University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, UNL Libraries

Libraries at University of Nebraska-Lincoln

The Changing Landscape of Contemporary Cataloging

Sue Ann Gardner
University of Nebraska-Lincoln, sgardner2@unl.edu

The Changing Landscape
of Contemporary Cataloging

Sue Ann Gardner

ABSTRACT. Intended to contribute to the current dialogue about how the emerging information environment is impacting cataloging issues, this survey paper covers a broad range of topics, such as how search engines compare with integrated library systems, and includes some thoughts on how cataloging processes may evolve to continue to remain relevant. The author suggests that there is a need for significant changes in integrated library system interfaces and infrastructures as well as some changes in cataloging practice. The value of descriptive vs. non-descriptive elements in the catalog record and some pros and cons of the MARC format are covered. doi:10.1300/J104v45n04_06 *[Article copies available for a fee from The Haworth Document Delivery Service: 1-800-HAWORTH. E-mail address: <docdelivery@haworthpress.com> Website: <http://www. HaworthPress.com> © 2008 by The Haworth Press. All rights reserved.]*

KEYWORDS. Future of cataloging, online catalogs

INTRODUCTION:
CATALOGS, CATALOGING AND THE WEB

Cataloging and catalogs are poised to experience a seismic shift in the near future. Because patrons are increasingly seeking information through search engines like Google, one potential goal in libraries is to re-tool the integrated library system (ILS) so that it will be the preferred

Sue Ann Gardner is Cataloging and Metadata Librarian, 322 Love Library, University of Nebraska-Lincoln, Lincoln, NE 68588-4100 (E-mail: sgardner2@unl.edu).

gateway to information discovery among patrons both in-house and elsewhere.[1,2] With the excellent quality of, and constant improvements to, search engines, catalogers are being forced to reconsider their role in today's information environment. This is a pressing concern because, in response to these developments outside the library profession, declarations are coming from authoritative sources about how we need to adjust our cataloging practices immediately, such as ceasing tracing of series information,[3] seemingly whether those decisions are in the best interests of resource discovery or not.

Since the advent of modern cataloging over 100 years ago, catalogers have created surrogate records for the items held in, or accessed through, a library. The surrogate offers a description, and a classification number, and authority-controlled subject terms to assist with collation of like items. The call number, which contains the classification number, determines the physical location of a piece in a collection, thereby physically collating items by primary subject. Catalog records provide authority-controlled name terms (personal, corporate, geographic, and meeting). In the United States since the 1980s, most of these records have been created using the machine readable cataloging (MARC) format, which includes a plethora of coded fields which are not part of the traditional international standard bibliographic description (ISBD) format but potentially may be manipulated in the electronic environment.

As remote-access digital products became increasingly available in the early to mid-1990s, the discussion in libraries turned to "access versus ownership," and catalogers expanded their cataloging to provide access to items that were not physically held. In the past several years, cataloging has also realigned to include methods for describing an ever-increasing number of non-text digital products such as streaming video, streaming audio, remote access still images, continually updated Web sites, geographic data sets, and other formats.

Additionally, new manifestations of older works now appear more frequently due to renewed user demand, as is described in Anderson's "The Long Tail."[4] The concept of "out of print" is mutating. In all, the changing economic models in publishing and distribution will affect catalogers' work. Just as a book no longer needs a substantial local readership in order to justify its purchase, our local patrons may be treated to access to ever-individualized resources, and the bibliographic gateways that we choose to employ will reflect this.

With the arrival and rapid evolution of the Web, the novel *Flatland* comes to mind, in which the protagonist is a two-dimensional figure. He first describes the social structure of Flatland as though it were absolute

and unchanging. Then he is lifted out of his two-dimensional milieu, into three dimensions, and that causes him to understand that there is more to his world than he can see.[5] The Web feels like this. It is certainly not linear, and the depths are so deep and the relationships are so variously disconnected and entwined and unexpected, that there is no analog to the information milieu of even the early 1990s.[6]

CATALOGS AND CATALOGERS: NEW ROLES OR OBSOLESCENCE?

