-- remaining poems go here -->
</anthology>
```

Close inspection of the above encoded text reveals that the syntax is too verbose and unstructured. In other words, the presence of "over tagging"—or how this example makes no assumptions about the rules governing the document; for instance, whether or not a title can appear in other places besides preceding the first stanza. Without structured document definition, every element must exist with a start and end tag and thus the syntax becomes too lengthy. Rules which state which elements can be nestled makes the encoding both easier to compose and decipher.

Creating such rules are the first stage in formalizing the structure of a SGML document— the document type definition, or DTD. In formulating a DTD, depending on the document and designer, the structure may be exceptionally complex or relatively simple. What is essential to remember when creating a DTD is that every document type definition is an interpretation of a text. In other words, "there is no single DTD which encompasses any kind of absolute truth about a text" ("Defining SGML"). SGML DTDs are conducive to environments where a uniformity of a document structure is in demand. In such settings, documents are viewed as "untrained" material to adhere to a rigorous set of rules. Placing these rules on a particular discipline of documents not only reduces the burden of creating markup elements for a scholar, but it also imposes a strict interpretation of the discipline inherent in the proposed document structure.

. What needs to be understood is that SGML was created for institutions that would be publishing and exchanging documents *en masse*, such as large publishing houses or governmental sectors. In addition, SGML was created to develop quality control tools within production workflows (Cave 147). To further elucidate on the functionality of SGML, its inherently weaknesses and strengths needs discussion.

The major forte of SGML is that it imposes no fixed tag set. So, if a certain institution within an industry is employing a SGML DTD conducive to that environment has a unique "widget" it must encode, no obstacles are present. In other words, if an industry needs to represent a new information object, freedom to do so is permitted. SGML imposes a rigorous standardization that must be met. Errors are not allowed and must be corrected. Corrections may indeed be labor intensive but will ensure the

validation of the document. SGML's strong emphasis on the separation of form and structure leads to programming advantages—Footnote—see SGML FAQ book
SGML rule implementation fosters uniformity of syntax and discourages, as DeRose puts it, "variations on a theme" (3).

However, the weaknesses of SGML have prevented it from being implemented by a wider population. SGML's major flaw is that its syntax is *extremely* complicated. SGML has too many strict sub-rules that confuse implementers; and seeing as the syntax was created with a publishing viewpoint and fundamentally ignorant to computer science and formal language theory, designers are greatly limited (DeRose 4). Moreover, due to the complexity of its design, SGML systems are uncommon and expensive.

Ostensibly, SGML was created for the management of a grand scale of documents subject to revision and multiple printings. For government and large scale publishing houses, the standardization, relative extensibility, and the rigorous rule structure of SGML proved invaluable for bibliographic description and control. SGML was too complicated to be extended to the laymen. In order for the potential of the internet to be fully realized, alterations needed to be made to the standard general markup language.

Hyper Text Markup Language

A common misconception is that the World Wide Web (WWW) is synonymous with the Internet. To state the difference tersely: the Internet is a global network of millions of computers capable of exchanging data, news, and opinions; the World Wide Web is "an interface and navigation tool that provides a means of structuring Internet

documents and relating them to other documents so that the maximum use can be made of the Internet resources" (December and Randall 8).

The WWW is a brainchild of a physicist from Switzerland named Tim Berners-Lee. While working in a computer service section at CERN, the European Laboratory for Particle Physics in Geneva, Berners-Lee came up with the concept to promote productivity within his institution, unaware that it would eventually be practiced on a global level. Similar to the theories that engendered ARPANET, Berners-Lee sought to create an infrastructure that would enable researchers from various sites to share resources and files. The revolutionary nuance to Berners-Lee's concept is that-- besides being able to download files to individual computers-- one could actually link the text in the files themselves—a move from sole downloading to linking amongst the downloaded. Berner-Lee's concept revolutionized a bibliographic concept first established over four hundred years previous: cross-referencing. The origins of the term "web' was derived from the idea that "documentation of a scientific and mathematical nature would thus be represented as a `web' of information held in electronic form on computers across the world" (Longman 1998).

In the early 1990s, simultaneously with Berners-Lee concept, emerging developments in communication technologies would directly correspond and compliment Berners-Lee's conception. For one, hypertext, a concept originally envisioned in the 1940s by Bush and discussed earlier in this paper, was beginning to develop. Bill Atkinson, a computer programmer for Apple Computers, developed an application called Hypercard, which, through clicking and linking, enabled a user of a personal computer to navigate through a tour of information using on-screen buttons. Initially however the

hypertext links could only be made to files on the same computer—the concept of global hyperlinks was yet to be developed.

Another internet improvement that complimented Berners-Lee's concept was when the Internet's came up with a new idea of naming computers: the domain name system. As a substitute for much longer and complicated IP addresses, names were comprised of a serious of letters separated by dots, a much simpler means of identification. A program called Distributed Name Services (DNS) maps the domain names onto the IP address. DNS was revolutionary in that it made the Internet accessible to those who were not highly computer savvy and or literate.

The only hurdle in the path of expanding Berners-Lee's concept to a global reach-- and thus developing his web of information *world wide*-- was the networking of hypertext links. Berners-Lee realized that any hypertext tool that would be used for worldwide dissemination would be hindered by the fact that many different platforms were linked to the Internet: PC's, Macintoshes, UNIX, and other simple terminals. Moreover, a variety of desktop publishing methods, such as SGML, Microsoft Word, and LaTex were in practice. This presented a complex problem of interoperability: "Commercial hypertext packages were computer-specific and could not easily take text from other sources; besides, they were far too complicated and involved tedious compiling of text into internal formats to create the final hypertext system" (Longman 1998).

Berners-Lee sought to develop a platform equally simple and effective. Berners-Lee created some simple software to create documents and a simple protocol for retrieving foreign document through hyperlinks. The protocol invented was HTTP,

which stands for Hyper Text Transfer Protocol; and the text format for HTTP was coined HTML, Hyper Text Markup Language. Both the former and latter were easy to create and access, thus encouraging others to build upon and further the technology.

HTML was based on the standards of SGML, derived from in order to develop a framework for defining documents. In essence, HTML is an application of SGML conforming to SGML rules to advocate how a language should be constructed. HTML incorporated SGML's main strength—that being that the language was independent of the browser or viewing software. Many elements of HTML were adopted directly from SGML. Unlike SGML, however, the elements of HTML are predefined so as to ensure simplicity and encourage implementation. An element unique to HTML—and what separates its functionality form SGML—is the anchor element—elements with HREF attribute that enabled global textual cross-referencing.

Berner-Lee's concept intrigued intellectual interest almost instantly. Academics and software engineers worldwide began to take interest in the prospect of the new information utility, quickly developing and expanding its practicality and use.

As a derivative of SGML, theoretically speaking, HTML's strengths are SGML's weaknesses, and vice versa. HTML basically has three major advantages over SGML that correlate to its success. Firstly, the language syntax is limited and flexible. Unlike SGML, there is no "standard" formalization to respect. Ostensibly, minor syntax errors are permitted—browsers are able to interpret improper arrangement and closing of tags. In addition, HTML works without DTDs and SGML declarations—which are complicated and require more than an intuitive understanding of the schema. Secondly,

the tag set is rudimentary and can conceivably be learned very quickly. Lastly, HTML is Open source.

HTML's simplicity also results in various weaknesses. The HTML tag set is fixed and cannot be expanded in order to include document objects important to distinct disciplines. HTML documents not being formulated by DTDs results in a variety of incompatible HTML displays that frequently appear on the web. Though formatting data is controlled by HTML, there is little control of structure of HTML documents. In other words, HTML does not support the designs needed to represent database schemas or object oriented hierarchies. Another glaring weakness of HTML is its ability to support syntax errors, which in effect limits the syntax language. In other words, it makes it difficult to properly handle information, and furthermore foster new technologies or improvements to the syntax language.

Inundated by web sites at such a rapid pace, revisions to HTML were imminent. The World Wide Web Consortium (W3C) was formed to police standards, practice, and research for many areas of the Internet. From 1994 to 1999, the HTML format was revised four times, from HTML 2.0 to its present incarnation HTML 4.1. Important to the new versions was the inclusion of meta-tags—a tool for bibliographic control. Essentially, in a HTML document, meta-tags are indexed by search engines and determine the "meaning" or purpose of the document. Meta-tags are embedded within the HTML document, within the <head> and </head> elements, and are not visible in the document unless one was to access the document's source2. Meta-tags in HTML

2 To view the source of a web page on IE, right click a blank space in the document and the click on "View Source" from the menu.

documents generally take two forms: (1) HTTP-EQUIV; and (2) META-TAGS with NAME.

HTTP-EQUIV meta-tags, as the name suggests, are equivalent to HTTP headers. Their main function is to “control the action of browsers, and may be used to refine the information provided by the actual headers” (*A Dictionary of Meta-data Tags* 2003). One HTTP-EQUIV tag is the “Content-Script-Type” which defines the programming language used by the web page. A web page employing javascript would thus have the following meta-tag appearing in the <head> of the document:

<META HTTP-EQUIV="Content-Script-Type" CONTENT="text/javascript">

Other HTTP-EQUIV meta-tags include when the page expires, what language the text is written in, what content it includes and content style, cookie and cache control, amongst a few others.

Meta-tags with the “NAME” attribute do not correspond to HTTP headers and essentially exist as keywords for search engine spiders to index (**Ramiscal 46)**. These keywords not only represent the description of the document but also the attributes such as author, creator formatter and unofficial copyright information. For example, if this very research paper was published on the web, some of the NAME attributes would appear as such:

<META NAME=”robots” CONTENT=”index,follow”>
<META NAME=”keywords” CONTENT= “Library Science, bibliographic control, bibliographic description, SGML, HTML, XML, MARC”>
<META NAME=”description” CONTENT=”A masters thesis research project on the impact information technology has had on the standards of library cataloging”>
<META NAME=”author” CONTENT=”Brian Rocco”>

Without question, HTML's greatest strength is its simplicity. HTML is easy to learn and use, and directly motivated the Web revolution. Paradoxically however, its greatest strength is also its greatest weaknesses. Its lack of standards and ease of use, academics and purists feared, would result in anarchy: a mess of millions of documents flooding the web without any means of organization. Initially envisioned for scientific and academic means of communication, the simplicity of HTML welcomed all disciplines. Soon enough, the dot com explosion resulted in what the purists feared. The Internet and World Wide Web was being invaded by commercial industries—today over one *billion* web pages exist on the World Wide Web-- Google indexes 1,326,920,000 pages! The algorisms behind powerful search engines such as Google provide impressive search results. Still, unlike its predecessor and inspiration SGML, HTML lacked standard definition and control. Nothing's perfect, though perfection is invariably desired; the purists wanted more from HTML, something as functional but with more structure; something as formulated as SGML but not as complicated—the future of a the semantic web was dependent upon such criteria.

eXtensible Markup Language

The buzzword word—or acronym rather— in the Internet world today is XML. Being labeled as "the next big thing," XML is the next step the Internet and the Web must take to reach its full potential. Along with its associated technologies, the promise of XML "is for both human-machine and machine-machine communications, with XML as the 'lowest-common-denominator' language that all other systems, proprietary or

open, can use" (Mercer 4). In other words, XML permits designers to customize tags, which in effect facilitates the transmission, definition, and interpretation of data between applications and organizations.

Ostensibly, XML combines the complexity and hierarchal capabilities of SGML with the simplicity of HTML. XML is a subset of SGML, and its intention is not to replace HTML, which is commonly mistaken. In fact XML is not a metalanguage like HTML but rather a "toolkit for creating, shaping, and using markup languages" (Ray ix). In actuality, XML eventually will be fully immersed into HTML, and produce a purer version of HTML, XHTML. (Ibid).

But it is important to understand the differences between XML and HTML. Jon Bosak of Sun Microsystems breaks it down in three major respects: (1) Information providers can define new tag and attribute names at will; (2) document structures can be nested to any level of complexity; (3) any XML document can obtain an optional description of its grammar for use by applications that need to perform structural validation (*XML, Java and the Future of the Web* 1997).

Unlike HTML and its graphics and design capability, XML harkens back to a text-based structure and an emphasis on storing and representing information. The operative words are "storing" and "representing" data, and not "presenting" data, which is what made HTML so popular and efficient. XML works by separating style from content, simple in concept but revolutionary in theory. As HTML solely presents the data, XML addresses the context or meaning of the data. For example, the word "bill" can be tagged as a "name, a charge, a paper currency, a proposed law, or the mouth of a bird. Tagging the data enables machine interpretation with great precision" (Hogan

2004). Creating context becomes even more important when working with numbers because numbers have no inherent context. For example, 500 might be a great dollar price for a new entertainment set; but it would be a lousy number for the days required for delivery. Therefore, the tag that puts the 500 in context is vital.

XML tags not only create context, they also, which is equally important, create structure. An XML document should be marked up so as there is no open interpretation as to how the data should be interpreted. Names, order, and the hierarchy of the elements must be unambiguously structured. This has, among others, two major advantages: one being that it enables XML to be easily read by both machines and humans; second, the "purity of the information structure" will not get in the way of format conversions (Ray 12).

Unlike HTML, the syntax of the XML document too must be pure, and meet standardization. As previous discussed, browsers, when reading HTML documents, can easily read over and interpret syntax errors. XML documents, like SGML documents, must be well formed and valid. A "well formed" XML document means that XML tags are properly written: they are nested properly—they strictly adhere to the basic syntax rules. A valid XML document is a more complex concept. Besides being well formed, it must also reference a DTD or schema that supplies the kind of information important for the document's intent (Mercer 49). If the XML document is either not well formed or valid, the XML processor instead of displaying the page will generate an error message indicating that the document is either ill-formed or invalid—and elude th the line number where error occurred.

XML structure is dependent upon either DTDs or XML schemas. Before composing an XML document, one must first either find or build a DTD or schema. DTDs and schemas essentially support XML documents the same way; but schemas are the recommended way to go (Mercer 59). The DTD or XML schema tells XML processors which element and attributes are allowed in your document. There are many industry standard DTDs already developed for use in various disciplines. Frequently authors using industry DTDs will have to add an element or attributes to correspond to their interests. This is why XML is an critical improvement to HTML-- it is extensible, hence the name. Like HTML, XML is composed of element and attributes; but unlike HTML the designer of an XML document can create new elements and attributes without changing the XML standard. Working with a predefined DTD, the author can still add elements and attributes to specifically tailor their needs. In a sense, adding elements to a DTD means the author is creating a new DTD. So in effect, when building a DTD it must be placed within an accessible area so that an application can read the XML and can validate the document against the DTD (Mercer 116).

The main weakness of DTDs is that they are written in the Extended Backus Naur Form (EBNF) (Mercer 86). XML schemas allow the author to compose the XML document using various languages that correspond to the XML format. Writing XML schemas instead of employing a DTD has the following advantages: the processor will only have to understand XML; the schema can be manipulated programmatically; and XML schemas can provide much better support for extensibility, datatypes, and namespaces (Ibid). Namespaces allow one to expand the XML vocabulary. A namespace is a group of element and attribute names (Ray 42). A problem can occur

when a word has different meanings in different context. As stated earlier, the word "bill" has several contextual meanings. In order not to confuse the XML parser, the author specifies which namespace the elements and attribute are derived from. Schemas and DTDs are not required by XML, but they will make the elements and attributes unambiguous.

An important framework created by the W3C to compliment XML is the Resource Description Framework (RDF). An RDF describes a web site's metadata and provides interoperability between applications that exchange information on the Web (W3C "RDF"). RDF uses XML as a syntax to exchange information. A complete analysis of the importance and functionality of an RDF can be read on XML.com in the article "What is RDF" by Tim Gray at http://www.xml.com/pub/a/2001/01/24/rdf.html. As mentioned throughout this research project, the metadata of the web, based on HTML configuration, is anything but controlled. XML in accordance with RDF aims to standardize the representation and collection of metadata.

The Emergence and Importance of XML

On today's web, XML has already reshaped the production and communication of numerous institutions and organizations. In particular, XML has had an enormous impact on both business-to-business (B2B) and business-to-manufacturing practice. By implementing XML, businesses now have "speedier work-order processing, faster inventory reconciliation, and better corporate visibility into manufacturing" (Singh 62). After reading this introduction to XML, it is easy to understand why and how XML is

being eagerly adopted by business institutions. The ability of XML to function independently across international companies with different software and hardware configurations characterizes the World Wide Web much more efficiently than its predecessor, HTML.

Today electronic printing and publishing industries salivate at the cost and production benefits of employing XML. Prior, the exchange and communication of information between printers, publishers, and vendors had to cross multiple platforms. Locating, structuring and formatting documents pre-XML disadvantaged productivity. To alleviate the interoperable dilemma, the publishing community designed XML Book Industry Transaction Standards (XBITS) to create a standardized electronic format. The best known XBITS is DocBooks. As a single source, "DocBook allows authors and editors to rapidly exchange drafts electronically and then, using the same DocBook XML files, immediately compose into a wide variety of final formats" (Poe 14). Indeed, the advantages of interoperability for the publishing community are obvious

In 2002, Amazon.com, the commercial web powerhouse, announced that their service would employ "XML-based services technology to make the contents of its multimillion-item catalog freely available for use by any Web site or software application" (Akin 70). In effect, by converting to XML, the over 30,000 members of the Amazon Web Services developers' program can freely link their applications or websites to Amazon's enormous catalog.

As XML application signals a change in the way web-based information is being processed and shared, numerous other commercial sites worldwide have begun to recognize the power inherent in XML. But perhaps Microsoft Corporation's inclusion of

XML-realted technology pushes the metalanguage mainstream. Office 2003 now enables users to save documents and create schema-based templates with XML support. Powered by XML, any data in a Microsoft Office document now can be calculated without requiring reformatting. Present in corporate, education, and home offices worldwide, the software juggernaut is playing an important role in the standardization of XML, and as a result, the way in which web-based applications exchange information.

By changing the way in which information can be converted and exchanged, XML simplifies the development and management of information. If the implementation of XML syntax and its related technologies continues to grow as expected, the architecture of the World Wide Web will be considerably transformed, affecting all industry: education, business, and government. As CNET news.com stated in an article on their site (accessed 2/14/04) : "The pervasiveness and widespread adoption of XML has in fact changed the economics of technology adoption--while in the past it might have been economically feasible to work in a vendor-proprietary, closed environment, the reverse is now the case."

Summary

Comprising various forms corresponding to available technologies, the creation of metadata has been practiced for centuries. Although the theory and means of metadata creation has drastically evolved from century to century, a few basic fundamental cataloging truths seem to transcend time. Essentially, what has not changed is the concept of metadata, that being data on data—the description and control of-- used to locate a resource. But as technological advancements occur, new theories arise and

challenge existing principals. Today, amidst the digital revolution, the standards of documentation seem to be rolling over as quickly as they are accepted, affecting the way in which the world formats and exchanges information.

CHAPTER THREE

METHODOLOGY

Overview and Restatement of Problem

Frequently, with the rise of a new technology comes the rethinking of established theories. Fixed and productively functional for over forty years, the rules and standards of library bibliographic control and description are presently threatened by academic challenges. The Internet and Worldwide Web has challenged the library milieu to review fundamental cataloging standards and practices. The library community currently faces a period of transition where the reevaluation of the theory and practice of cataloging must be reconsidered, and perhaps even drastically altered.

With the digital revolution progressing at such a rapid pace, the corresponding output of information objects are presently circulating without a standard means of control. Much literature has been written that explores concepts of control and theorizes on standardization. An analysis of the corresponding literature on the theory and practice of the "new" document is vital for the information control. Chapter Three delineates the method and processes to be implemented for the evaluation and selection of resources associated with the reinterpretation of information organization in the library community. This chapter will discuss what research tools will be used; how the selection of resources will be established; and how the data will be collected, recorded, processed and analyzed.

Restatement of Hypotheses and Research Questions

This research project will attempt to prove that information technology has reinvented library cataloging theory and practice; and will examine the following four research questions:

5. How functional are the MARC record format and AACR2 rules for bibliographic description in the contemporary library environment that houses "modern" forms of documents?
6. When migrating records to new standards and schemas, what factors will librarians face in order to enable interoperability?
7. How important is it for librarians to both understand and implement eXtensible Markup Language (XML)?
8. What are the new metadata standards available and appropriate for special collection library cataloging implementation?

Research Methodology

This project will employ the method of content analysis. In the case of this research project, content analysis involves the evaluation of concepts hypothesized by authors of various resources associated with information technology and library cataloging. Content analysis could be used, as Weber states in his monograph on utilizing the research methodology, to research and study various trends, data, and patterns (Weber 9). This research project will perform a content analysis in order to "describe trends in communication content" and "reflect cultural patterns of groups, institutions, or societal attention" related to the impact information technology is having on the theory and practice of library cataloging (Ibid). As quickly as the digital revolution has evolved, so too has the production of literature on the implementation and

surfacing theory of metadata. It is thus important to analyze the "trends in communication" within the library environment. Resultant of the project's four research questions, the resources to be evaluated will undergo a qualitative content analysis by the researcher.

Annotated Bibliography as a Tool for Content Analysis

This project will utilize an annotated bibliography as a tool to perform a content analysis on the research topic. A bibliography is defined as a "list of materials or items usually restricted in some way, such as by subject, form (e.g., periodicals), or coverage (e.g. items published before 1900)" (Rubin 193). As a rule, bibliographies are intended to lead the user to the sources they identify. The main difference between an annotated bibliography and a plain bibliography is that with the former each citation is succeeded by a concise description of the resource to inform the reader of the relevance of the cited source.

The annotated bibliography to be created as a result of this research project will work as a tool of access for librarians and library science students interested in theories of contemporary cataloging; as well as persons interested in metadata, documentation, and the cataloging of computer network resources. This project hypothesizes on the chaotic structure of web sources and the impact it is having on the theory and practice of modern library cataloging. If one were to enter XML as a search term in Google, the result would be over thirty-million hits. In a sense, this annotated bibliography exists as a compass to navigate through the chaos. As a result, the annotated bibliography will provide a concrete content analysis of the resources available for access. The annotations will utilize the following format:

Clarke, Kevin S. "Updating MARC Records witn XMLMARC." *XML in Libraries.* New York: Neal Schuman Publishers, 2002; p3-17.

Discussed in this article is the potential that XML has to exceed the impact MARC records has had on librarianship. The author commentson the fact that librarians unconditionally deal with structured information—structured so as to meet interoperable and international standards of documentation. Clarke states that librarians are in danger of becoming marginalized or segregated from the evolving online resource search tools if they continue to catalog records using the now antiquated MARC record format. Moreover, Clarke argues that the XML can free libraries from the constraintsof library system vendors and create a more flexible, platform independent system for maintaining MARC records. Clarke doesn't immediately suggesta complete conversion form MARC to XML, though he states that in time a complete version will occur. Included in the article is a discussion on MARCUTL—the language used for migration of MARC to XML and a detailed examination of the conversion process.