Library catalogs are just one aspect of this vast digital store of human creation. Where do catalogs fit into the new environment? What will our role as catalogers be in the digital era? Since catalogers produce surrogates of items, with endeavors underway such as Google's Book Search Library Project in which the full text of millions of texts are being digitized,[7] a more pointed question is: how is surrogate creation relevant in today's information environment? And, assuming that some of what catalogers do is still relevant, what about it, or which aspects of it, makes it relevant? How do we need to re-think our approach to resource discovery?

Perhaps it has not reached the level of consensus, but there is growing concern that the successor to AACR2,[8] *Resource Description and Access* (RDA), is in need of a fresh conceptual basis to reflect the changing nature of information-seeking behavior, and to take advantage of new technologies and new approaches to database manipulation. Karen Coyle, digital library consultant, and Diane Hillman, of Cornell, in "Resource Description and Access: Cataloging Rules for the 20th Century," contend that RDA is poised to "keep us rooted firmly in the 20th, if not the 19th century." They note that cataloging rules were created in the context of a card catalog, and in an era when texts were the primary format cataloged.[9] Michael Gorman has suggested that the tool that will succeed AACR2 should be abbreviated such that only a basic cataloging framework is presented, irrespective of format, and that guides for cataloging various formats (such as cartographic materials, sound recordings, etc.) be created separately by specialists in the field as interpretations of the basic framework.[10] Through all of this, our basic tenet still holds, namely, "The highest principle for the construction of cataloguing codes should be the convenience of the users of the catalogue."[11]

Many other librarian-authors have weighed in on the changing-nature-of-discovery discussion, including Karen Calhoun of OCLC,[12]

Deanna Marcum of the Library of Congress,[13] Thomas Mann of the Library of Congress,[14,15,16] the librarians at the Indiana University library,[17] the Bibliographic Services Task Force of the University of California Libraries,[18] and Karen G. Schneider of ALA *TechSource*,[19,20,21] just to name a few. Though the authors all acknowledge that a new information age is upon us, and they assert that librarians want to and must remain relevant in today's fast-changing digital milieu, there are generally two lines of thought. Marcum,[22] Calhoun,[23] and others[24] assert that much of traditional cataloging has already become obsolete and that librarians must adopt radical new practices to ensure that a libraries' products and services remain relevant into the information age. As of this writing they have not offered many specifics about what cataloging should entail, and their focus has generally been on the novice information seeker. Schneider,[25,26,27] Markey,[28] Mann,[29,30,31] and others[32] believe that the products of traditional cataloging have value in the current information market and may be harnessed to serve information seekers now and into the future. Moreover, Mann believes that our niche must include the advanced user, the scholar, as opposed to the novice.[33]

SEARCHING IN PUBLIC ACCESS CATALOGS
vs. WEB SEARCH ENGINES

Karen Schneider, of ALA *TechSource*, recently wrote a three-part blog addressing many of the catalog-related issues of the day. In one part she has a list of "features your OPAC wishes it had," including relevance ranking, stemming, field weighting, spell checking, flexible default query processing, faceting, and several other capabilities.[34,35,36] This points to the fact that some of the changes that are required need to be made not to cataloging *per se* but to the infrastructure of our catalogs and ILSs, such as those provided by Innovative Interfaces, Ex Libris, and others.[37] Innovative Interfaces, for example, has lately incorporated some of the above-listed capabilities into their system. Other ILS vendors are surely honing their products and services likewise, and other companies, like Endeca,[38,39] are focusing on these issues full bore.