Selection of Subjects

This research project aims to explore the impact the rise of information technology has had on the theory and practice of cataloging in an academic or special collections library. Therefore, the researcher will limit his scope of research to academic and scholarly journals affiliated with the Library Science discipline; as well as research papers given at conferences sponsored by or associated with the ALA, Library of Congress, ISO, or other established Library Science affiliations associated with technical services.

The researcher obtained access to both CUNY and SUNY Buffalo's OPAC and collective database resources. The following Library of Congress subject headings will be employed when searching both OPACs:

- Cataloging of computer network resources
- Metadata
- Information organization

- Descriptive Cataloging—Rules
- Cataloging—Standards
- HTML (Document Markup Language)
- XML (Document Markup Language)
- SGML (Document Markup Language)
- Libraries—Automation
- Information Storage and Retrieval
- World Wide Web
- Descriptive Cataloging Rules

Similar subject terms will be used when searching the two research databases, Academic Search Primer (Ebsco Host) and Wilson Web, Library Literature and Information Full text. When searching for information on the World Wide Web, resources will only be collected from organizations affiliated with the standardization of metadata: W3C <http://www.w3c.org>; Text Encoding Initiative <http://www.tei-c.org>; Library of Congress <http://www.loc.gov>; Dublin Core Initiative <http://dublincore.org>; and the ISO <http://www.iso.ch/iso/en/ISOOnline.openerpage>

Since the research project is concerned with theory and practice of library cataloging, and the fact that resources on the World Wide Web are suspect, no other avenues of the web will be explored.

The annotated bibliography will comprise ninety-five research articles, as well as five annotations of authoritative meatdata web sites. Articles may appear in a monograph of a collection of articles, or as individual articles form in a journal.

Research Instrument

The Checklist

In order to effectively control the compilation of resources collected for this research project, a checklist was created by the researcher. As a template, the checklist organizes the resources documented. In essence (and ironically enough considering this project's intent), the checklist is actually metadata on the metadata researched. Fields of bibliographic information for each resource in the template include: statement of responsibility; title; publisher; subject heading assigned. Three remaining fields on the checklist are: Source, Article Type, and notes. The most important field is that titled "Article Type." Under this heading will be four check boxes (AACR2/MARC; XML; Migration/Interoperability; Digital/Special collections) for the identification of the article type. A distinction between "types" of articles will provide selection standards for collection. Each article type corresponds or is associated with individual research question. From there, the notes area will dissect the article and its relationship to the research question. The checklist for the researcher will act as a ruler against which items selected will be measured. The complete checklist can found at Appendix A on page X.

Data Collecting and Recording

As stated, a checklist will be used by the researcher to organize and collect the data of resources to be included in the annotated bibliography. The checklist aspires to record all pertinent bibliographic information, as well as how the information pertains to the research topic. The heading on the checklist "Article Type" is the point of reference for classification. The resources will be alphabetized and then dived into two subsections: (1) articles and papers; (2) Internet links.

In order for a resource to be included, a qualitative content analysis of the resource will be conducted by the researcher. For an article to be included in the first subsection, the scholar must address concerns relating to any one of the research questions posed. Subsection two consists of five authoritative and internationally recognized experts associated with the discipline of modern bibliographic cataloging.

The first subsection will contain journal articles and conference papers pertaining to the article type. Provided for each individual resource will be basic bibliographic information: Statement of responsibility, title, publication date, physical format. Following each citation will be a concise description of the resource and how it pertains to the subject heading and research question involved.

The second subsection will provide Internet links to sites with rules for markup and metadata syntax and schema (including examples of sites employing said syntax or metadata), and directions and or specifications needed for use or migration.

Data Analysis

Overall analysis of an individual discussion, proposition or theory contended--and how it pertains to the overall research question-- will not be ascertained until Chapter V.

The resources collocated within each heading corresponds to a specific research question. Individually, each resource will be systematically read and analyzed for interpretation. Trends, patterns, and recommendations proposed and contended for each research question will then be formulated and processed.

Limitations

Due to information technology's proclivity to evolve at such a rapid pace, the research that will be conducted must keep up to date with the latest—and even near future—advances. Throughout the course of the project, it is possible that immediate annotations to the annotation itself will be needed; however major revisions on the overall theories contended is highly unlikely to occur before the project is completed.

Though the researcher has access to only a limited number of database conglomerates, each do indeed comprise a vast majority of library science journals currently published. The project's intent is not to cover the entire "universe" of literature on metadata, but rather to analyze a small group of literature on metadata in the library environment.

CHAPTER IV

ANNOTATED BIBLIOGRAPHY

Scholarly Journal Articles

1 Arms, Caroline R. "Some Observations on Metadata and Digital Libraries." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 245-61.

As a result of the Internet, many varieties of digital collections have been made available through networked connection for particular communities and purposes. The author of this article examines how metadata elements needed to enable users to find, identify, select, and gather resources; often do not match the elements and rules for bibliographic cataloging of materials traditionally held by libraries. The author contends that in order for varied types of resources from varied sources to have coordinated access, a need for a level of commonality among metadata schemes is crucial. The article examines how the American Memory digital project gathered together metadata from heterogeneous sources. In addition, the author examines several initiatives to develop rich structured metadata schemes for specific domains and others that find simple approaches to sustain resource discovery across domains. The author details trends and commonalties, and examines various metadata schemes, including XML, Dublin Core, and the TEI Header.

2 Banerjee, Kyle. "How Does XML Help Libraries?" *Computer in Libraries* 22 no8 Spring 2002 : 34-42.

The author of this article remarks on the excitement XML has generated within the library cataloging community. Stated with a sense of warning, the author contends that XML is already drastically transforming the way in which information is managed and delivered. The article discusses inherent weaknesses present in current cataloging schemas and how, with XML, they can be eradicated. An introduction to XML and the importance of some of its related technologies, i.e. Document Object Model (DOM), and extensible Stylesheet Language (XSL), is detailed. The article further explores the underlining application theories of XML and how it can improve various services for the library environment. With examples of both DOM and XSL, the article champions the simplicity of creation and user-end utilization of XML. At the end of the article is a valid XML document example and a XSL style sheet for view and study.

3 Banerjee, Kyle. "Improving Interlibrary Loan With XML." *XML in Libraries.* New York: New York: Neal Schuman Publishers. 2002: 31-41.

In this article Kyle Banerjee, a librarian at Oregon State Library, discusses the difficulties, both of time and manual labor, when processing requests for interlibrary loans. Using XML, the article discusses how the university revised ILL ASAP, their locally developed program used for interlibrary loan transactions. The project learned that XML was useful for presenting small amounts of structured data, e.g. loan requests. However, the library leaned that XML is both memory and CPU intensive, and results in slow processing. Moreover, the author warns, until the technology of XML matures, considerable trial and error implementation will be required when attempting a software cross-over.

4 Beacom, Matthew. "Crossing the Digital Divide: AACR2 and Unaddressed Problems of Networked Resources." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 135-153.

This article scrutinizes how the advent of the World Wide Web has initiated profound changes in how the library community uses information. Due to the changes the Web has forced upon cataloging, librarians, as the author contends, must reevaluate sources of information and reexamine cataloguing procedures and practice. The author examine four primary concerns associated with how technology has forced librarians to reevaluate cataloging theory. Firstly, the author examines the change in how knowledge is packaged. Secondly, the change in how knowledge is published an disseminated. Thirdly, the change in how access to knowledge is controlled. And lastly, the change in how librarians help others use knowledge before it is packaged, published, and restricted as networked resources. The author examines the adjustments needed to be made to AACR2 in order to meet the needs of new resources engendered by the digital revolution.

5 Bishoff, Liz, and William A. Garrison. "Metadata, Cataloging, Digitization, and Retrieval: Who's Doing What to Whom: The Colorado Digitization Project Experiment." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 377-88.

In order to increase user access to the wide array of resources constituted by Colorado's archives, historical societies, libraries, and museums, the Colorado Digitization Project was created. The project digitized various resources; then provided distribution via the Web, only to discover that there were serious over overlapping of content in the holdings. The new goal of the project was then to find ways to bring together the resources held by widely dispersed cultural

heritage institutions into one virtual collection. The project committee felt that the most important issue to address was the creation of metadata. They decided to develop a union catalog, and to support the process and accommodate local preferences, they developed a set of metadata guidelines (descriptive, functional, and administrative) that did not require the adaptation of one standard such as MARC or EAD. Instead, they established a set of core elements derived form Dublin Core elements, which would support cross-database searching.

6 Bothman, Robert. "Cataloging Electronic Books." *Library Resources & Electronic Services* 48 no1 January 2004: 12-19.

As the rules and standards for cataloging electronic sources continue to change, librarians must continue to educate themselves on the appropriate metadata. This article discusses the electronic record as a unique manifestation, and provides instruction on the application of current cataloging rules. The article explores MARC 21 elements used to describe the e-book, as well as the variable data fields associated with LCCN, ISBN, and general classification. Each element field is fully addressed, with examples. In effect, the author shows how the cataloging of e-books require extra steps beyond those required for a print book. The key points to cataloging an e-book, according to the author, are to cite the source of the title and source of edition statement if it differs from the source of title; state the mode of access or system requirements necessary; and add the extra codes required for the electronic nature of the e-book.

7 Broun, Kevin. "Integrating Internet Contact." *netConnect.* Accessed 21 February 2004. http://www.netconnect.com

This article explores the now mature technology Rich Site Summary (RSS), and argues for its acceptance by and inclusion to the library environment. For the reader, the author first defines RSS and describes it functionality on the Internet today. The author then details the role of RSS in the National Institute Health library. The author discusses how productive the RSS technology has been for intranet services in the library. Of paramount importance for this library, RSS enables information feeding of specific content from specific web hosts. Due to RSS being formatted in standard XML, the author notes how easy it is to be formatted by various software components. In addition, RSS does not require hands-on staff effort. The author comments on the rise of "blogging" in and out of the library environment and what place RSS has in the striving web application. Moreover, the article details issues of parsing RSS documents, publishing RSS content, and vendor software support.

8 Burke, Gerald, and Carol Anne Germain and Mary K. Van Ullen. "URL's in the OPAC: Integrating or Disintegrating Research Libaraies' Catalogs." *The Journal of Academic Librarianship* Vol 29, Number 5 September 2003: 290- 299.

Discussed in this article is the impact Online Public Access Catalogs are having on the catalogues of research libraries. As an authoritative source of access, errors or inaccurate information in research libraries' catalogue, as the author's contend, should not exist. However, the increasing practice of including Web sites in the catalog raises new issues of a catalogue's accuracy and reliability. An analysis of the Association of Research Libraries' (ARL) catalogue examined the rate of errors of catalogued URL's. A survey methodology was conducted by the researchers to discover the frequency of URL errors or inaccuracies in OPACs. The result did tend to vary, but a significant proportion of the researched library catalogs had a significant number of errors. Further discussed by the authors are the impact these errors can have on the reputation of a research library as an authoritative source. In addition, the authors provide some guidelines to follow to avoid URL errors within an OPAC.

9 Calhoun, Karen. "Redesign of Library Workflows: Experimental Models for Electronic Resource Description." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 357- 76.

Examined in this article is the potential for and progress of gradual transition from a highly centralized model for cataloging to an "iterative, collaborative, and broadly distributed model for electronic resource description" (359). The article's intention is to prepare library mangers for some cataloging experiments being conducted in order to help them conceptualize new methods for defining and planning for e-resource description. The article examines how a library could build a coherent system for discovery and retrieval of networked resources. Seeing as the role of metadata plays an important role in the process, especially in enabling cross-functional collaboration, the author calls for the reuniting of functional groups into "virtual teams that can integrate the e-resource description process, speed up operation and improve service" (361). What the author feels is important for the collaborating process to be successful is the integration of labor associated with the cataloging work. She examines the traditional division of labor, and changes that need be addressed.

10 Campbell, Jerry. "Access in a Networked World: Scholars Portal in Context." *Library Trends* Vol. 52, No.2, Fall 2003: 247-255.

The potential of network technology and its impact on the library community is studied in this article. As the author states, parallel with the rise of network speed and capability was the mass production of electronic resources created by proprietary systems, making access across disparate systems near impossible. This article details the scholar portal project, an effort to create a search and retrieval tool to temporarily solve the problem of information exchange between disparate systems. The impact the power of search engines has played on library users search methods is examined. Moreover, the project explores platform

mapping, and how the many faces of metadata, i.e. HTML, XML, Dublin Core. Though these standardizations have slowed downed the proliferation of metadata variations, they have not resolved the problems associated with "islands of disconnected data and information" (251).

11 Caplan, Priscilla. "International Metadata Initiatives: Lessons in Bibliographic Control." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 61-79.

Explored in detail in this article is the proliferation of metadata element sets used for resource description. Specifically, the author examines the following subset of metadata in great detail: the TEI header, EAD, Dublin Core, and VRA Core. The author examines why each of these schemas developed, its major points of difference from traditional AACR2/MARC resource description and library cataloging, and what advantages they offer to their user community. Moreover, the author discusses the challenges for implementers of these schemes, as well as possible future developments. Furthermore, the article delves into identifying some commonalities among the schemes, and attempts to offer some lessons for developers using the particular metadata schemes. It concludes by examining the implications of commercial metadata standards upon traditional library bibliographic cataloging.

12 Carliner, Saul. "Designing Better Documents." *The Information Management Journal,* September/October 2002: 42-50.

This article explores the impact information design is having on the definition of the "document." Before going into the transformation of the document, the author investigates information design and defines it by stating that it "represents a focus on the fusion of content, structure, and appearance of documents" (43). From there, the article explains the ways in which information design practice can improve information assets, including interoperability, security, legal implications, and findability. The article delineates the relationships between information management, information architecture, and information design. Combined, all of these disciplines and standards affect the way in which information designers classify content, which in effect influences the standardization of the document form.

13 Clarke, Kevin S. "Updating MARC Records witn XMLMARC." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: 3-16.

Discussed in this article is the potential that XML has to exceed the impact MARC records has had on librarianship. The author comments on the fact that librarians unconditionally deal with structured information—structured so as to meet interoperable and international standards of documentation. Clarke states that librarians are in danger of becoming marginalized or segregated from the evolving online resource search tools if they continue to catalog records using the

now antiquated MARC record format. Moreover, Clarke argues that the XML can free libraries from the constraints of library system vendors and create a more flexible, platform independent system for maintaining MARC records. Clarke doesn't immediately suggest complete conversion form MARC to XML, though he states that in time a complete version will occur. Included in the article is a discussion on MARCUTL—the language used for migration of MARC to XML and a detailed examination of the conversion process.

14 Chan, Lois Mai. "Exploiting LCSH, LCC, and DDC to Retrieve Networked Resources: Issues and Challenges." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 159-176.

Examined by the author is whether existing tools of vocabulary control for structured bibliographic organization can continue to function when dealing with Web resources. The author contends that a library's OPAC should be viewed as a part of the library infrastructure rather than as a separate entity. What the author examines is how the nature of the Web and characteristics of networked resources affect subject access; and then proceeds to analyze the requirements of effective indexing and retrieval tools. She explores the traditional tools in a networked environment and what adjusts need to be made in order to correlate with the technological advancements; mapping temrs and data form different sources; and integrating subject access tools. Moreover, she examines how these differently affect a Library of Congress Subject Headings, Library of Congress Classification, and Dewey Decimal Classification.

15 Chan, Lois Mai, and Marcia Lei Zeng. "Ensuring Interoperability among Subject Vocabularies and knowledge Organizations Schemes: a Methodological Analysis." *IFLA* J 28 no5/6 2002 : 323- 328.

The authors of this article examine how information resources on the web are indexed with different vocabularies and organized under the auspices of different metadata schemes. The authors, using a methodological analysis approach, researched the ways in how to retrieve the best results in cross-domain searching. Before discusses the outcomes of their own research, the article explores the concern for vocabulary compatibility and investigates methods previously used for achieving and improving interoperability. The article details the process of setting up an interoperable system—from mapping vocabularies, to link storage and management, to specific institutional concerns such as integrating the views of different cultures.

16 Chandler, Rob and Karen Anderson. "An XML DTD for Subject Related Resources." *Australian Academic and Research Libraries,* 33 no3 Summer 2002: 189-201.

This article was written to support the project completed by the authors on the analysis and definition of a XML document validation template used to describe the data elements that are to be supported by the resources section of the Information and Knowledge Management Research Group's (IKMRG) website. The resources analyzed consisted of annotated URL's and stored documents. A second object of the research was to produce a metadata validation template for the website by researching for the analysis and definition of an XML DTD (Document Type Definition) suitable for their collection. The article develops an understanding of XML and of the process involved in the analysis of metadata related to electronic document archiving and retrieval systems. The article goes into a thorough discussion of implementing XML and some of its related technology.

17 Chen, Chao-chen, et al. "The Design of Metadata for the Digital Museum Initiative in Taiwan." *Online Information Review*, Volume 26 no5 2002: 295- 306.

This article examines the critical role metadata has on a digital project. On a generic level, the article first explores issues concerned with metadata creation, with respect to analyzing attributes of the collection, as well as understanding the users information needs and information seeking behavior. Next, the articles details the steps, both the success and failures, of the digital project. In particular, it describes the importance and difficulties of interoperability that arise when dealing with disparate language and platform issues. Developed under the Digital Museum Initiative and sponsored by the National Science Council of Taiwan, the paper examines the development of a Chinese metadata, an XML/metadata management system called Metalogy. The article provides an in depth analysis of the procedures undertaken for the design of appropriate metadata suitable for the project.

18 Chepesiuk, Ron. "Organizing the Internet: Yhe 'Core of the Challenge." *American Libraries,* Vol. 30 Issue 1 Jan99: 60- 64.

Explored in this article is the internet metadata scheme, Dublin Core, and what effect it will have on the library environment. In particular, the article investigates how the Dublin Core data elements can work together with the MARC record format. As the author expresses, beneficial to implementing Dublin Core elements for cataloging resources is the fact that there are no set of "rules" that limit a cataloguer in the age of multimedia. The author discusses how practical Dublin Core elements are for cataloging electronic resources—something MARC has difficulty describing and controlling. The author contends that the flexibility of the Dublin Core elements will allow for adjustments within specialized communities with their own needs. Furthermore, the author posits that Dublin Core is not going to endanger the MARC record format; but rather MARC will continue to be valuable to the cataloging community as a template or framework for metadata collection.

19 Connell, Tschera Harkness, and Chandra Prabha. "Characteristics of Resources Represented in the OCLC CORC Database." *Library Resources & Technical Services*, 46 no2 Ap 2002: 39-49.

With more and more libraries providing access to Web resources through OCLC's Cooperative Online Resource Catalog (CORC)—and thus by the extension of the WorldCat database—the ability to provide a database at its maximum potential highly depends upon if the construction of the database meets interoperability standards. The authors of this article examines the characteristics of Web resources in CORC in terms of their subject matter, the source of the content, publication patterns, and the units of information chosen for representation in the database. The study examines 414 resources from a sample that belonged to the social science discipline of an academic library. As a result of the study, the authors concluded that, by following AACR2 standards, definitions to aid in the handling of Web resources are much needed so clearer specifications can be developed for databases representing resources described by AACR2.

20 Cree, John Scott. "Cataloguers—Empowering the User?" *Cat Index*, no 149 Autumn 2003: 11-16.

The author of this article examines the active role the cataloguer, and his/her corresponding technologies, has on the user. The author begins by presenting an overview of a history of cataloging practice and how it affected the user. This provides a historical overview for the reader and conveniently enough leads the discussion to contemporary cataloging practices and how the resources affect the theory, and how in turn the users influence theory. An examination of system migration, electronic documents, and system subject analysis, and how each is affected by technological advancements is investigated. Importantly, how the cataloguers places the users needs before the institution's concerns is critical for the creation of a friendly and constructive catalog. The future of cataloging, in particular the impact electronic documents is playing, is then discussed and left open for future considerations and developments.

21 Crowston, Kevin and Barbara Kwasnik. "Can Document-genre Metadata Improve Information Access to Large Digital Collections?" *Library Trends* Vol 52 no. 2 Fall 2003: 345-361.

Examined in this article are the problems of information-retrieval, and how the bibliographic concept of document genre can be of assistance. The author contends that the need for explicit identification of genre seems particularly important for large digital collections because any search usually retrieves documents with a "diversity of genres that are undifferentiated by obvious clues as to their identity" (345). Moreover, due to the fact that most genres are

characterized by both form and purpose, identifying the genre of a document offers information as to the document's purpose and how it correlates to the user's situation. The article delineates the possible role of genre identification in the information-retrieval process. The authors propound that genre identification would enhance searching because they feel the topic alone is not enough to define an information problem; and moreover, search results containing genre information would be easily understandable. The authors then proceeds to address the issues of studying the effectiveness of identifying genre in large digital collections.

22 Danskin, Alan. "Today MARC Harmonisation, Tomorrow the World Wide Web: UKMARC, MARC21, XML and ONIX." *Catalogue &Index* no143 Spr 2002: 4-11.

This essay discusses the impact that the decision by the British Library to adopt MARC 21 in reference to UKMARC, and implications the decision has on the theory of network cataloging. It explores why the British Library decided to switch to MARC 21 now rather than twenty or even ten years earlier. The benefits of co-operative cataloguing for the British Library's implementation of MARC21 are analyzed. In particular, the essay examines the two key problems the British library faced before converting to MARC21: authority control and conversion between formats. Lastly, the article examines the impact ONIX, the publisher and bookseller production subscription format, will have on the library community. The author feels that it "could potentially mean that some of the descriptive component of cataloging is 'outsourced' to the book trade, allowing libraries to focus on the value added intellectual processes of subject indexing and collocation" (10).

23 Delsey, Tom. "The Library Catalogue in a Networked Environment." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 43-55.

In this article Delsey explores how the development of Web technology has enabled libraries for the first time to link records in their online catalogues directly to the digital resources described. The article provides an overview of how technology has changed the relationships between the library catalog, the catalog user, alternative sources of bibliographic data, and the resources described in the catalog. The paper looks at the catalog's role from a technical vantage point, and how libraries can support various interfaces to the catalog. Specifically, the author examines the impact of technology on the interface; the user interface; the resource interface; the abstract/index interface; and the union catalog interface. From there an examination of reassessing data requirements and a reexamination of the interfaces, and the repositioning of the catalog is explored.

24 Dillon, Martin. "Metadata for Web Resources: How Metadata Works on the Web." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 1-19.

Discussed in this article are the various meanings of metadata, both on and off the Web, as well as the various uses in which metadata has been deployed. The majority of the paper focuses on the Web and the roles metadata has on the Internet. Specifically, the paper discusses metadata used in resource discovery. The paper examines both the Dublin Core data elements benefits and weaknesses. The question to use or adapt either the MARC/AACR2 or the Dublin core schema is examined. In addition, the rise and potential benefits of the XML schema are examined. The concepts of authority control and selection in relation to web resources is fully explored. Dillion proceeds by stating that the future of the semantic Web is dependent on the realization, by various professions and institutions, of the regulation of metadata schemas used for both traditional and digital information objects.

25 Dorman, David. "Technically Speaking." *American Libraries*, Vo. 2 no 167, August 2002: 84-88.

Highlighting the developments in library technology, this article focuses on the importance and impact IFLA's newest cataloging standards, FBRB, is having on the theory of contemporary library cataloging. The article explores how FRBR extends Lubetsky's—contributing author of AACR2 and leading theoretician of cataloging in the 20th century—idea of the Works as a basic unit of cataloging. The FRBR standard defines three bibliographic levels: Work, Expression, and Manifestation, all in which can have its own MARC record. FBRB helps eradicate the problem of searching through various MARC record for basically the same edition of a work. With FRBR, all these records still exist, but are now linked to a higher level MARC record, the Expression, and thus enables a more exact and proficient retrieval process. The article also discusses the growth of a few other library services due to information technology, including virtual reference.

26 Downing, Thomas A. "An Initial Survey and Description of How Selected United States Government Libraries, Information centers, and Information Services provide Public Access to Information Via the Internet." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 335-42.

This article examines how the United States Government agencies have created innovative Internet applications, providing public access to online information services. The author issued a survey to selected agencies to discover that most agencies provide multiple services to meet differing approaches for retrieving

information. The survey also revealed that many information providers not only informed users, but entertained via use of the rapidly evolving innovative applications. Downing, however, feels it is not clear as to what degree United States Government agencies ensure permanent access to online information published at these sites. Downing relates how his findings can both hamper and progress an academic library OPAC. He discusses how the use of innovative applications added to an OPAC can be user-friendly and inviting; however, he warns of the impractical cost of software, hardware and technical support needed to install and maintain.

27 Dunshire, Gordon. "Joined Up Indexes: Interoperability Issues in Z39.50" *International Cataloging Bibliographic Control* 32 no3 J1/S 2003: 46- 51.

Based on the findings of Co-operative Academic Information Retrieval Networks for Scotland and Scottish Networks (CAIRNS) and SCONE, this paper discusses issues in the interoperability of indexes to metadata records in distributed information networks. The CAIRNS Cataloging Issues Working Group identified a number of factors affecting cross-searching of metadata indexes authors, titles, subjects, and control numbers, including local cataloging policies, content standards, and index structures. The SCONE project investigated issues in subject indexing at the collection level, in particular the relationship collections with specific subject content and general collections where subject mappings are appropriate. The author then examines the results of these findings in a cross-domain context and discusses technical difficulties that mat incur, and practical solutions for recovery.

28 Ellero, Nadine. "Panning for Gold: Utility of the World Wide Web for Metadata and Authority Control in Special Collections." *Library Resources & Technical Services* 46 no3 July 2002: 79-91.

Described in this article is the use of the World Wide Web as an important name authority resource and tool for a special collection's analytic-level cataloging for the discovery of the names of people deceased and alive. The author found that current tools and initiatives such as the Name Authority Component of the Program for Cooperative Cataloging (NACO) and the Library of Congress Name Authority File are useful. The article examines the digitization project undertaken by the Historical Collections and Services Department of The Claude Moore Health Sciences Library. The project created authority data for names with information from resources found on the web. Detailed are the steps the project took in acquiring specific personal data, and how to match it with a proper authority file. Looking further into the future, the author concedes that future projects similar to the one they conducted will be benefit greatly with the help of XML and RDF implementation.

29 Fichter, Darlene, and Frank Cervone. "Documents, Data, Information Retrieval, & XML." *The Information Management Journal*, October/November 2002: 22-40.

This article communicates how the irregular structure of documents has resulted in the fundamental problem of information retrieval—that being the inconsistency of search results. The author delves into the benefits XML has on the structuring of documents—how it separates style form content, something its markup language predecessor HTML lacked. Of paramount importance to the implementation of XML is the creation or application of DTDs; this article fully explores how the ability to create implicit data elements will be extremely beneficial to myriad library collections. Lastly, the author comments on how HTML will, for the near future, remain as the markup language used for the construction of library Web pages. However, the author contends that if libraries do not reformat regular HTML pages into XHTML, a risk of data migration in the near future is imminent. Fitcher contends that not only will the XML format eradicate the difficulties encountered when attempting to catalog improperly structured documents, but it will hwlp foster an assortment of difficulties of data structuring.

30 Fichter, Darlene. "Migrating Native Law Cases form Html to Xml." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: 135- 148.

For the University of Saskatchewan Library, Darlene Fichter was in charge of a project to better the access and delivery of library materials. The project decided to convert HTML documents to XML format. The project team felt that a conversion to XML would "ensure the utility and longevity of the digital repository" (138). The project's first step for conversion was to discover if they needed to develop their own DTD, which indeed was needed, and a custom DTD was create. The process of creating a custom DTD was more time-consuming and challenging than initially expected. From their the mapping of HTML elements to XML was undertaken, and is thoroughly detailed in the article. The project discovered that HTML files with consistent coding and a range of elements can easily be converted to XML. Imperative for a successful conversion however, the project learned that a programmer to parse the HTML documents would save a considerable amount of time.

31 Franklin, Rosemary Aud. "Re-inventing Subject Access for the Semantic Web." Accessed February 21, 2004: http://www.emeraldinsight.com/1468-4527.htm

As the digital revolution exploded and the production of Web pages proliferated, a lack of authority control over documentation was glaring. This article explores the beginning models of an authority system with subject access tailored after

library science principals of bibliographic control and cataloging. The author fully investigates the utilization of Dublin Core metadata standards. Employing the Dublin Core metadata schema, according to the author, enables "semantic and syntactic relationships to be defined" (94). Moreover, the author comments on how, with the use of a controlled vocabulary, such as Library of Congress Subject Headings, can add to the descriptive facets of relating ideas. Following standardization such as this, the author contends, will ease information discovery and the filtering of data that lacks authority control. This system design enables scalability and extensibility, something that the present infrastructure lacks.

32 Gil, Ester, and Paul Ross. "Database-Generated Web Pages for Managing Electronic Resources and Beyond." *Colorado Libraries,* 29 no4 Winter 2004: 33-42.

This article examines an article written by the authors two years previous, in which recommendations were made for developing database-generated pages for management of subscription based databases. This article's aim is to present the changes made by the University of Denver Penrose Library in managing electronic resources; and to see if the previous article's recommendations were followed. As it turns out, the recommendations of the previous article were indeed followed, and proved not only to be helpful for the previous database project; but it enabled the database designers more leverage to implement the future adjustments the author's recommended. Also included in the article is the implementation of a complicated database to manage research guides. Provided by the author are several software independent strategies for the design and construction of effective bibliographic databases.

33 Gourley, Don. "Integrating Systems with XML-based Web Services." *XML in Libraries.* New York: Neal Schuman Publishers, 2002 : 181- 195.

In the essay Don Gourley explores how XML can be used to combine together various mechanisms of a system to enable well formed, flexible user services. Seeing that a conversion to XML would result in standardization portability, and longevity of documents—and moreover ensure system integration and interoperability—under the guidance of Gourley, the Washington Research Library Consortium embarked on a project to integrate independent systems with XML-based Web services. This essay details the projects goals, i.e. providing network-accessible interfaces to share ALADIN services that make ALADIN patron information available to other applications; and describes how these goals were attained. What is notable about this project is their use of Sample Object Access Protocol (SOAP) for XML messaging. A detailed examination of SOAP implementation is discussed.

34 Graham, Rumi Y. "Subject No-hits Searches in an Academic Library Online Catalog: An Exploration of Two Potential Ameliorations." *College Research Libraries* 65 no1 January 2004: 36-56.

This article illustrates a study that explored ways in which users' subject-searching problems in a local online catalog can be minimized. For the study, the author weekly examined catalog transaction logs in order to identify topics of subject searches retrieving no records, which had no applicable information resources. For the topics thus discovered, the author explored two potential ameliorations of the no-hits search results through use of authority record cross-reference and "pathfinder" records which provided terse instructions on search refinement. The author describes the outcomes of the study, possible concerns regarding the methods used, delineates further steps needed to determine whether the ameliorations made a difference to users' search experience, and suggests related areas for additional research needed.

35 Greenberg, Jane. "A Comparison of Web Resource Access Experiments: Planning for the New Millennium." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 343-56.

This article examines how over the past few years the bibliographic community has initiated a series of experiments aimed to improve access to the increasing number of valuable information resources that are being placed on the Web. The author explores the five leading experiments in this field: OCLC's CORC project; UKOLN's BIBLINK; ROADS; DESIRE projects; and the Nordic metadata project. The research employed a multi-case study methodology with five evaluation criteria to be examined: organizational structure; reception; duration; application of computing technology; and use of human resources. The author proceeds to define the Web resource access experimentation environment, reviews the study's research methodology, and discusses key, important discoveries. In conclusion, the paper creates a strategic plan and invites contributions and suggestions on the framework.

36 Highsmith, Anne, and *et al.* "MARC It Your Way: MARC.pm." *Information Technology and Libraries* 21 no1 March 2003: 19-27.

MARC.pm is written PERL for the reading, manipulating and writing of USMARC data. This article examines MARC.pm, a piece of open source software developed by librarians for librarians. A description of exactly what MARC.pm is and a series of descriptive pieces penned by librarians who have employed MARC.pm is offered. Included in these discussions are the process of

editing electronic resource holding; working with netLibrary records; using MARC data in a web application; and the Joint Admisnistered Knowledge Environment (jake), fused with the MARC format. Each discussion dissects the project undertaken, and provides examples of code structure. In addition, the article offers various URL's corresponding to the PERL language and its role in formatting the MARC.pm language structure. Though PERL is somewhat a challenging computer language, the author concedes that it can be extremely resourceful for librarians and the library community.

37 Huthwaite, Ann. "AACR2 and Its Place in the Digital World: Near-Term Solutions and Long-Term Direction." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 221-235

This paper reviews the developments that have challenged certain principals of AACR2 that have arisen since the proliferation of electronic documentation. In particular, the notion of the idea of "resource in hand"—the physical object—is considered in appropriate for cataloguing remote access electronic resources. The concerns expressed by the Joint Steering Committee for Revision of AACR (JSC) and the International Conference on the Principles and Futrure Development of AACR2 are examined by the author. He then further examines the progress on theses developments, including the outcomes of the JSC meeting. The paper also examines various possibilities for the long-term direction of AACR; and the relationship between AACR2 and metadata schemes. It determines that both sets of metadata standards will continue to play a role, and that AACR will even continue to be used for electronic resources of lasting value. However, the author recognizes the implications of

38 Hunter, Jane L. "A Survey of Metadata Research for Organizing the Web." *Library Trends,* Vol. 52, No. 2, Fall 2003: 318-344.

As a comprehensive survey that researches how the many different forms of metadata can contribute to the organization of the web, this essay is an exceptional introduction to the vastly expanding field of digital information control. The article provides an overview of the key metadata research issues. It then proceeds to examine current projects that are exploring new methods, and the developing new technologies that aim to improve the process of discovery, access, and retrieval of information on the internet through metadata. The article discusses, in lengthy, informative detail, key research areas such as XML; web services; semantic web technology; metadata harvesting; multimedia metadata; rights metadata; automatic metadata extraction; personalization/customization; search engines; broadband networks; mobile and wireless technologies; authentication; annotation systems; and preservation metadata. The benefits and difficulties of each schema is discussed.

39 Jacso, Peter. "XML and Digital Librarians." *Computer in Libraries* 22 no8 S 2002: 34-39.

In this article Peter Jasco contends that XML is the ideal vehicle for resource discovery, information retrieval, and information exchange on the Web. After postulating that idea, the author examines what role and place XML can have in the library environment. According to Jacso, XML is not here to replace the MARC record format; but rather to enhance the staple schema of library cataloging. The flexibility of XML, he argues, can add to and foster the inadequacies of MARC. Furthermore, the article discusses the cataloging practices of digital libraries and the implications XML can have on the Dublin Core metadata elements that are commonly employed by virtual library collections. The article fully discusses the ramifications of XML on both schemas, and presents images of an original MARC record next to an XML reformat, as well as a Dublin Core record corresponded with its XML conversion. Moreover, for XML conversion, the article presents some open source resources that can be used for experimentation.

40 Jonsson, Gunilla. "The Basis for a Record: in the Light of Functional Requierments for Bibliographic Records." *IFLA* January 29 no 1 2003: 41-48.

Due to the intensification of revisions to various cataloging codes, the author of this article examines two main impetuses she feels is spurring the rethinking of the theory and practice of library cataloging. Firstly she feels that due to global networking, today's cataloging environment must be integrated on a global level. And secondly, the emergence of electronic publishing has created a wide array of problems that cataloguers previously had never encountered. What the author feels is important is the need to apply a common conceptual framework to the cataloging art, which, in the author's opinion, the Functional Requirement of Bibliographic Records (FRBR) is beginning to address. Backed by FRBR's standards, the author examines three cataloguing codes, AACR2, RICA (the Italian rules), and RAK (the German rules), and what impact electronic resources is having on the content-carrier dichotomy and the concept of different manifestations of resources.

41 Kaplan, Michael. "Exploring Partnerships: What Can Producers and vendors Provide?" *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 389-413.

This article examines the impact electronic resources are having on the theory and practice of library cataloging. Of importance is the evolution of metadata standards that have arisen with electronic resources. The author explores

"metadata marriage' and ways in which schemas can be melded together for resource description. The author examines metadata for electronic books, and difficulties with ordering form vendors such as netLibrary. He examines a new technology adjunct for the Internet and Web data collection created by Herbert van de Sompel and colleagues at the University of Ghent. The technology, called SFX for special effects, the author claims will be revolutionary for research done on the Web. SFX is a framework for context-sensitive linking between Web resources and is the means to uniting and linking disparate electronic resources to related subjects. The framework has the potential to expand library catalogues and integrate them with surrounding disciplines.

42 Khursid, Zahiruddin. "The Impact of Information Technology on Job Requirements and Qualifications for Cataloguers." *Information Technology Libraries,* 22 no1 March 2003: 18-26.

This article discusses how information technology is affecting a library system's computer hardware and software, CD-Rom, internet, along with other domains such as MARC 21, CORC, and the other rising metadata schemes—all of which are changing the job functions of cataloguers. The author explores how libraries are now coming up with new requirements for a position as a cataloguer. The article's aim is to review job advertisements published in *American Libraries* and *College and Research Libraries,* and assess the impact the use of information technology in libraries is having on job qualifications for cataloguers. The article examines cataloging position titles with relation to degrees required and required skills; and traces trends of requirements and "highly desired qualifications" expressed by various institutions. In conclusion, the author contends that his research proves that the preparation of cataloguers for today's environment requires both library educators and administrators to revamp library cataloging courses and training programs.

43 Kim, Kyung-Sum. "Recent Work in Cataloging and Classification, 2000-2002." *Library Resources & Technical Services,* 47 no1 J1, 2003: 96- 108.

Provided by this article is a review of publication on cataloging and classification that has appeared during the two year span of 2000- 2002. Reflected in the coverage is the impact electronic sources are having on cataloging theory and practice. In addition, another characteristic of the paper is its international scope. It examines the impact "the connecting bridge" of the internet protocol. Part one of the paper focuses on the following categories: descriptive cataloging, authority control, classification, subject cataloging, cataloging non book material, electronic resources and metadata, and lastly, international cooperation. The second part of the paper switches the focus to concentrate on other cataloging-related issues, such as management, education and training.

44 Lagoze, Carl. "Business Unusual: How "Event-Awareness" May Breathe Life Into the Catalog." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 269-286.

The author examines how new digital context has created a dramatically new context than that which was handled for years by the traditional cataloging model. What the author explores is the fact that the traditional catalog depended on relatively fixed resources delivered by a unwavering set of role players, i.e., publishers, authors, and information intermediaries; and the new digital context is characterized by "fluidity in both content and those who provide it." The author proposes new roles for the catalog based on the digital context and a data model that meets the needs of these new roles. As the author contends: "An 'event-aware' model of cataloging, one that recognizes digital resources as inherently dynamic, will allow the research library to adapt to the realities of the digital millennium."

45 Larson, Carolyn and Linda Arret. "Descriptive Resource Needs from the Reference Perspective: Report on a Survey of U.S. Reference Librarians." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 287-310.

The authors discuss the results of a survey they conducted and address various topics from the perspective of research providers. For instance, the most advantageous levels of library and metadata descriptions, using both full MARC records and structured Dublin Core records for content retrieval of Web-based resources; descriptive needs for professional reference, i.e., more subject data, more summary information; additional descriptive elements to facilitate user requests; examination of the inclusion of traditional concepts such as authority files, uniform titles, specialized thesauri; and problems that "arise as reference providers navigate the current 'continuum' of resource discovery form catalog through 'middleware' tools (such as pathfinders, finding aids, abstracting and indexing services, and databases) to content" (287). The authors contend that due to the mass production of web document and information objects, the importance of a library staff learning and implementing the Dublin Core data elements is imperative.

46 Lewis, Walter, Gail Richardson, and Geoff Cannon. "Expediting the Work of the Indexer with XML." *XML in Libraries.* New York: Neal Schuman Publishers, 2002 : 77- 86.

Indexing electronic documents with XML technology is discussed in this article. XML technology enables an indexer to contextualize names and phrases into basic categories such as people, places, and things. The project used the Text Encoding Initiative to encode the electronic documents that were to be indexed. Three XML tools, XMLAttributes, XMLNames.pl, and Index Extraction, all written in Perl, were then used to manipulate and retrieve the encoded texts. The article goes into depth on how they employed each of these tools. Among other advantages, the project discovered that this method was essential for indexing documents that were imaged rather than transcribed. Furthermore, the sharing of indexes between institutions is facilitated by an XML-structure.

47 Library of Tasmania. "E Initiatives and Networked Infrastructure." *Australian Academic Resources Library* 34 no4 D 2003: 292- 309.

Being an area of expansion and continuous development for the National library of Australia, the library of Tasmania, with contributions from members of the CASL Digital Issues Working Center, examined the delivery of library and information services in the electronic environment. The group sought to examine the identification of resources and services where access and availability can be enhanced by electronic delivery; and where cooperation and collaboration will increase the range of library and information resources available online to the public. The article provides a summary of areas of electronic library service in Australia, and comments on the problems and successful advances made in their digitalization and digital assets management and the delivery of the services via the Web.

48 Ludwig, Mark. "Breaking Through the Invisible Web." *Library Net Connect*, Winter 2003 : 8-12.

The author of this article discusses the fact that more and more web content—including library content—is irretrievable. This is increasingly a resultant of the use of proprietary database interfaces where the data cannot be found by popular Internet search engines. The author urges for librarians to bring the "invisible web to the surface so that our public can discover our resources through a query on a popular search engine" (8). In order to accomplish this, the author suggests several strategies. Firstly, the author states that library environment must recognize the technological paradigm shift—in other words, libraries should construct their catalogs with web-based technology—with a simple user friendly interface sans any library "jargon—rather than with integrated library systems (ILS). Next, the author discusses the importance of HTML meta-tags and attracting search engine spiders. Furthermore, the article discusses future implications to consider for a web-based library catalog, i.e. moving beyond proprietary relational databases, and recognize the flexibility and power of XML technology.

49 Luther, Judith. "Trumping Google: Metasearching's Promise." *Library Journal* (1976) 128 NO16 o1 2003: 36-39.

This article discusses the impact Google and its powerful algorithm search capability is having on library patrons, and in effect, the library community. In regards to the patrons, the author contends that Google's power and ease of use has changed their expectations when seeking information. Moreover, Google is free of library terminology, which can be intimidating to a majority of library users. The author contends that metasearch, a software that allows the users to enter simple keywords in a simple interface and retrieve resources from multiple full-text and bibliographic databases at once, needs to be added to library OPACs. The article details how metasearch works, on both the front and back end. Also included is a discussion on the root of librarian resistance to metasearch; as well as the impact it can have on proprietary database publishing houses; and future implications metasearch can have on the Web and library environment.

50 Mann, Thomas. "Is Precoordination Unnecessary in LCSH? Are Web Sites More Important to Catalog than Books? A Reference Librarian's Thoughts on the Future of Bibliographic Crontol." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 87-134.

This article examines the reasons why the precoordination of LCSH subject headings, both in the LCSH thesaurus and in OPAC browse displays, continues to be necessary. In particular, the author examines how LCSH headings depend on their word order in ways that cannot be captured by "postcoordinate Boolean combinations or by word proximity searches that drop relational prepositions as stop words" (112). Also, how a vast network of linkages between LCSH headings and the LCC classification scheme depends on precoordination. In addition to how displays of precoordinated strings enable researchers to recognize entire collections of relevant options that previous to precoordination could not recognize combinations. And lastly, the author examines how the precoordination of terms is unquestionably linked to a wide network of cross-references that would vanish without it.

51 Mathias, Eileen. "Anatomy of a Digitization Project." *Library Journal (1976)* Winter 2004: 2-7.

The author of this article describes the digitization project conducted at the Ewell Stewart Library of the Natural Sciences in Philadelphia, Pennsylvania. Examined in the article are the issues the library faced in undertaking the project, the decisions made, both the successful and unsuccessful facets of the project, and the future plan's of the library. Before starting any digital project, the author contends, it is of utmost importance to determine who the audience will be. From there, choices must be made on software and hardware components before

deciding on what appropriate metadata scheme to employ. This project found that XML with ContentDM to be an appropriate combination fr their collection. Once all the underlining facets are together, a plan for migration of the prior encoded data must then be comprised. The remainder of the article goes into a lengthy, and informative, detailing of their project and "bringing it to the web" (6). Included in the discussion are various components of digital imaging discipline.