As an ILS user, however, it is clear to me that these systems still have a long way to go to offer the information seeker a seamless search experience. Until I can put a sophisticated search string into a search box in an ILS and be taken to a link to the full text of the article that I want, or to the bibliographic record I want, I will continue to use Google to look for certain resources, especially those that I know exist. If I know that I

want to access Schneider's *ALA TechSource* blog, I enter "Schneider 'ALA TechSource'" in a Google search box and one issue of it is the first link that comes up. Clicking on that link quickly takes me to any of her blogs on *TechSource*. If I know that I want to read Markey's "The Online Library Catalog: Paradise Lost and Paradise Regained?" from *D-Lib Magazine*, I put the article title, in quotes, into the Google text box and the first link takes me to the full text of the article. I can find full text articles from many peer-reviewed publications via Google, free of charge, as information commons increasingly serve as the platform for journal digitization projects, and all with one quick, knowledgeable entry in its text box. When this method does not take me to the full text, I often use Google to get a relevant citation and then go to my ILS to seek out a specific resource that is not available free on the Web.

Google, in conjunction with the ILS, can serve the advanced user in many instances. Does Google serve the novice user nearly as well? Probably not, but then neither does a standalone ILS. As long as information seekers know to, and are able to, quickly toggle between a search engine and their local ILS, this arrangement may be a healthy, symbiotic one. It is important to note that a Google search, as of now, will not provide an information seeker with access to the products of the scholarly "deep Web." So, one goal might be to entice the user to remain longer in the local ILS, so that he or she may be able to search that portion of the Web in conjunction with the non-digital resources he or she will readily be pointed to. If this is the goal, then how can we meet it with a re-working of ILSs, catalogs, and cataloging? If Google is the starting point for most information seekers now, is there a way to entice them to enter the ILS *first*, and go out to a source like Google, from there?[40] Or should we, as Schneider suggests, attempt to "dis-integrate the catalog, weave it into the Web, and push forward to the future"?[41]

At least one endeavor is underway to attempt to sidestep the limitations that vendors impose on search capabilities, namely, the University of Rochester's eXtensible Cataloging (XC) Project. From their "About" page, they report that: "The University of Rochester's River Campus Libraries is studying how best to develop an open-source online system that can unify access to traditional and digital library resources . . . the University will begin planning and requirements analysis activities for a new system known as eXtensible Catalog (XC)."[42] It is ambitious, and it will be interesting to see how they progress.

THE PURPOSES OF THE LIBRARY CATALOG

Catalogs have traditionally been used for both inventory and discovery. Schneider suggests that cataloging should primarily serve the needs of users to discover resources and, secondarily, local inventory purposes.[43] While this may be true in the abstract, I believe that libraries will necessarily maintain their local nature so their ILSs should retain their inventory functions. After all, libraries comprise "collections" of owned or accessed items that are selected for specific user populations. Just because we can offer access to so much more than we could even two decades ago does not obviate the need to serve the population that sits under the umbrella of the institution that, frankly, pays for access to all of these resources. In short, cataloging is an activity bound by the financial support of the parent institution. Fiduciary responsibility extends to inventory of collections, virtual or otherwise. Where I agree with Schneider on this point is that the inventory function of the ILS should be covert, and not interfere with searching.

WHAT DO WE CATALOG, AND WHO DO WE CATALOG FOR?

With remote access to resources so readily available, it is tempting to think that our patron base is growing ever larger in each library, though I think that bibliographers and administrators need to continue to make a strong case for "collecting" for their local population. It is true that libraries and, indeed, all information seekers are more connected than ever, though I do not see that our local population base is increasing dramatically. As long as a library is connected to a parent institution such as a university, municipality, or company, our role as librarians is that our collection development and, thus, our cataloging continue to be relevant for that physical community.

If Hypothetical Midwest University has an agriculture college, the librarians there will continue to "select" resources for the agriculture faculty and students. It is both our traditional and current role to bring together resources that are of use and interest to our readers. This is why I think that librarians, and catalogers more specifically, will have a place in the future information age. Users certainly have greater and greater access to resources on their own and, if they want to and are able to, they can navigate information sources independent of a library setting. Ever since libraries have had open stacks, and since media have

become freely available for sale, this has been the case. But what librarians do is slog through much of what is available and expertly select broadly and deeply for a known user populations. Without this selection process, our patrons are left to select and navigate on their own. We do some of the selecting work for them so that they are free to focus on research, learning, or whatever their intention is, instead of spending time on gathering and organizing information resources.