52 Mathias, Eileen C. "Using a Web OPAC to Deliver Digital Collections." *Online Information Review* Volime 27 Number 1 2003: 28-36.

The Ewell Sale Stewart Library of the Academy of Natural Sciences completed a major digital imaging project and the author of the article describes the project, choice that were considered for Web delivery of text and images, as well as reasons for choosing the image management function Innovative interfaces. Included in the article is a description of the data entry process as well as a review of the Millennium Media management project; in addition, an analysis of the evolving image metadata standards. The article details all the software components, hardware necessities, and markup languages employed. The difficulties associated with which platform or software to choose is covered. Furthermore, suggestions, recommendations and warnings for particular platforms and protocols are stated.

53 Mayes, Bessie. "Beyond MARC: New Trends for the Library of the Future." *OLA Quarterly* 9 no1 Spr 2003: 2-8.

Contended in this article is the importance of the library profession to keep pace with the new technologies created for public use by industries across the globe. The author argues that libraries must adapt to changing trends in order to assure its place in the evolving world of information science. Indeed, the author feels that new technologies have, and will continue, to enforce the library community to reevaluate cataloging practice and theory. The article examines MARC's inherent weaknesses engendered by the digital revolution, and offers suggestions for inclusion of XML within the MARC-based catalog coding. In conclusion, the article propounds the importance role XML is beginning to play in the library environment; and moreover, how essential the implementation of XML will be for the future.

54 McCallum, Sally. "Extending MARC for Bibliographic Control in the Web Environment: Challenges and Alternatives." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 245-61.

This paper questions if the MARC record format can accommodate changes to AACR2 and developments in alternative bibliographic control tools like Dublin Core, XML, and RDF. The author examines MARC 21 as a proven standard for

application with library OPAC's and WebPAcs worldwide. What the author finds problematic is MARC'S interoperability or compatibility of the data it carries in today's digital environment. The paper "deconstructs the MARC format" as well as the Dublin Core element, XML, and Resource Description Framework (RDF), separating structural issues from content-driven issues. From there, the author examines the pressures from new types of digital resources , the library community's response to these pressures, and future transformations that may take place. Include at the end of the article is a concise table comparison of the basic resource description metadata fields of HTML, Dublin Core, MARC, TEI, ISO 12083.

55 McDonough, Jerome. "Encoding Digital Objects with METS." *XML in Libraries.* New York: Neal Schuman Publishers, 2002 : 167- 80.

METS, the Digital Library Federation's Metadata Encoding and Transfer Standard, is an initiative that is aiming to develop interoperability standards based on XML. The author of this essay explains how to use METS to create digital objects that can be shared easily between disparate systems. The project's goals were to support encoding of both audio and video material; allow for use of various descriptive and administrative metadata sets; and provide the ability to link between portion of a document and to other documents. The essay discusses the METS schema in detail, delineating the components of the document format. The project proves that by using METS with Metadata Object Description Schema (MODS) allows for the opportunity for the library environment to move beyond MARC's inherent constraints.

56 Medeiros, Norm. "Making Room for MARC in a Dublin Core World." *Online* 23 no 6 n/d 1999: 57-60.

This article examines the role the Dublin Core Initiative and its data elements are playing on metadata implementers. As the article states, although Dublin Core can be used to describe materials in traditional formats, its main impact is its ability to adequately describe Internet resources. The article touches upon the heated debate within the library community of Dublin Core replacing the MARC record format. Due to the enormous amount of records already created in the MARC format, Medeiros feels that the traditional bibliographic structure is greatly limiting a collection and needs revision; this is not to say that Dublin Core will fully replace MARC soon. However, a need to establish a superior means of online access has drawn a considerable amount of attention to incorporating the Dublin Core elements into the library catalog. The author explores the benefits of the library community using Dublin Core data elements in their cataloging practice. As the author's research concludes, combining MARC and Dublin Core is greater than the sum of their parts.

57 Missingham, Roxanne. "Portal Down Under: Discovery in the Digital Age." *EContent*23 no 2 April/May 2000: 40- 48.

Examined in this article is the fact that with millions of web pages available for access, the challenge to make information readily accessible through a controlled catalogue needs to be professionally addressed by the library community. As search engines provide millions of hits, finding the right resource can be a difficult task indeed. Explored in this article are options for controlling information portals to create user-friendly and authoritative web resource search results. The project undertaken for this article based their results on international standards of bibliographic control, and felt that the Dublin Core metadata elements were the ideal metadata standard to implement. The information portals created aimed to provide convenient and effective access to information through a single gateway; the results, successes and failures of the project are examined, with a detail of the lessons the project learned.

58 Moen, Willaim E. "Interoperability for Information Access: Technical Standards and policy Considerations." *The Journal of Academic Librarianship*26 no2 March 2000: 129- 136.

This article discusses how technical standards affect the flow and processing of information, especially in today's network environment where interconnection and information exchange between systems requires a broad range of standards. Exemplified by the author are the following standards inherent into today's environment: network interfaces, operating systems, standards defining data formats, storage devices, and computer-to-computer communications protocols. The article then addresses libraries and their standards and the challenges of networked environment standards. In order to explore how each influences the other, the article looks at two issues; firstly, who develops the standards that affect critical library functions, especially information access; and secondly, the evolutionary nature of key technical standards. The article then proceeds to examine the ways in which interoperability are a technical goal for information access; and theorizes on how information policies—not technical interoperability—prescribe information access and sharing.

59 Moen, William E. "Resource Discovery Using Z39.50: Promise and Reality." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 185-206.

Whether or not the ANSI/NISO Z39.50 is an antiquated library technology is discussed in this article. In order to evaluate Z39.50 and its position in today's library, the author first offers an historical background of the Z39.50 development. The author also discusses the complexity of information retrieval problems Z39.50 addresses, and how the goals for its use have changed over the years. A discussion of how the constant goal of Z39.50 developers was to enable

interoperability between diverse systems and resources. The author then explores how Z39.50 enables this interoperability, yet lists reasons why its use has been deficient in achieving interoperability. The paper concludes with an overall assessment of Z39.50 in the network environment, its strengths, weakness and future challenges the standard faces in serving as an information retrieval tool for the library community.

60 Myrick, Leslie. "Harnessing Oracle and XT for Finding Aid Dissemination and Search." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: p45-57.

In this article, Leslie Myrick details the actions, success and failures, of two archival divisions and three Special collections archives building an infrastructure for the production and dissemination of an EAD-encoded finding aids markup in XML. After the conversion of the document, their transformation into HTML was achieved through James Clark's XT transformation engine. The final step needed was to enable cross-collection search, which was accomplished by using Oracle interMedia Text. This conversion resulted in the finding aids being more browsable, searchable, and sensible. The article goes into full detail on the conversion and production of the finding aid. As a result, the project claims that by employing an XSLT-based transformation results in the validation of use of dynamically generated HTML. Moreover, the project asserts how Oracle conforms to open-source tools such as Perl, CGI, and XT.

61 Plosker, George R. "Industry Revolution: Implications for Librarians." Accessed 21 Feb 2004. http://onlinemag.net

Discussed in this article is the impact the insurgence of the internet has played upon the librarian profession and library environment. According to the author, as he terms it, the 'Information Industry Revolution," has created new challenges for the the library profession. Points of interest discussed by the author of the article includes: the Open Web and the information professional; industry consolidation; product quality; changing roles and responsibilities; engineering solutions; marketing and communications programs; and most importantly the relationship between . In addition, the article goes into a brief discussion on the challenges of marketing library services in the age of the World Wide Web and creating interoperable metadata schemas applicable for the internet environment.

62 Podge, Gail. "Metadata Made Simpler." Accessed: February 21 2004. http://www.niso.org/standards/index.html.

A product of the National Information Standards organization, this article is a comprehensive introduction, evaluation and guide to metadata with regards to the internet infrastructure. The article introduces the concept of metadata and further discusses its importance. From there, the article delineates metadata element sets

used in library environments (Dublin Core, TEI, EAD, GILS, VRA, ONIX), and what impact they are having on the theory of standard library cataloging. The article then discusses issues of interoperability and exchange, and provides examples of mapping between schemas and cross-walk creation. Also included in the article is a discussion on the importance of metadata registries, which can provide information on the definition, origin, source, and location of data.

63 Qin, Jian. "Representing and Organization of Information in the Web Space: From MARC to XML." *Informing Science* Volume 3 No 2. 2000 : p83- 87.

This article describes the historical connection between the MARC record format and the evolving metadata schemes for representing information. In particular, the essay delves into the current implications the development of XML is having on the metadata milieu. The article establishes the notion of how the library as an institution has been transformed as a result of network capabilities, i.e., the library catalog is no longer a tool for its own collection—it has evolved into a global network node. The article argues that this transformation of the "physical" library in turn transforms theory and procedures implemented for years by the library community. MARC records, the author contends, are static and inappropriate for today's digital environment.

64 Quam, Eileen. "Informing and Evaluating a Metadata Initiative: Usability and Metadata Studies in Minnesota's Foundations Project." *Government Information Quarterly*18 no 3 2001: 181- 194.

Minnesota's Foundations Project aimed to improve access to environmental and natural resources information. In reconfiguring the metadata for web resources, the project chose to utilize the Dublin Core data elements. This paper explores the projects intent and outcome. Three studies on the project were conducted: needs assessment, Bridges web site interface, and usability of controlled vocabulary in Dublin Core metadata. Of importance for the project, and hence the paper, was the controlled vocabulary in relation to keyword searches; and how effective the Dublin Core elements would be. To insure consistency in keywords, the project implemented two types of controlled vocabularies, a thesaurus and the Content Classification Engine (CCE). The paper goes on to discuss the usability study on metadata and controlled vocabulary. In conclusion, the project found that the metadata needs to be maintained, as well as the vocabulary and other related components, as in this case CCE.

65 Radebaugh, Jackie. "MARC Goes Global—and Lite." *American Libraries* 34 no 2 February 2003: 43-46.

In this article, Radebaugh discusses the backbone of library cataloging for the past half century: the MARC record format. The author delineates how the MARC record format has adapted to its surrounding technological environment. MARC tags, fields, and subfields, now accommodate not only books, but magazines, CD-

Roms, art, music, and objects. In addition, as the article declares, the MARC format has even expanded its template to make room for the exploding universe of web pages and other electronic documents. In particular, web created XML DTD's have been created to make MARC records available for institutions outside of the library community. Moreover, the latest edition to the MARC format—MARC 21 Lite Bibliographic format—is a version of MARC reduced to only essential elements, enabling libraries around the world to easily access, implement, and share MARC records. Radebaugh contends that the MARC record format continues to adjust and redefine itself as new information objects continue to find place in library collections.

66 Reynolds, Regina Romana. "Partnerships to Mine Unexploited Sources of Metadata." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 439-462.

The idea of the library catalog playing a role as a portal to Web resources is discussed in this article. How library standards of bibliographic control can be employed in a Web interface is the immediate challenge for a catalog to act as a portal. The author feels that non-AACR2 metadata schemes are important for the library community to adapt to. The article contends that in order for the library community to fully realize the potential of newer metadata-based catalogs, new partnerships and new sources of cataloging data have to be explored and exploited. Metadata plans that support emerging identifiers such as the Digital Object identifier is a potential source of bibliographic data which libraries can convert to or enhance MARC records. In addition, the article examines ways in which registration procedures could be modified to better provide libraries with the much needed data.

67 Rhyno, Art. "Building XML Databases with Zope and Castor." *XML in Libraries.* New York: Neal Schuman Publishers, 2002:117- 131.

In this article, the author begins by commenting upon how options for storing XML in a database are limited when compared to managing other information stores that use XML. Rhyno then goes on to discuss the problems and success of employing two open-source database systems for storing and serving XML data. The database project, conducted at the University of Windsor, adopted two different approaches for storing the XML in each database. Their project consisted of two different types of content: historical documents marked up in EAD, and the other type encoded by MARC and RDF. The EAD content was administered using Zope, an open-source Web application server. Several XML related products exist for Zope implementation, including DOM (Document Object Module), which is recommended by W3C to standardize how programs Interact with Web documents. For the MARC RDF documents, Castor, a Java-based XML Object/Relational Database mapping tool. The author details the process of conversion for both types. The project proved that storing XML

in a database is feasible without requiring intensive labor; moreover, the author contends that by combining a database with XML tools, libraries can "leverage mainstream technologies to create applications and use database technology to manage large XML collections" (124).

68 Rosenblum, Brian. "Transforming Word Processing Documents into XML: Electronic Scholarly Publishing at the University of Michigan." *XML in Libraries.* New York: Neal Schuman Publishers, 2002 : 149-163.

Discussed in this article is the need, and problems, associated with transforming word processing documents to XML. Seeing as software like Microsoft Word is used considerably worldwide, the idea of a highly structured file is essential. The author, Brian Rosenblum of the University of Michigan, examines the call for a structure for word processing documents that could be readily migrated to XML. The project aimed to convert the word processing document into Text Encoding Initiative (TEI)-based XML. The project was concerned with converting journal articles in their word processing form into XML compliant documents. The main challenge the project faced was with the implementation of DTD. A need to find a balance between customization and standardization is discussed.

69 Saupert, Alenka. "Catalogers' Common Ground and Shared Knowledge." *Journal of the American Society for Information Science and Technology* 55(1) 2003: 55-63.

This paper explores the complexity of subject descriptors in the library environment. Due to the fact that subject descriptors of documents in library catalogs are provided by in-house cataloguers; while subject descriptors of documents in bibliographic databases are usually provided by outside indexers, the articles traces the inconsistency in subject descriptors at his libaray; and attempts to provide practical options to combine the variations. By creating a hypothetical model of the process of subject cataloging, the author explores the steps, tools, considerations, and comprises cataloguers must realize when providing subject descriptors for their catalogs. For future consideration, the author then comments on the importance of combining these in-house and out-source indexing service so as to better the overall process.

70 Schadle, Steven, and Karen Calhoun. "Preparing Cataloguers for the 21st Century." *Tachnicalites,* Volume 23m No 6, November/December 2003: 6- 11.

Due to the rise of electronic sources and various changes made to cataloguing theory, this article discusses the new educational trends that cataloguers of the 21st century need to comprehend. The article discusses how ALCTS in August of 2001 appointed a task force (ALCTS LC Action Plan Task Force) to develop a set of "core components" for technical servces; further develop the education and training needed for 21st century cataloging; and promote the understanding and use of new forms of metadata. The article then goes into detail on the task force's recommendations for the educational enhancement and training of library

cataloguers of the 21st century, complete with an introduction to the various important metadata schemas and markup languages.

71 Schmidt, Heidi. "Using XML to Federate Collections: The Legacy Tobacco Documents Library." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: 89- 101.

With no unified set of descriptors, Schmidt details in this article the flexibility of XML when dealing with a large collection of documents. Due to a settlement between NAAG and six major tobacco industries, millions of tobacco industry documents were made available to the public. As a result, a grant was obtained to create an electronic archive for the documents with a search interface for the entire collection—more than 20 million pages. The article discusses how, in the course of a year, they were able to make the documents available to the public through a common search interface by employing XML technology. The XML schema enabled the project to create an extensible architecture for the Tobacco Library itself.

72 Shuler, John. "XML, UDDI, SOAP: The 'Verbs' and 'Nouns' of 'Semantic Electronic Government Information" *The Journal of Academic Librarianship* 27 no 6 2001: 467-9.

This article explores how public agencies continue to integrate civic services and programs within electronic information networks by creating interoperable machine and software standards and consistencies among platforms, all stitched together by a friendly user-interface. The author of this article feels that the library community, in particular academic institutions, can learn a lot from how these public institutions catalog and map resources. The author comments on the evolving protocols of intelligent information "agents" used by public institutions: XML, UDDI, and SOAP. The author fully explores all three. As stated, XML, (extensible Markup Language), is a framework that facilitates the transfer or recognition of programmable information more easily among differenet computer platforms. Universal Description, Discovery, and Integration (UDDI), is a protocol that enables businesses to quickly and dynamically find and transact with one another. The Simple Object Access Protocol (SOAP), is a protocol that acts as the glue between heterogeneous software components. The author feels that all three of these protocol would a huge benefit for library environment.

73 Sokvitne, Lloyd, and Jan Lavelle. "Creating a Unified E-Government Portal Using Xml." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: 59- 75.

The authors, who work at the State Library of Tasmania, discuss how they employed XML to combine access to numerous government services into one simple portal. By using XML, the project realized that XML is a powerful and flexible data storage medium, and in effect, used XML as the basis for other web services. After the conversion, the library was able to extract data from MARC

records and convert them into XML files that could be used by their search engine and retrieval system. Now in XML format, the catalog was now able to produce a quick development of a Web site that far exceeded the ability of the traditional MARC configuration. However, due to XML's lack of maturity as a technology, the program did experience several disadvantages of employing the new system, which they discuss. In particular, they experienced difficulty making data changes without the capabilities of a relational database, a design that does not correspond to XML technology.

74 Southwell, Kristina. "How researchers Learn of Manuscript Resources at the Western History Collections." *Archival Issues* 26 no2 2002: 91-109.

As the author states, with today's use of bibliographic databases and search engines, researchers discover manuscript resources many different ways. In order to track researching trends, the purpose of this article is to gather statistical data on which access points are most commonly used. The information then gained, the author hopes, can be used to evaluate the effectiveness of a collection's access points and lead to improved services for all researchers. The authors used a survey to collect their data. The results of the research concluded that typical manuscripts users identify primary tools through Internet sources, followed by electronic citations in books, articles, and bibliographies. The researchers concluded that the survey helped identify the needs of the average patron, which in effect is a key to a successful partnership between manuscript repositories and the manuscript research community

75 Stankus, Tony. "Making Sense of Serials: The Digital Object Identifier and Crossref: Consensus Driven Successes." *Technicalities* 22 no.2 March/April 2002: 9-13.

The author of this article explores how individual journal articles—separate from the entire journal—are now reliably traceable on the web as individual information objects. The role Digital Object Identifiers (DOI) is playing on the retrieval and journal articles is examined. Discussed are three specific points that need to be understood in order to comprehend how DOI works, and thus encourage its use. Firstly, the Rule of the DOI and the Right of the DOI are "readily managed by even the smallest and most minimally technologically-capable of publishers" (11). Secondly, the DOI encoding schema champions inclusion, meaning publishers can add to the schema to correspond to specific information objects. And lastly, DOI is supportive of nonprofit organizations. The article then proceeds to discuss CrossRef, the non-profit protocol that facilities the exchange between texts in disparate electronic journals by readers whose libraries had paid electronic subscriptions to both journals; and how it can be beneficial for the library environment.

76 Suleman, Hussein, *et al.* "Building Digital Libraries From Simple Building Blocks." *Online Information Review*, Volume 27, no 5. 2003: 301- 310.

Explored in this article is how metadata harvesting has been established by the Open Archives Initiative as a mechanism for connecting a data provider to a purveyor of services. A rising framework, The Open Digital Library model, is a framework that attempts to divide the services into fitting components corresponding to the philosophy of the Open Archives Initiative. As the author contends, this framework has been implemented by a variety of projects has been evaluated for its "simplicity, extensibility, and reusability to support the hypothesis that digital libraries should be built from simple Web service-like components instead of as a monolithic software applications" (301). The article provides and overview and background of the OAI; the concepts behind ODI; and several case studies that delineate specific digital library projects.

77 Tatem, Jill M. "EAD: Obstacles to Implementation, Opportunities for Understanding." *Archival Issues* v. 23 no. 2 (1998) : 155-69.

For the archival discipline, the innovation and acceptance of the markup language Encoded Archival Description (EAD) has revolutionized the field's cataloguing and documentation creation. This article, written in 1998, the "early" days of EAD, is important because it discusses the then negative perception of EAD's complexity and usefulness, or lack thereof. The article attempts to encourage the acceptance of EAD by expressing that an improvement to the ease of EAD's use depends on advances in authoring and browsing software—in other words, the rate of technological advancements will soon indeed compliment EAD creation and use. As mentioned, when this article was written, EAD was still in its "infancy"; the article provides an informative overview of its use, relative advantages, compatibility, all with an attempt of encouragement and patience for the then immature markup schema.

78 Taylor, Stephanie. " A Quick Guide to XML." *Interlending & Documentation Supply*. Volume 31 Number 3, 2003: 187- 193.

This guide to XML offers a quick and comprehensive overview of XML. The article discusses the history of the markup language—how and why it was developed—how it works, and where future developments might lead. Also included is an ample description of XML's related technologies and their function and purpose within the XML construct. In addition, the application of XML and its flexibility of use are detailed. The author investigates how the XML protocol can benefit many aspects of a library catalog. At the end of the article there is a very informative appendix that lists frequently used terms with regards to the XML technology.

79 Tennant, Roy. "The Engine of Interoperability." *Library Journal* (1976) 128 no20 2003 : 23-29.

Of paramount importance for the interoperability any cataloging project is to create and follow a standardization of a metadata form. In this article Roy Tennant discusses the importance of interoperability and the challenges that librarians must address in this age of technology. Tennant briefly discusses metadata standards that are vying for cataloguer's undivided attention, i.e. the Dublin Core data elements, The Metadata Object Description Schema (MODS) and the Metadata Encoding and Transmission Standard (METS). In addition, Tennant discusses the emerging protocols that provide service for metadata transmission. Tennant brings to the forefront the Open Archives Initiative (OAI) Protocol for Metadata Harvesting. As Tennent explains, "the protocol defines a method by which digital repositories can 'expose' their metadata for harvesting and indexing in centralized discovery services" (25).

80 Tennant, Roy. "The Importance of Being Granular." *Library Journal.* 127 no9 May 15, 2002 : 32- 36.

In this article, Tennant discusses how the library community is dependent on metadata, and, in a sense, prisoner to its proclivity to mutate and morph. Tennant introduces the term granularity and its implications to metadata theory. Granularity refers to "how finely you chop your metadata." The granularity of a metadata schema is vital to understand and Tennant goes into detail on the importance of granularity. Tennant goes onto comment on how the granularity and extensibility of the MARC format is limited. But this is not to say that MARC is the only metadata schema that faces granularity difficulties. According to Tennant, no metadata scheme is granularly perfect. It is suggested that a cataloging project review individual schemes and how their granularity would best correspond to the collection to be catalogued.

81 Tennant, Roy. "MARC Exit Strategies." *Library Journal* (1976) 127 no19 15 2002: 27-29.

Delineated in this article are the weaknesses of the MARC record format, followed by strategies and advice on how to convert the record format to a newer, more productive form. Tennant urges librarians to recognize MARC's inherent weaknesses, and begin rethinking cataloging principals. He discusses how the current standards are centered on the book as a physical object, and warns cataloguers of the rise of the "virtual" document and how that will affect the cataloging system. Moreover, Tennant argues that MARC is a "flat" cataloging structure which will not accommodate hierarchical structuring, the en vogue nuance of today's world of bibliographic description. The two exit strategies described in the article focus on starting from scratch and migration. Tenneat briefly covers the economic and labor nuances of both.

82 Tennant, Roy. "MARC Must Die." *Library Journal*, (1976) 124 no 22 July 23 2002: 64-68.

The future of the MARC record is examined in this article. The author contends that the format, in today's digital world, is nearing obsolescence. Tennant states that the problems with MARC are serious and extensive; and that it's beginning to lag behind current bibliographic modes of description. Aside from its obscure coding, which is only readable by machines and trained cataloguers, Tennant feels MARC's level of granularity is not competent enough for cataloguing today's rich field of resources. Moreover, Tennant argues that if the libraries continue to use the MARC record, instead of migrating to a more universally and productive metadata form, it will seriously compromise the library institution as a whole. In the age of technology libraries need to associate formats with other institutions on an international level, and the MARC format hinders cooperation.

83 Tennant, Roy. "Not Your Mother's Union Catalog." *Library Journal*, (1976) 128 no 28 Ap 15 2003: 24-28.

In this article, Tennant brings to point how OCLC World Cat and RLG's Eureka—two of the world's largest databases—plan to change their cataloging format dramatically. Tennant feels that these changes will indeed have an impact on the future of library cataloging. As Tennant states, OCLC is reengineering WorldCat so thoroughly that "nothing remains unquestioned, including the MARC record" (24). What has inspired the theory behind these cataloging industries was the outcome of a conference in 1998 by the International Federation of Library Associations and Institutions (IFLA). As a result a published report was created, named Functional Requirements of Bibliographic records (FRBR)—aa Tennant describes it: "a revolutionary recasting of the bibliographic record on behalf of the library user" (25). Basically, FRBR defined such principals as the hierarchical dimensions of a creative product: work (distinct creation), expression (realization of a work), manifestation (physical embodiment), and item (a single exemplar). As the reformatting done by WorldCat proves, uses FBRB principles provides catalogues to "ingest" virtually any metadata, creating a revolutionary level flexibility for the library catalog.

84 Tennant, Roy. "Publishing Books Online at eScholarship." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: 101-114.

Roy Tennent discusses in this article how XML could be utilized to replace some of the functions of SGML, in particular, encoding ebooks. According to Tennent, and as a result of the ebook project Tennant head at the University of California, San Francisco, the capabilities of XML exceed what SGML could previously offer; one example being the ability to enable the user to control dynamic displays. The ebook encoding project's primary concern was to create an easy-to-use searching within a book, as well as across books. The article details the

paths the project took in creating the searchable indexes, explaining in detail the software and server specs, encoding schema, as well as the creation of user-defined style sheets. The author states that "XML can be as much an art as it is a science," meaning that due to the immaturity if XML as a technology, creating a forceful search service calls for a lot of imagination and flexibility.

85 Terris, Olwen. "Cahos in Compromise—Cataloging and Indexing in a Digital Age." *Multimedia Information Technology* 29 no 3 August 2003: 79-85.

With the rise of digital material emerging form a library setting, Olwen wonders if emerging librarians were adequately educated on cataloging and indexing principles during graduate school. The author feels that if the knowledge of the cataloguer is limited, the ability to identify and retrieve "relevant value-added" information will be severely compromised." With large amounts of electronic resources being described by people with limited ability, the potential for resource loss is examined. The author feels that powerful search engines must have an authority-driven structure behind them if they are to be of any assistance for semantically structuring the web; and the corresponding metadata structure relevant to theses resources, the author feels, is a complicated issue, which he terms as a chaos of compromise.

86 Thomas, Sarah E. "The Catalog as Portal to the Internet." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 21-37.

This article examines how the library community can reinterpret the catalog so that it can serve effectively as a portal to the Internet. Questioned by the author is if the catalog can be an appropriate model for the discovery and retrieval of highly dynamic, rapidly multiplying, networked documents. The author examines how web portals are beginning to replace catalog as the dominant index for library users. As the author contends, the library catalog represents stability, dependability, reliability, and quality—all of which are threatened by the erratic control of internet sources. The necessity for the library community to administer authoritative control on Web resources is argued by the author. In effect, the author contends that the library catalog, or permutations of it, can help. The article proceeds to explore the changes the library community needs to make in order to obtain these goals.

87 Tillet, Baraba. "Authority Control on the Web." *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium.* Washington, D.C.: Library of Congress, 2001: 207-220.

This paper examines the International Federation of Library Association's attempt to enable interoperability across multiple authority files and to link and provide switching displays of authorized headings on an international scale. The addition of library catalogs to the blend of information being searched on the Web has

created a stir in the Web's lack of authority control. The author examines how the combinations of Unicode and new technologies are opening up access to all scripts and languages. The author explores how crosswalks can link Dublin Core metadata and cataloging records in MARC and other formats with XML and other relative communication structures, thus and expand the opportunities for contributing authority records to an international pool. The author states how a pool of authority records for bibliographic entities to use on the Internet is of interest to not only libraries but other publishing industries.

88 Todd, Chris. "Metadata Mayhem: Cataloguing in the National Library of New Zealand." *The Electronic Library.* Vol 21 Number 3, 2003: 214-222.

Employing a chronological approach to the cataloging of electronic resources within the National Library of New Zealand, this essay examines the history of metadata practice within its institution; followed by an examination of current approaches to cataloging published digital materials and the transformation of the catalogue record as a result of the conversion of metadata schema. Lastly, the article details prevalent issues that are still under discussion within the institution. In the creation and conversion of new metadata schemes, the University used Dublin Core elements expressed in XML. Discussed are the benefits and difficulties of such a process. An example of a difficulty faced was mapping documents relating to the data and presentation requirements of the Web. The article also details the fields that were added to MARC records, and the transformation of Discover records, for use in their ENCompass database system. File format examples are printed in the article for close inspection and study.

89 Trippe, Bill. "Can XML Drive Taxonomies and Categorization?" *EContent* 26 no11 N 2003: 30-4.

This article explores how XML capabilities can foster and improve the standards and implementation of taxonomies and categorization. The author fully explores content management and the technological standards associated, and what implications they are having on contemporary organization and control of documentation. The article researches the impact XML has had on the disciple of information specialists. Explored by the author are the many functions XML can produce for any institution engaged in categorizing resources; discussed are the creation of taxonomy, thesauri, and keyword lists, and the benefits to construction them with XML and its related technologies.

90 Veen, Theo van. "Searching and Retrieving XML Records Via the Web." *XML in Libraries.* New York: Neal Schuman Publishers, 2002: 17-27.

Access to information within the Koninklijke Bibliotheek infrastruce is examined in this article. The database contains roughly twenty catalogs of different collection in which the access is provided via web applications written in Perl. The reliance upon one particular programming language has led to "increasing

danger of divergence." In fear of a lack of interoperability, the infrastructure aimed to rid themselves of reliance upon and the complexity of the Z39.50 protocol. With the new infrastructure, responses to requests were to be output in XML only. In doing so, they would create a standardization entrance for the search and retrieval of XML records by creating user interfaces using XSL exclusively; and meanwhile conform to international standards of search and retrieval. The outcome of the project revealed the power of XML when working with a large database store(s) of information.

91 Weiss, Amy K. "A History and Analysis of the Cataloging of Electronic Resources. *Library Resources & Technical Services* 47 no4 O 2003: 171- 187.

This article examines how the cataloging rules of computer-based materials continue to undergo modification and revision. The article focuses on analysis and comparison of the various codes and guidelines issued for the cataloging of electronic resources. The creation of new rules, the article contends, has been spurred by introduction of "new physical carriers, the preeminence of materials accessed remotely versus those with physical carriers, the need for guidance in cataloging specific instance of computer-based materials, and the evolution of the theoretical concerns underlying cataloging codes." Based on this history of constant change, the author feels that it is easy to predict that many more changes for the cataloging of computer-based materials will continue to occur. But the author warns that theses changes may result in confusion; and moreover, that, whenever possible, the older rules and standards should be fully examined and tested for they may be readily adaptable for the new types of electronic resources.

92 Weiss, May K. "Proliferating Guidelines: A history and Analysis of the Cataloging of Electronic Resources." *Library Resources and Technical Services* 47 no4 O 2003: 171-187.

This article provides an analysis and comparison of the range of codes and guidelines for implementation of electronic resources. The author examines how the creation of new cataloging rules was necessitated by new physical carriers, materials accessed remotely, rise of computer-based material, and the evolution of theoretical purposes of cataloging codes. She then proceeds to determine if the proliferation of the creation of new rules and standards for cataloging electronic resource needs to be suppressed, as she feels that older rules and standards can be applied and adapted to new types of electronic resources. AACR2 cataloguing rules are dissected and its weaknesses are determined and strengths highlighted. In particular, the author examines the much debated concept of AACR2's handling of content vs. carrier. In addition, the Dublin Core data elements are examined as well as the usefulness of ISBN.

93 Wells, Katherine L. "At Our Fingertips" Online Cataloging Tools." *Mississippi Libraries* 64 no4 Winter 2000: 117-122.

Discussed by the author of this article how the rise of the internet has not only evolved cataloging practice; but as the tools used for cataloging—in other words, the internet has changed the dynamic of the profession twofold, both in intellectually and manually. This article explores the benefits internet capability has on the cataloging profession and provides various links and recommendations for cataloging implementation. In particular, the author provides and examines links to general cataloging information, like the Internet Library for Librarians; sites on handling on bibliographic description, such as sites that provide rules for AACR2, amongst others; sites on classification procedures, like The Library of Congress Cataloging Policy and Support Office (CPSO); bibliographic formats and standards for OCLC and MARC tools; as well as a discussion on online tools for cataloguing procedures when handling non-book material.

94 White, Layna. "Museum Implementation of Encoded Archival Description." *Art Documentation,* 21 no1 Spring 2002: 13- 20.

Due to the desire to meet new standards and improve overall reference and access services, the Grunfield Center for Graphics Arts Museum set out to update their levels of description for data on exhibitions and various other forms of documentation. Before covering the process, the article introduces the Encoded Archival Description (EAD) and the usefulness of its revolutionary form of hierarchical handling of object description. The author then details the act of implementing EAD; the inherent data structure standards, of in this case, Dublin Core; the corresponding data content standards; and data value standards. The author than evaluates the use of EAD for a Museum collection. Although the author found EAD to be more suitable to foster description for her libraries collection, she feels that the Museum community would benefit more from the schema if a specific Museum DTD was created.

95 Wisser, Katherine M., and Jennifer O'Brien Roper. "Maximizing Metadata: Exploring the EAD-MARC Relationship." *Library Resources & Technical Services* no2 Ap 2003: 7—76.

The markup language, Encoded Archival Description (EAD) has provided a new way for the library community to represent collections of manuscripts and archives. Firstly, this article reviews, previous to EAD, the representational practices of archival cataloging in order to highlight the advantages of employing EAD. For a collection of their finding aids, the authors of this article, librarians at the North Carolina State University Libraries, sought to test the process of cross-walking EAD to MARC. By creating a EAD template for the finding aids, MARC records were produced that only required minor editing and revision. As the author's contend, the creation of a bridge between EAD and MARC has stirred much theoretical debate in the library community. With this project, the

researchers have attempted to level the tensions of opposing opinions, and produce options, suggestions, and procedures to emulate. In conclusion, the article ascertains that an inclusion of EAD within the MARC world, not only improves resources description, but creates a cooperation and partnership between not only two collections, but two types of library professionals, cataloguers and archivists.

Authoritative Web Sites

96 *Dublin Core Metadata Initiative.* Dublin Core Metadata Initiative. Accessed 15 March 2004 <http://dublincore.org/>

Since its inception in 1995, the use of The Dublin Core Metadata Initiative has been recognized and implemented on a global level. Simply put, Dublin core was initially intended to facilitate the discovery of information objects on the web. Its flexibility and easy to implement data schema has attracted many. The mission of the Initiative is to develop metadata standards across domains; establish framework for the interoperation of metadata sets; and facilitate the development of discipline specific metadata sets to encourage metadata standards and interoperability. The Web site provides an overview and history of the Initiative; a description and discussion of its metadata elements; encoding guidelines; as well as examples of DCMI documents and projects that have utilized Dublin Core elements.

97 *Library of Congress.* "MARC Standards" Accessed 20 March 2004. <http://www.loc.gov/marc>

This page of the Library of Congress Web site includes an extensive overview of both the traditional usage and recent developments of the MARC record format. The Web site includes the basic MARC record structure, i.e., a description of the MARC with a tutorial; bibliographic formats and code listings; ordering information, MARC mapping; amongst other similar topics. Most importantly, the Web site provides links to the impact information technology is having on MARC reconfigurations. For instance, links to MARC in XML; MARCXML "Slim" schema; and the MODS schema can be accessed. Moreover, the site provides a link to the implication FRBR has had on the MARC record format. Included is a download for a free FRBR display tool.

98 *National Information Standards Organization.* National Information Standard Organization. Accessed 20 March 2004 <http://www.niso.org/>

The National Information Standards Organization (NISO), is an association that develops standards for libraries as well as for publishing and information services.

As stated on the Web site, "(a)n NISO standard is developed through consensus, defines model methods, materials, or practices for libraries, bibliographic and information services, and publishers." NISO, a non-profit association accredited by the American National Standards Institute, identifies, develops, maintains, and publishes technical standards to manage information for environments associated with construction and control of new digital media . NISO standards apply both traditional and new technologies to the full range of information-related needs, including retrieval, re-purposing, storage, metadata, and preservation. NISO publishes its standards freely on their Web site. For instance, links to ANSI/NISO Z39.50 Information Retrieval: Application Service Definition and Protocol Specification, along with the other thirty two approved standards can be directly downloaded via their site.

99 *The Text Encoding Initiative*. 2004. TEI Consortium. Accessed 26 February 2004 <http://www.tei-c.org/>.

Initiated in 1987 and presently hosted by Brown University, University of Oxford, University of Bergen, and the University of Virginia, The Text Encoding Initiative is an international and interdisciplinary encoding scheme that represents literary and linguistic texts for online research. The TEI began as a research effort cooperatively organized by three scholarly societies: the Association for Computers and the Humanities, the Association for Computational Linguistics, and the Association for Literary and Linguistic Computing, that was funded solely by substantial research grants from the US National Endowment for the Humanities, the European Union, the Canadian Social Science Research Council, the Mellon Foundation, and others The TEI standard has been implemented for online documentation by libraries, museums, publishers, and various scholarly endeavors. The web site provides information on projects and institutions; tutorials and guidelines for implementing TEI; and links to TEI specific and generic free software for exploiting the TEI scheme.

100 *World Wide Web Consortium*. 2004. The World Wide Web Consortium. Accessed 22 Februsary 2004 <http://www.w3.org/>.

The World Wide Web Consortium (W3C) is an authoritative web resource that provides specifications, guidelines, software and tools for interoperable technologies. Ostensibly, W3C is a forum for information, commerce, communication, and collective understanding. The Web site comprises various introductions and tech reports on web related subjects. Included on the site are in depth activity reports, e.g., information on budding projects like the Semantic Web; relevant news stories; along with numerous technical reports and publications covering the latest advances in software, hardware, metadata, and markup languages. The site also consists of numerous introductions and tutorials for various markup languages, schemas, and metadata standards.

CHAPTER V

ANALYSIS AND EVALUATION

Chapter Overview and Restatement of Problem

Fixed and productively functional for over forty years, the rules and standards of library bibliographic control and description are presently being contested by authorities in the field of cataloging. Frequently, with the rise of a new technology comes the rethinking of established theories. The Internet and World Wide Web produce manifestations of documents, challenging the library milieu to review its fundamental cataloging standards and practice. The library community currently faces a period of transition where the reevaluation of the theory and practice of cataloging must be reconsidered, and perhaps even drastically altered. As metadata schemes and theories corresponding to these new forms of bibliographic description continue to proliferate amongst the library community, individuals seeking to comprehend the new standards, and libraries in search of an appropriate form of metadata for their collection, may find the abundance of resources intimidating and unstructured. The researcher has addressed this situation by identifying, analyzing, and organizing resources related to metadata for the library environment. After analyzing the content in relation to metadata in over forty library journal articles, a tool of access in the form of an annotated bibliography was created in order to introduce concerns and facilitate access to the journal articles selected.

As is evident in the articles annotated, the fundamental practice of metadata in the library environment presently faces a period of transition. The researcher believes that the annotated bibliography created in chapter IV will not only provide an overview of current metadata schemas that are appropriate for the library environment; but will also

enable a thorough examination of current metadata schemas' advantages, weaknesses, and future possibilities.

This content analysis study was conducted in order to address the following research questions:

9. How functional are the MARC record format and AACR2 rules for bibliographic description in the contemporary library environment that houses "modern" forms of documents?
10. When migrating records to new standards and schemas, what factors will librarians face in order to enable interoperability?
11. How important is it for librarians to both understand and implement eXtensible Markup Language (XML)?
12. What are the new metadata standards available and appropriate for digital and special collection library cataloging implementation?

This chapter will analyze and evaluate each question by examining the findings supported by the articles annotated. The citations that appear in the following bibliographic essays will correspond to the number that appears before the article annotated in Chapter IV. Where appropriate, following the interpretation of the research questions posed, the reader will find a record or data elements of the metadata schema discussed.

Response to Research Question One: How functional are the MARC record format and AACR2 rules for bibliographic description in the contemporary library environment that houses "modern" forms of documents?

Findings

The technological advancements of the late 1960s—mainframe computer processing of Library of Congress standards via machine readable cataloging—resulted in the reconstruction of both theory and practice of library cataloging. Empowered by network capabilities, the MARC syntax, MARC data elements and the Anglo American Cataloging Rules (AACR) enabled libraries to share resources and data. During this period, memory, storage, and processing power were all rare and expensive. These factors, however, did not impede the utilization of the MARC record, one of many reasons why the MARC syntax was so beneficial for the library environment.

Today, with the advances in memory, storage capacity, processing power, and the development of digital information objects, the usefulness and practicality of the MARC record format and AACR rules of description are being seriously questioned. Due to various new forms of "documents," the limitations of the MARC record format are glaring. A vast majority of scholars agree to the fact that these limitations gravely jeopardize the library catalog. Hypothetically, if possible by the snap of the fingers, most cataloguers would replace the MARC syntax with XML. But seeing as libraries on a global level have been cataloging with MARC for more than 40 years, the conversion and/or migration of the literally millions of records is highly problematic. The conversion would not only exhaust time and money, it might also require, for certain institutions and cataloguers, intensive training. Scholars readily agree on MARC's weaknesses, but there are several approaches to ways in which they should be handled. As the title of Tennant's article "MARC Must Die" (82) infers, a few scholars contend that the implementation of the syntax should immediately cease. Others, cognizant of the

fact that there are an enormous output of MARC records already created, proscribe to the adjustments the Library of Congress has made with the syntax, contending that these alterations are indeed sufficient for resource description of digital information objects. A review of the scholarly opinions on the present and future role of the MARC record format expressed in the articles annotated is important to examine.

Previous to the digital infrastructure, bibliographic description was handled efficiently by AACR2. Matthew Beacon's article (4) addresses how the World Wide Web has created profound changes in how librarians use information. Beacon contends that there has been a significant change in the way knowledge is now packaged, published and disseminated, and how access to knowledge is controlled. Theses changes have highlighted imperfections to the AACR cataloging description and present librarians with adequate means of description that previous were not extant. According to Beacon, these limitations call for a reevaluation of the handling of various AACR rules of description. Connell and Praba agree, contending that changes needed to be made to ACCR are lengthy but realistic (19). Furthermore, Gourley feels that these changes are just a necessary but ephemeral "stop-gap" (33). In other words, in due time the MARC format will become obsolete.

Widely agreed upon is the idea that the continuation of the use of the MARC record format depends on alteration to its form. Some scholars feel that due to economic pressure—the fact the vast majority of sources are already configured via MARC, and the time it would take to migrate or convert records to a new format—cataloguers are forced to continue to utilize the MARC format and make the necessary alterations. Sanctioned by Library of Congress and supported by scholars such as Kevin Clarke in his article

““Updating MARC Records with XMLMARC” (13), there is a present movements for migration to XML.

But according to scholars such as Roy Tennant, there are many serious and extensive problems with the MARC record format that may jeopardize the library institution. For one, although the author admits that the MARC format is a competent standard, “it lacks essential checks and balances to assure that appropriate granualarity—how finely the individual elements are chopped—is achieved when coding a record” (81). Tennant discusses how the current standards are centered on the book as a physical object, and warns cataloguers of the rise of the “virtual” document and how that will affect the cataloging system. Moreover, Tennant argues that MARC is a “flat” cataloging structure which cannot accommodate hierarchical structuring, the powerful trait in today’s world of bibliographic description.

Interpretation

What some critics feel is the greatest disadvantage of continuing to utilize the MARC record format is the fact that no other profession outside of the library environment uses it. With a structure that is arcane, and only readable by computers or the trained, some scholars feels that the continued implementation of the MARC record format will in effect segregate the library community from other institutions, i.e. publishing firms, government agencies, media aggregators, and various other information professions. As the network infrastructure expands physical and intellectual barriers, certain scholars feel that libraries will endanger their place and role in the research environment if they do not adhere to contemporary standardizations implemented by organizations globally.

Response to Research Question Two: When migrating records to new standards and schemas, what factors will librarians face in order to enable interoperability?

Findings

Stated simply, interoperability enables use of content produced on one system to be used on other systems. Due to the impact of the networked infrastructure, the flow and processing of information is greatly affected by specific technical standards that were previously handled in the library environment sufficiently by the Z39.50 protocol. According to Moen, the Z39.50 standard was intended to address resource queries within a limited community (i.e. libraries), but has now evolved to incorporate advances in technologies and approaches (58). However, as Moen contends, the future of Z39.50 is uncertain due to its cost of complexity. The continuation of the protocol requires extensive programming knowledge of languages like C++ and technical tools like Abstract Syntax Notional One (ASN.1) and Basic Encoding Rules (BER). With the processing power and memory available today, setting up a Web-server with full-text indexing search engine is, relatively speaking, simple. Moreover, absent in Web-server databases and present in Z39.50 is the need for various plug-ins, which may lead to interoperable difficulties.

As libraries look to adapt their holdings to the standardizations arising due to advancement in information technology, they must recognize current trends evolving out of other information professions. In particular, as discerned by Dorman is his article "'Furber' Debuts at ALA Exhibits" (25), the International Federation of Library Associates (IFLA) newest cataloging standard Functional Requirement for Bibliographic Records (FRBR) creates for the library environment a cataloging schema that is