What is fantastic about the current environment is that a researcher, for example, whose line of inquiry is not a main thrust at a university, can search the catalog and collection at a library whose parent institution does have that emphasis. So we acquire some new patrons in that way, but this is purely incidental and, due to economic realities, cannot be our focus. Just as is true for interlibrary loan borrowers, we do not select for these incidental patrons, they select for us. What we need to do in our libraries is advertise what our areas of focus are, if we have any, and garner a new patron base in that way.

Inventory needs aside, resource discovery is the essential function of the ILS, it is true, and Lorcan Dempsey, of OCLC, in his blog, notes that the user experience is predicated on "ranking, relating and recommending," much as is done on sites such as Amazon.com. He notes that ILSs would do well to incorporate these capabilities to enhance resource discovery and make using the ILS more enjoyable for users.[44]

While librarians need to keep selecting and cataloging for local populations, we need to re-work the way we format the metadata in the catalog record, we need to present the product of our metadata creation in a more user-friendly way, and we need to improve searching infrastructure.

WHAT MAKES A CATALOG RECORD USEFUL IN THE AGE OF FULL TEXT?

After reading many of the recent forward-looking expositions on these issues, I am left wondering: What will my work as a cataloger be like one, five, or ten years from now? How will creating catalog records differ from how I have done it the past 18 years? Because, if it does not actually suffer from obsolescence, it is sure to change.

To answer this, consider: What do catalog records offer that a full text scan of a text does not offer? This is salient because digital full text, as the Google Book Search Project is providing for millions of items in several libraries, in some ways offers a superior search tool to any cata-

log record-surrogate. In fact, Google asserts that their scans represent a sort of surrogate of the originals. That claim, apparently, is actually the basis of their argument in their lawsuit to sidestep copyright issues. Google insists that it is, in effect, creating a "catalog" in their scanning of full texts.[45] This seems, to me, conceptually a stretch if not patently absurd. Eric Morgan, in his "A 'next generation' library catalog," adeptly notes that *data* (i.e., full texts) are distinct from *meta*data (i.e., catalog records, and other surrogates).[46]

To answer the question, then: What do catalog records offer that the electronic full text can not? The list comprises primarily non-descriptive bibliographic elements, ones that require analysis and intellectual input on the part of the cataloger, such as:

- classification number
- authority-controlled subject headings
- authority-controlled names (personal, corporate, conference, and geographic) and titles (uniform and series)
- other/alternate title information.

Additionally, most current cataloging employs:

- the MARC format, in which the metadata in each field in each record is highly reliable and transferable (a title is a title)
- MARC coded fields that correspond to relevant data about a resource, its format, etc.
- AACR2, which gives us well-established rules for cataloging/metadata creation.

It stands to reason that these cataloger-supplied elements and conditions are well-suited to augment the content of full texts in the retrieval process, though not everyone agrees with this assumption. Nancy Fallgren, of the University of Maryland, contends that "[t]he traditional bibliographic access points of author, title, and subject now constitute a small proportion of the data that can be retrieved with full-text keyword searching."[47] However, as Karen Markey points out, the result of the analysis that is proffered to provide cataloger-supplied metadata such as those listed above should place those cataloger-supplied elements at a **higher rank** in search algorithms than randomly-derived terms from deep within texts.[48] Declaring that the traditional access points constitute a small proportion of the data/metadata is like dismissing diamonds

because they constitute just a small proportion of the slurry in which they are found. They may represent but a fraction, but they are precious bits.