exceedingly more flexible than any schema previously practiced by the library milieu. FRBR's creation was greatly influenced by this notion of "library cataloging liberation." In other words, the limits of the MARC format have segregated library cataloging data from the rest of information providers. FRBR, in essence is, as Tennant describes: "a revolutionary recasting of the bibliographic record on behalf of the library user" (25). In effect, the utilization of the FBRB principles enables catalogues to "ingest" virtually any metadata, creating a revolutionary level of flexibility for the library catalog. FRBR will not only help foster interoperability of the library catalog, but it will also eradicate various weaknesses of the MARC format. FRBR enables the creation of hierarchical dimensions of a creative product: work (distinct creation), expression (realization of a work), manifestation (physical embodiment), and item (a single exemplar), something which MARC fails to do. For numerous reasons, this hierarchal structure, based primarily on XML principals, is the trend currently being utilized by various institutions world wide. The FRBR structure not only results in the granularity of search results; but it will also create a framework that will enable the library to interact with other non library catalogs.

As a result of the proliferation of metadata schemas, another concern for the library community is creating interoperable platforms amongst the library community itself. As one library employs Dublin Core to cataloging its special collection, another will invariably choose a different schema. How to achieve the best retrieval results in cross-domain searching has developed into an important challenge for librarians. In order to accomplish such searches "it is important to render the different knowledge organization systems, such as controlled vocabularies and classification schemes,

interoperable within a single search apparatus" (15). Interoperability between and among different subject vocabularies—both controlled and uncontrolled—is critical for cross-domain searches. As Chan explores, librarians and information professional must create the mapping and integration of existing knowledge organization system. Moreover, when mapping together various collections or creating cross-walks between, librarians must recognize specific institutional concerns such as integrating views of different cultures.

Similarly, the flux of Web pages and Web-related resources have been recognized by such institutions as IFLA as being problematic for the establishment of interoperable standards. The addition of library catalogs to the blend of information being searched on the Web has called attention to the Web's lack of authority control. Tillet comments on how the combinations of Unicode and new technologies are opening up access to all scripts and syntax languages; which in effect will ease the creation of cross-walk creation and thus expand the opportunities for contributing authority records to an international pool (87).

Interpretation

As networked communication continues to rapidly eradicate physical barriers of resource collections through virtual access, librarians must continue to not only keep current with compatible software and hardware configurations, but also recognize impending adjustments that may be needed for metadata, schemas, and languages. Current standards, such as Z39.50, are currently being revaluated and assessed. Librarians must realize that these standards and new applications must meet international

compliance. Of paramount importance for today's digital environment is the compatibility with an array of software, hardware and metadata configurations.

Response to Research Question Three: How important is it for librarians to both understand and implement eXtensible Markup Language (XML)?

Findings

In such a short period of time, information technology has had an enormous international affect on all aspects of communication. Through the Internet and World Wide Web, HTML, with its simple design capability, became the new information medium, allowing organizations to readily disseminate textual and/or graphical information. Derived from the exceptionally more complicated markup language SGML, HTML's greatest attribute is its simplicity, enabling individuals and organizations with limited or no language programming knowledge to disseminate information. However, as the infrastructure exploded, information professional began to realize that HTML's greatest attribute—its simplicity—was also its greatest weakness. What contributes to the simplicity of HTML implementation is its lack of standard definition and control, both of which result in unstructured documents with little or no authority control.

As a subset of the exceptionally more complicated Standard general Markup Language (SGML), eXtensible Markup Language (XML) was designed with more precise rules and fewer features and options in order to facilitate production and dissemination. Unlike HTML and its graphic and design capability, XML harkens back to a text-based structure and an emphasis on storing and representing information. The operative words are "storing" and "representing" data, and not "presenting" data, which

is what made HTML so popular and efficient. XML works by separating style from content, simple in concept but revolutionary in theory. As HTML solely presents the data, XML addresses the context or meaning of the data. XML permits designers to customize tags, which in effect facilitates the transmission, definition, and interpretation of data between applications and organizations. Indeed, XML is important because it is becoming the language of the Web; a language that is much more structured than HTML, a language that enables interoperable exchange of information—a concept that rings "library science."

As business and governmental agencies recognize XML's benefits, so to has the library environment. Discussed in the articles annotated are the many advantages XML can create for numerous library services. A majority of scholars and information professional agree that, within a few years time, it is very possible that XML will be fully immersed in every HTML document, either in the XHTML form or attached as an XML document. Of critical importance for the 21st century librarian is the understanding, implementation, and nuances of XML and how it will affect bibliographic description and control.

XML, unlike HTML, has many associated technologies and protocols that function with the markup language to enable interoperable capability. Important to understand is the function of an XML document's DTD or associated XML schema. As Anderson and Chandler explain in their article, XML DTDs and schemas tell XML processors which element and attributes are allowed in a document (17). This allocation of elements is contingent upon the specific, intended discipline. Although each DTD or schema defines the element, XML allows for inclusion of elements as one see fit.

Shuler examines a few other protocols that are presently deployed by government agencies, and comments on why the library community should emulate their practice (72). For instance, Shuler examines the Simple Object Access Protocol (SOAP), a protocol that acts as the glue between heterogeneous software components, and infers the advantages it would provide for the library network.

The scholar Stankus explores how a library's serial collection can greatly benefit by formatting it with a XML structure. Stankus explores how an individual journal article—separate from the entire journal—is now reliably traceable on the web as individual information objects. In conjunction with XML, Digital Object Identifiers (DOI) can be used to facilitate retrieval of journal articles. Three specific points need to be understood in order to comprehend how DOI works. Firstly, the Rule of the DOI and the Right of the DOI are "readily managed by even the smallest and most minimally technologically-capable of publishers" (75). Secondly, the DOI encoding schema champions inclusion, meaning publishers can add to the schema to correspond to specific information objects. And lastly, DOI is supportive of nonprofit organizations. Stankus proceeds to discuss CrossRef, the non-profit protocol that facilities the exchange between texts in disparate electronic journals by readers whose libraries had paid electronic subscriptions to both journals—which can obviously be beneficial for the library network.

Another advantage XML presents for the library environment, as expressed by Banjerlee in her article, is its ability to proficiently handle the structured information of loan requests, and thus improve interlibrary loan services (3). Although Banjerlee contends that the XML reformatting is a great improvement, she admits that XML is both

memory and CPU intensive, which results in slow processing. Moreover, the author warns, until the technology of XML matures, considerable trial and error implementation will be required when attempting a software cross-over.

Interpretation

The ability of XML to create the structure of an hierarchical relationship, which in effect establishes the context of data, has already had a major impact on the information professionals outside of the library community. As professed by numerous scholars, it is imperative for the library community to adapt to and implement XML, the markup/syntax language of the 21st century. In doing so, XML can not only improve numerous library services, it can revolutionize every facet of data collection, data searching, and data retrieval. It is also important to consider the implications XML will have on the future of the library discipline. If libraries do not readily adapt to XML, and if XML and its related technologies continue to rapidly progress—drastically affecting the ways in which information is created and retrieved—the library environment is in jeopardy of excluding itself from distinguished information and research professions.

Response to Research Question Four: What are the new metadata standards available and appropriate for digital and special collection library cataloging implementation?

Findings

As separate collections devoted entirely to archival or special collections continue to prosper in today's digital environment, it has become increasingly obvious that traditional library cataloging standards are inappropriate for collections that comprise myriad forms of information objects. Digital information objects, as has become

apparent, cannot be adequately handled by traditional library bibliographic encoding. What many scholars find problematic is MARC's lack of interoperability and compatibility of the data it carries in today's digital environment.

Information technology has enabled libraries to not only create an Online Public Access Catalog (OPAC), but also a "digital library." A digital library essentially enables the entire collection— information objects in the form of digital image, sound and or text, along with the metadata used for control and description—to be accessed remotely. Depending upon the collection and user needs, various metadata schemes and software/platform configurations come into play. As expressed by Campbell in his article (10), parallel with the rise of network speed and capability was the mass production of electronic resources created by proprietary systems, making access across disparate systems near impossible. Moreover, the proliferation of HTML documents has resulted in a lack of authority control. Although platform mapping has somewhat controlled the lack of metadata authority, it has not resolved problems associated with islands of disconnected data and information. The rise and evolution of various new information objects have raised a considerable amount of issues for libraries; foremost being the cataloging of both the metadata of the new information objects as well as the migration of the old.

New metadata schemes and cataloging theories have evolved in order to foster catalog description for collections holding modern information objects. Many librarians have found The Dublin Core Metadata Initiative (DCMI), a scheme originally created for the description of information objects to facilitate discovery on the web, flexible and suitable for non traditional library collections. But as librarians recognize the advantages

of using Dublin Core elements, they must first analyze the metadata scheme already employed; then consider the difficulties and challenges that will occur during migration to Dublin Core. As Chepeseuick contends in his article, Dublin Core will not replace MARC, but compliment the format (18). As the author expresses, one advantage to implementing Dublin Core elements for cataloging resources is the fact that there are no set of "rules" that limit a cataloguer in the age of multimedia. The author posits that Dublin Core is not going to endanger the MARC record format; but rather MARC will continue to be valuable to the cataloging community as a template or framework for metadata collection. The author discusses how practical Dublin Core elements are for cataloging electronic resources—something MARC has difficulty describing and controlling. The author contends that the flexibility of the Dublin Core elements will allow for adjustments within specialized communities to met with their own needs.

As is exemplified by the digital project discussed in the article by Bishoff and Garrison, the benefits of Dublin Core are self evident for a collection that attempts to combine resources from an assortment of collections—in their case combining all the resources from an archival, historical, library, and museum into one digital collection (5). As their collection comprised different metadata schemes (e.g. the library used MARC, and the historical used Dublin Core itself), the mapping of Dublin Core elements to other metadata elements sets are possible, and moreover, independent of the Dublin Core itself. By establishing a set of core elements derived from Dublin Core elements, the project was able to create proficient cross-database (cross-collection) searching.

Archivists have had a tradition of archival description which greatly differs from traditional bibliographic description. While bibliographic description is centered on a

single publication, archival description centers on record groups. Thus the MARC record format fails to create a relationship between records, or in other words a hierarchical arrangement. With the extensible power of XML, archivists have created Encode Archival Description (EAD), a document type definition appropriate for the handling of groups of records in a hierarchical relationship. As Tatem describes in her article, archival documentation must consider both the concepts of *respect des fonds,* also known as the principle of provenance, and the principal of original order (77). Unlike the "flat" design of the MARC record, EAD fosters these two important principles for archival documentation and description.

Interpretation

In today's digital environment, library collections continue to comprise a myriad form of information objects that cannot be properly controlled or described by traditional library cataloging procedures. The flexibility and compatibility of the Dublin Core data elements are far more appropriate than MARC for the description of information objects such as photographs and other non traditional library resources. Moreover, seeing as the the Dublin Core Metadata Initiative (DCMI) was created for the description of information objects to facilitate discovery on the web, its future impact on the library catalog are essentially limitless.

For archival repositories, the creation of EAD had revolutionized the cataloging practice. EAD is not only able to handle individual information objects of various form—documentation, photographs, sound recording-- it also creates a hierarchical relationship that is imperative for archival representation.

CHAPTER VI

SUMMARY, CONCLUSIONS, RECCOMMENDATIONS

Research Project Summary

For years the MARC record format met no comparable alternative for describing the content of library resources. However, recent technological advancements-- namely breakthroughs in memory, storage, and processing power-- have forced librarians worldwide to reevaluate the standards of bibliographic control practiced for close to forty years. New technological advancements have begun to highlight the inherent weaknesses of the MARC encoding schema, as well as, to a certain extent, create new dilemmas with the format. Moreover, the World Wide Web is widely recognized as a powerful information tool; but is also admitted universally to be hindered by a lack of bibliographic control—or to put it a bit more dramatically: disturbed by bibliographic chaos. This research project examines the effects the digital revolution has had on the library community. In effect, the research projects attempts to prove that information technology has reinvented library cataloging theory and practice.

As metadata schemes and theories corresponding to these new forms of bibliographic description continue to proliferate amongst the library community, library science students seeking to comprehend the new standards, and librarians in search of an appropriate form of metadata for their collection, may find the abundance of resources intimidating and unstructured. The researcher has addressed this situation by identifying, analyzing, and organizing resources related to metadata for the library environment. After analyzing the content in relation to metadata in over forty library journal articles, a

tool of access in the form of an annotated bibliography was created in order to introduce concerns and facilitate access to the journal articles selected.

The articles annotated addressed four research questions for modern library cataloging:

13. How functional are the MARC record format and AACR2 rules for bibliographic description in the contemporary library environment that houses "modern" forms of documents?
14. When migrating records to new standards and schemas, what factors will librarians face in order to enable interoperability?
15. How important is it for librarians to both understand and implement eXtensible Markup Language (XML)?
16. What are the new metadata standards available and appropriate for digital and special collection library cataloging implementation?

Conclusions

The Internet has had an enormous impact on almost every facet of business, entertainment, education, and communication. Institutions and industries worldwide have readily adapted to the digital revolution by continuing to reconfigure both hardware and software dependencies as technology progresses. The digital revolution has by no means reached its summit— institutions must constantly keep up-to-date with technological progress. The library community presently faces a period of transition where the reevaluation of the theory and practice of cataloging must be reconsidered, and perhaps even drastically altered. In particular, the MARC record format was conditioned

to describe *physical* objects, i.e., a text, a map, a sound recording device. However, the digital revolution has created innumerable resources that have no physical properties—they just exist within the internet ether. Libraries today are faced with the daunting challenge of not only cataloging new forms of resources, but also rethinking the standards they have successfully practiced for a number of years. Moreover, due to the global scope of the Internet environment, revisions to standards must be agreed upon on a domestic and international level.

Undoubtedly, eXtensible Markup Language (XML) will continue to have an enormous impact on the World Wide Web infrastructure. Already greatly influencing business and government information professionals, the library environment must make amends to adapt to the flexible syntax. It is also important to consider the implications XML will have on the future of the library discipline. If libraries do not readily adapt to XML, and if XML and its related technologies continue to rapidly progress—drastically affecting the ways in which information is both created and retrieved—the library environment is in jeopardy of excluding itself from distinguished information and research professions. With the MARC record and its costly proprietary configurations and arcane structure, the library environment has already segregated itself from information professionals for years. And now with Internet infrastructure enabling world wide access of catalogs and holdings, the revision of library cataloging theory and practice is imminent. If the library community were to continue with MARC encoding, the future of the discipline as an authoritative means of access to information access is suspect.

With XML being an underlining protocol, there are many new forms of metadata that have evolved that are suitable for library cataloging. In particular, the Dublin Core Metadata Initiative, Encoded Archival Description, Text Encoding Initiative, and the Functional Requirements for Bibliographic Resources are all playing major roles in the revision of library cataloging standards. As all these new forms of metadata continue to progress, the library environment must track these technologies, evaluate their effectiveness and, before long, consider migration or reformatting.

The MARC record format, though practical, resourceful and productive for over forty years, is nearing obsolescence. As new information objects continue to sprout and costs continue to wan on storage, memory and processing power, MARC and AACR2 as a reliable and effective protocol for bibliographic control and description continues to fade. Before other metadata schemes fully mature, it is now imperative for the library community to evaluate the migration or an overall reformatting of the MARC record format.

Recommendations

Due to the impact information technology is having on the theory and practice of library cataloging, the purpose of this research project was to explore, introduce, and suggest recommendations for new metadata standards. Obviously, a suggestion to reformat or migrate an entire collection of records encoded by the MARC format to another metadata scheme is easier said than done. Out of the scope of this project is the actual migration and conversion process and what specific economic, labor, and training factors would come into play. Additional research on numerous institutions' metadata

migration would be resourceful and practical for libraries that are considering undertaking a catalog conversion project. A researcher could explore an institution's undertaking of a particular metadata schemes and what positive and negative factors came to fruition. A survey approach could be employed to discover and analyze different approaches conducted by libraries when migrating or reformatting catalog records. This could provide a template for the conversion process, and warnings and lessons learned. Individual metadata schemas could be examined, as well as the software and hardware reconfigurations necessary. A research project that fully explores the process of catalog conversion would be an invaluable resource for the library community.

As this research project fully examined the impact information technology is having on library cataloging, another recommendation for additional research would be for a researcher to investigate the opposite—in other words, what impact library catalog—in particular subject and name authority files—could have on the further development and construction of the Semantic web. Inundated by web documents without structured form or metadata, the development of the Semantic web could greatly benefit from the ways in which the library community has for years deployed controlled vocabulary sets to create unique, authoritative records.

APPENDIX A
Example of a MARC Record

000 01494cam 22003134a 450
001 13361323
005 20040226152835.0
008 031003s2003 gw a b 101 0 eng
906 __ |**a** 7 |**b** cbc |**c** orignew |**d** 1 |**e** ocip |**f** 20 |**g** y-gencatlg
925 0_ |**a** acquire |**b** 2 shelf copies |**x** policy default
955 __ |**a** pc17 2003-10-03 RUSH to ASCD |**c** jg12 2003-10-06 to tech1 |**d** jf09 2003-10-06 to sl |**e** jf12 2003-10-07 to Dewey |**a** aa20 2003-10-08 |**a** ps13 2003-12-17 1 copy rec'd., to CIP ver. |**a** jf00 2004-01-15; |**f** jf04 2004-02-26 to S.L. |**e** jf12 2004-02-26 to BCCD
010 __ |**a** 2003059239
020 __ |**a** 3540203613 (softcover : alk. paper)
040 __ |**a** DLC |**c** DLC |**d** DLC
042 __ |**a** pcc
050 00 |**a** QA76.76.H94 |**b** R88 2003
082 00 |**a** 006.7/4 |**2** 22
111 2_ |**a** RuleML 2003 |**d** (2003 : |**c** Sanibel, Fla.)
245 10 |**a** Rules and rule markup languages for the Semantic Web : |**b** Second International Workshop, RuleML 2003, Sanibel Island, FL, USA, October 20, 2003 : proceedings / |**c** Michael Schroeder, Gerd Wagner (eds.).
260 __ |**a** Berlin ; |**a** New York : |**b** Springer, |**c** c2003.
300 __ |**a** 171 p. : |**b** ill. ; |**c** 24 cm.
440 _0 |**a** Lecture notes in computer science, |**x** 0302-9743 ; |**v** 2876
500 __ |**a** Held in conjunction with the International Semantic Web Conference
504 __ |**a** Includes bibliographical references and index.
650 _0 |**a** Document markup languages |**v** Congresses.
650 _0 |**a** Semantic Web |**v** Congresses.
650 _0 |**a** Metadata |**x** Standards |**v** Congresses.
700 1_ |**a** Schroeder, Michael.
700 1_ |**a** Wagner, G. |**q** (Gerd), |**d** 1957-

APPENDIX B
An Example of an XML Record

Taken from http://www.W3C.org

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

- <CATALOG>
- <PLANT>
  <COMMON>Bloodroot</COMMON>
  <BOTANICAL>Sanguinaria canadensis</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$2.44</PRICE>
  <AVAILABILITY>031599</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Columbine</COMMON>
  <BOTANICAL>Aquilegia canadensis</BOTANICAL>
  <ZONE>3</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$9.37</PRICE>
  <AVAILABILITY>030699</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Marsh Marigold</COMMON>
  <BOTANICAL>Caltha palustris</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Mostly Sunny</LIGHT>
  <PRICE>$6.81</PRICE>
  <AVAILABILITY>051799</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Cowslip</COMMON>
  <BOTANICAL>Caltha palustris</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$9.90</PRICE>
  <AVAILABILITY>030699</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Dutchman's-Breeches</COMMON>
  <BOTANICAL>Dicentra cucullaria</BOTANICAL>
  <ZONE>3</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$6.44</PRICE>
  <AVAILABILITY>012099</AVAILABILITY>
    </PLANT>
```

```
- <PLANT>
  <COMMON>Ginger, Wild</COMMON>
  <BOTANICAL>Asarum canadense</BOTANICAL>
  <ZONE>3</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$9.03</PRICE>
  <AVAILABILITY>041899</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Hepatica</COMMON>
  <BOTANICAL>Hepatica americana</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$4.45</PRICE>
  <AVAILABILITY>012699</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Liverleaf</COMMON>
  <BOTANICAL>Hepatica americana</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$3.99</PRICE>
  <AVAILABILITY>010299</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Jack-In-The-Pulpit</COMMON>
  <BOTANICAL>Arisaema triphyllum</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$3.23</PRICE>
  <AVAILABILITY>020199</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Mayapple</COMMON>
  <BOTANICAL>Podophyllum peltatum</BOTANICAL>
  <ZONE>3</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$2.98</PRICE>
  <AVAILABILITY>060599</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Phlox, Woodland</COMMON>
  <BOTANICAL>Phlox divaricata</BOTANICAL>
  <ZONE>3</ZONE>
  <LIGHT>Sun or Shade</LIGHT>
  <PRICE>$2.80</PRICE>
  <AVAILABILITY>012299</AVAILABILITY>
    </PLANT>
- <PLANT>
```

```
  <COMMON>Phlox, Blue</COMMON>
  <BOTANICAL>Phlox divaricata</BOTANICAL>
  <ZONE>3</ZONE>
  <LIGHT>Sun or Shade</LIGHT>
  <PRICE>$5.59</PRICE>
  <AVAILABILITY>021699</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Spring-Beauty</COMMON>
  <BOTANICAL>Claytonia Virginica</BOTANICAL>
  <ZONE>7</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$6.59</PRICE>
  <AVAILABILITY>020199</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Trillium</COMMON>
  <BOTANICAL>Trillium grandiflorum</BOTANICAL>
  <ZONE>5</ZONE>
  <LIGHT>Sun or Shade</LIGHT>
  <PRICE>$3.90</PRICE>
  <AVAILABILITY>042999</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Wake Robin</COMMON>
  <BOTANICAL>Trillium grandiflorum</BOTANICAL>
  <ZONE>5</ZONE>
  <LIGHT>Sun or Shade</LIGHT>
  <PRICE>$3.20</PRICE>
  <AVAILABILITY>022199</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Violet, Dog-Tooth</COMMON>
  <BOTANICAL>Erythronium americanum</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$9.04</PRICE>
  <AVAILABILITY>020199</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Trout Lily</COMMON>
  <BOTANICAL>Erythronium americanum</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$6.94</PRICE>
  <AVAILABILITY>032499</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Adder's-Tongue</COMMON>
```

```
  <BOTANICAL>Erythronium americanum</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$9.58</PRICE>
  <AVAILABILITY>041399</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Anemone</COMMON>
  <BOTANICAL>Anemone blanda</BOTANICAL>
  <ZONE>6</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$8.86</PRICE>
  <AVAILABILITY>122698</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Grecian Windflower</COMMON>
  <BOTANICAL>Anemone blanda</BOTANICAL>
  <ZONE>6</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$9.16</PRICE>
  <AVAILABILITY>071099</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Bee Balm</COMMON>
  <BOTANICAL>Monarda didyma</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$4.59</PRICE>
  <AVAILABILITY>050399</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Bergamot</COMMON>
  <BOTANICAL>Monarda didyma</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$7.16</PRICE>
  <AVAILABILITY>042799</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Black-Eyed Susan</COMMON>
  <BOTANICAL>Rudbeckia hirta</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Sunny</LIGHT>
  <PRICE>$9.80</PRICE>
  <AVAILABILITY>061899</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Buttercup</COMMON>
  <BOTANICAL>Ranunculus</BOTANICAL>
```