So, if descriptive cataloging can be modified to include portions of content from a full text, how do the cataloger-supplied, non-descriptive catalog record elements aid in searching? Aside from improved keyword searching, an answer seems to be, in part: **collation**, and **search result precision**. Topical collation, referring to the grouping of items on similar topics, is enhanced due to the use of authority-controlled subject headings and classification numbers. As is true now, papers, books, blogs, musical compositions, etc., by the same author will also all be found together if the names are authority-controlled. The lack of any attempt at authority control in the current metadata infrastructure of popularly-used search engines is partly what leads to incomplete, improperly ranked, or irrelevant results. Dempsey points out that inconsistent application of authority control, as in a large, combined database, nullifies its benefits.[49] This may be true, though if catalogers' time is freed such that we are no longer laboring over the transcription of descriptive elements, we could devote more time to controlled element input and revision.[50] Some automated authority control algorithms could play an important role in this endeavor as well.

Catalogers could begin to include additional metadata elements in catalog records to aid in resource discovery. For example, Markey suggests that we may offer what she calls "qualification metadata" that assist information-seekers in determining if a source will meet their needs. These elements would include data about whether a resource is "*In* a discipline . . . *To what extent* the author is an authority on the topic at hand . . . *of* a particular literary nature . . . *When* the particular subject took place," etc.[51] Since catalogers conduct subject and genre assessments already, assignments such as these would be just one step beyond our current purview.

SUBJECT HEADINGS AND OTHER SUBJECT-TYPE TAGS

In a 2002 study by Tina Gross and Arlene Taylor, and in one from 1997 by Henk Voorbij, it was determined that the presence of authority-controlled subject headings in bibliographic records augmented keyword searching substantially. From Gross and Taylor's abstract: "It was found that more than one-third of records retrieved by successful

keyword searches would be lost if subject headings were not present, and many individual cases exist in which 80, 90, and even 100 percent of the retrieved records would not be retrieved in the absence of subject headings."[52] In Voorbij's research, he found that nearly twice as many records are retrieved through keyword searching if controlled vocabulary is included in the records.[53]

Clay Shirky, in his blog, notes that LCSH and the Library of Congress and Dewey classification schemes are outdated and rife with bias. He argues that, now that items do not have to be categorized in order to shelve them, there is no good use for subject and classification assignment.[54] While it is inarguable that these classification schemes are biased and outdated (this is not news to catalogers; they are facts that catalogers meet head on every day in their use of these tools), their thoughtful, expert application and consistent use allows for collation of like sources. Because subject headings are assigned by people who assign them for a living, and do so as a rule ethically, expertly, and according to widely agreed-upon rules (namely, the Library of Congress subject cataloging manuals and AACR2), they are used as consistently as can possibly be expected.

What established subject headings provide, also, are scope notes, a list of not-used synonyms, broader and narrower term relationships, and *see* and *see also* references. However quaint they may be at times, and though classification numbers are often no longer needed to build a call number for the purpose of assigning an item a physical place on a shelf, classification and subject terms are a form of consistently-used "tag" which places resources with other resources on the same topic. Furthermore, LCSH, and other well-established controlled vocabularies and thesauri, are routinely updated. The updating process may need to occur even more quickly than it does currently–with some of the time spent now on descriptive cataloging spent instead on authority work–but these vocabularies grow and progress, nonetheless.

As is suggested on the LibraryThing (LT) site, in addition to cataloger-supplied authority-controlled subjects, names and titles (i.e., taxonomy), bibliographic records could also contain user-generated tags of various kinds (i.e., folksonomy).[55] These would be, necessarily, free-text, with no authority control of any kind.[56] The usefulness of user-generated tags remains to be seen. But there is potential for them to serve a role in searching. Likely they will ultimately serve as a method of sorting by genre, if nothing else. Just as inclusion of subject headings augments the full text in keyword searching, thesaurus term assignment coupled with free-text, user-supplied tag assignment is bound to result

in a more powerful combination in terms of searching than inclusion of just one tag type alone.