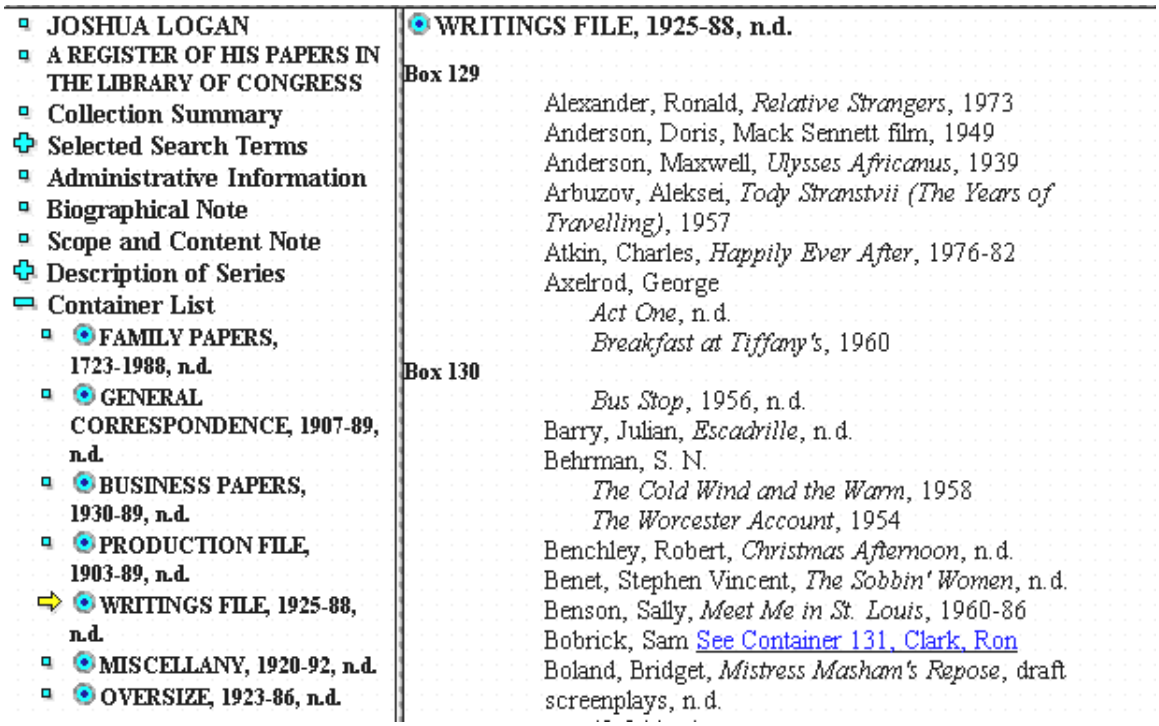
```
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$2.57</PRICE>
  <AVAILABILITY>061099</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Crowfoot</COMMON>
  <BOTANICAL>Ranunculus</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$9.34</PRICE>
  <AVAILABILITY>040399</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Butterfly Weed</COMMON>
  <BOTANICAL>Asclepias tuberosa</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Sunny</LIGHT>
  <PRICE>$2.78</PRICE>
  <AVAILABILITY>063099</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Cinquefoil</COMMON>
  <BOTANICAL>Potentilla</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$7.06</PRICE>
  <AVAILABILITY>052599</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Primrose</COMMON>
  <BOTANICAL>Oenothera</BOTANICAL>
  <ZONE>3 - 5</ZONE>
  <LIGHT>Sunny</LIGHT>
  <PRICE>$6.56</PRICE>
  <AVAILABILITY>013099</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Gentian</COMMON>
  <BOTANICAL>Gentiana</BOTANICAL>
  <ZONE>4</ZONE>
  <LIGHT>Sun or Shade</LIGHT>
  <PRICE>$7.81</PRICE>
  <AVAILABILITY>051899</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Blue Gentian</COMMON>
  <BOTANICAL>Gentiana</BOTANICAL>
  <ZONE>4</ZONE>
```

```
  <LIGHT>Sun or Shade</LIGHT>
  <PRICE>$8.56</PRICE>
  <AVAILABILITY>050299</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Jacob's Ladder</COMMON>
  <BOTANICAL>Polemonium caeruleum</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$9.26</PRICE>
  <AVAILABILITY>022199</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Greek Valerian</COMMON>
  <BOTANICAL>Polemonium caeruleum</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$4.36</PRICE>
  <AVAILABILITY>071499</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>California Poppy</COMMON>
  <BOTANICAL>Eschscholzia californica</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Sun</LIGHT>
  <PRICE>$7.89</PRICE>
  <AVAILABILITY>032799</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Shooting Star</COMMON>
  <BOTANICAL>Dodecatheon</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Mostly Shady</LIGHT>
  <PRICE>$8.60</PRICE>
  <AVAILABILITY>051399</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Snakeroot</COMMON>
  <BOTANICAL>Cimicifuga</BOTANICAL>
  <ZONE>Annual</ZONE>
  <LIGHT>Shade</LIGHT>
  <PRICE>$5.63</PRICE>
  <AVAILABILITY>071199</AVAILABILITY>
    </PLANT>
- <PLANT>
  <COMMON>Cardinal Flower</COMMON>
  <BOTANICAL>Lobelia cardinalis</BOTANICAL>
  <PRICE>$3.02</PRICE>
  <AVAILABILITY>022299</AVAILABILITY>
```

```
</PLANT>
</CATALOG>
```

APPENDIX C
Example of an EAD Record

Taken from: http://www.loc.gov/ead/tglib1998/tlindex.html



APPENDIX D
Dublin Core Metadata Elements
Taken from http://dublincore.org

Element Name: Title

Label: Title

Definition: A name given to the resource.

Comment: Typically, Title will be a name by which the resource is formally known.

Element Name: Creator

Label: Creator

Definition: An entity primarily responsible for making the content of the resource.

Comment: Examples of Creator include a person, an organization, or a service. Typically, the name of a Creator should be used to indicate the entity.

Element Name: Subject

Label: Subject and Keywords

Definition: A topic of the content of the resource.

Comment: Typically, Subject will be expressed as keywords, key phrases or classification codes that describe a topic of the resource. Recommended best practice is to select a value from a controlled vocabulary or formal classification scheme.

Element Name: Description

Label: Description

Definition: An account of the content of the resource.

Comment: Examples of Description include, but is not limited to: an abstract, table of contents, reference to a graphical representation of content or a free-text account of the content.

Element Name: Publisher

Label: Publisher

Definition: An entity responsible for making the resource available

Comment: Examples of Publisher include a person, an organization, or a service. Typically, the name of a Publisher should be used to indicate the entity.

Element Name: Contributor

Label: Contributor

Definition: An entity responsible for making contributions to the content of the resource.

Comment: Examples of Contributor include a person, an organization, or a service. Typically, the name of a Contributor should be used to indicate the entity.

Element Name: Date

Label: Date

Definition: A date of an event in the lifecycle of the resource.

Comment: Typically, Date will be associated with the creation or availability of the

resource. Recommended best practice for encoding the date value is defined in a profile of ISO 8601 [W3CDTF] and includes (among others) dates of the form YYYY-MM-DD.

Element Name: Type

Label: Resource Type

Definition: The nature or genre of the content of the resource.

Comment: Type includes terms describing general categories, functions, genres, or aggregation levels for content. Recommended best practice is to select a value from a controlled vocabulary (for example, the DCMI Type Vocabulary [DCT1]). To describe the physical or digital manifestation of the resource, use the FORMAT element.

Element Name: Format

Label: Format

Definition: The physical or digital manifestation of the resource.

Comment: Typically, Format may include the media-type or dimensions of the resource. Format may be used to identify the software, hardware, or other equipment needed to display or operate the resource. Examples of dimensions include size and duration. Recommended best practice is to select a value from a controlled vocabulary (for example, the list of Internet Media Types [MIME] defining computer media formats).

Element Name: Identifier

Label: Resource Identifier

Definition: An unambiguous reference to the resource within a given context.

Comment: Recommended best practice is to identify the resource by means of a string or number conforming to a formal identification system. Formal identification systems include but are not limited to the Uniform Resource Identifier (URI) (including the Uniform Resource Locator (URL)), the Digital Object Identifier (DOI) and the International Standard Book Number (ISBN).

Element Name: Source

Label: Source

Definition: A Reference to a resource from which the present resource is derived.

Comment: The present resource may be derived from the Source resource in whole or in part. Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

Element Name: Language

Label: Language

Definition: A language of the intellectual content of the resource.

Comment: Recommended best practice is to use RFC 3066 [RFC3066] which, in conjunction with ISO639 [ISO639]), defines two- and three-letter primary language tags with optional subtags. Examples include "en" or "eng" for English, "akk" for Akkadian", and "en-GB" for English used in the United Kingdom.

Element Name: Relation

Label: Relation

Definition: A reference to a related resource.

Comment: Recommended best practice is to identify the referenced resource by means of a string or number conforming to a formal identification system.

Element Name: Coverage

Label: Coverage

Definition: The extent or scope of the content of the resource.

Comment: Typically, Coverage will include spatial location (a place name or geographic coordinates), temporal period (a period label, date, or date range) or jurisdiction (such as a named administrative entity). Recommended best practice is to select a value from a controlled vocabulary (for example, the Thesaurus of Geographic Names [TGN]) and to use, where appropriate, named places or time periods in preference to numeric identifiers such as sets of coordinates or date ranges.

Element Name: Rights

Label: Rights Management

Definition: Information about rights held in and over the resource.

Comment: Typically, Rights will contain a rights management statement for the resource, or reference a service providing such information. Rights information often encompasses Intellectual Property Rights (IPR), Copyright, and various Property Rights. If the Rights element is absent, no assumptions may be made about any rights held in or over the resource

APPENDIX E

Authors Main: Additional:	Title:
Journal:	Publisher:
Source:	Subject heading assigned:
Article Type: • MARC/AACR2 1. Inherent Weakness w/ “modern” documents 2. Reformatting to XML 3. Working with Dublin Core • XML 1. Advantages over MARC 2. Mapping with other metadata • Migration/Interoperability 1. OPAC and web based documents 2. Woking w/ XML 3. Z39.50 • Digital/Special Collection 1. EAD 2. Dublin Core 3. Mapping w/ XML	Notes:

NAME INDEX

SUBJECT INDEX

Works Cited

Akin, Jim. "Amazon Everywhere." *PC Magazine.* New York: Ziff Davis Media, September 2003.

Bush, Vannevar. "As We May Think." Available: <http:www//ps.unisb.de/~duchier/pub/vbush/> Accessed March 1, 2004.

Caplan, Priscilla. *Metadata Fundamentals for All Librarians.* Chicago: American Libraries Association, 2003.

Clark S. *Back to Basics Part I Meta Tags.* Available at <http://www.webdeveloper.com/html/html metatags.html> Accessed March 2, 2004.

Cutter, Charles. *Rules for a Printed Dictionary Catalog.* United States: Rowman and Littlefield, 1970.

December, John and Neil Randall. *The World Wide WebUnleashed 1996.* Indianapolis: Sams, 1995.

"Defining SGML Document Structures." 24 Feb. 2004. <http://www.isgmlug.org/sgmlhelp/g-sg14.htm>

DeRose, Stephen. "XML and TEI." *Computers and the Humanities 33:* 11-30, 1999.

Furrie, Betty. *Understanding MARC Bibliographic: Machine-readable cataloging.* Washington D.C.: Library of Congress, 2003.

Garfield, Eugene. "From Sputnik to the World Wide Web-- A Retrospective View of Citation Indexing"<http://www.garfield.library.upenn.edu/papers/acrl2001.html> Accessed Februrary 25, 2004.

Gorman, Michael. *The Concise AACR2, 1998 Revision.* Chicago: American Libraries Association, 1998.

Hudgins, Jean. *Getting Mileage Out of Metadata: Applications for the Library.* American Library Associations : Chicago, 1999.

Haynes, David. *Metadata for information management and retrieval.* London: facet, 2001.

Lazinger, Susan, S. *Digital Preservation and Metadata: History, theory, Practice.* Libraries Unlimited: Engelwood, Colo, 2002.

Longman, Addisson Wesley. "Chapter 2: A History of HTML."

<http://www.w3.org/People/Raggett/book4/ch02.html> Accessed February 21, 2004.

Matthews, Robert. "A Question of Semantics." *Academic Search Primer*. Accessed 21 February 2004.

Mercer, Dave. *XML: A Beginners Guide.* New York: McGRaw Hill, 2001.

Pass, Gregory A. *Descriptive Cataloging of Ancient, Medieval, Renaissance, and Early Modern Manuscripts*. Chicago : Association of College and Research Libraries, 2002.

Poe, Stephan. "XML in the Publishing World." *E-Doc Magazine.* September/October 2003.

Ramiscal, Noel Guivani. "The Nature and Functions of Meta-tags: Covert Infringement of Trademarks and Other Isssues." *Prometheus*. Mar2000, Vol. 18 Issue 1, p39-48.

Ray, Erik T. *Learning XML.* Beijing: O'Reilly, 2001.

Singh, S. "Business-to-Manufacturing Integration Using XML." *Hydrocarbon Processing.* Ebsco Publishing, 2003.

Sperburg-McQueen, C.M. and Lau Burand. "A Gentle Introduction to SGML." *Guidelines for Electronic Text Encoding and Interchange.* TEI P3. <http://www-sul.stanford.edu/tools/tutorials/html2.0/gentle.html> Accessed 24 Feb. 2004.

Strum, Jake. *Developing XML Solutions.* Redmond, Washington: Microsoft Press, 2000.

Taylor, Arlene G. *The organization of Information.* Englewood, Co.: Libraries Unlimited, 1999.

Text Encoding Initiative. TEI Consortium. 22 Feb. 2004. <http://www.tei-c.org/> . *World Wide Web Consortium.* W3C. 22 Feb. <http://www.w3.org/>

XML From the Inside Out. Oreilly.com. 4 March 2004. http://www.xml.com/