CATALOG RECORD ELEMENTS THAT MAY NO LONGER BE SO USEFUL

Conversely, with full texts increasingly becoming available, the descriptive portions of catalog records could conceivably be cataloger-identified, but no longer cataloger-supplied:

- physical descriptive elements, such as pagination and size (file size, particularly, is a moving target in an online environment)
- transcribed title and statement of responsibility
- edition statement
- imprint data
- much information that appears in notes in catalog records, such as pagination for bibliographical references, noting of the presence of an index, source of title, summary, tables of contents, and such.

Current cataloging practice involves transcription of these descriptive elements. I can envision tagging elements in a full text as "title," "author," "publication date," and so on, and those tag designations would translate into fields in a catalog record. No transcription would be necessary, just identification of elements in a full text source. I hasten to add that this designation, as opposed to transcription, would still be a high-level cataloging function and not readily performed by a non-cataloger. As is true currently in cataloging (and other metadata creation), it is not the transcription of descriptive elements that is the difficult task, it is the identification of those elements. Distinguishing a title from a series title from head-of-title information can be daunting. So, though the method may change, the intellectual effort that would be needed to perform this task would remain.

ADDITIONAL CONSIDERATIONS: ISBD AND FRBR

Other changes on the horizon arise from the evolving of bibliographic catalog infrastructures and interfaces. The form of the traditional ISBD-formatted organization of the catalog record (i.e., order of

elements) is already largely irrelevant. To wit, the brief displays that most ILS interfaces employ bear no resemblance to the order of elements in an ISBD-formatted record. Another increasingly antiquated concept includes that of main entry, as Rahmatollah Fattahi[57] and others[58] have noted. While the concept of main entry is applicable in some instances, such as in bibliography creation, in the context of the ILS all authors may be considered primary now that cards do not have to be filed under the first author's last name. The concepts of "chief source" and "source of title" can now be effectively modified. What we consider to be an "alternate title" under the current rules may be equally applicable to an item as what we would call the title proper. Determining the details of this will take analysis and experimentation within the infrastructure of evolving ILSs and search engines. As I inferred earlier, I suspect that these concerns will not be addressed directly in RDA.

Some of the concepts outlined in the "Functional Requirements for Bibliographical Records [FRBR] Final Report" may be incorporated into a modern cataloging framework.[59] For instance, as various digital media become the default bibliographic entity, as opposed to hard-copy texts, catalogers are less apt to need to describe aspects of the manifestation of entities, which have become more fluid attributes for digital items, such as publisher and dimensions of the carrier. For digital products, the FRBR concepts of **work** and **expression** become more significant than those of manifestation or item. The inventory function of the catalog will require us to acknowledge manifestations and items in some fashion. However, the catalog is undeniably becoming more of a gateway to resources no matter where they reside.

THE USE OF MARC AS A METADATA SCHEME FOR BIBLIOGRAPHIC RECORDS

Because MARC is so widely used, it bears consideration in this discussion. If it is not already, the MARC format is at least potentially robust for the purposes of providing a means to metadata creation that leads to resource discovery. This is an intuitive assessment based on recognition of the utter consistency that the format provides the data, thanks largely due to the well-established set of rules (currently, AACR2) that governs its application. Through ultra-consistent, human-input tagging of elements in highly structured fields, MARC provides some measure of **context**, which is one criticism lodged against the free-

for-all search for terms in a full-text source.[60] However, outside the realm of MARC, the issue of context is addressed by faceting of terms,[61] meaning "manipulating search results to make it easy to browse by category," in full text databases, as Schneider explains.[62] Though use of highly consistent tagging is not *required* for effective faceting, use of controlled vocabulary is bound to augment it, as is the case for keyword searching generally.

Coupled with well-established rules for use, such as those found in AACR2 and the forthcoming RDA, MARC inarguably provides sheer consistency, and offers rich, extensible tagging. Schemata such as Dublin Core, while sufficiently extensible, lack the associated rule system that makes MARC such a vigorous metadata scheme. MARC's other strength is that it has a veritable army of practitioners (i.e., catalogers) who are well-versed in its use, and in its vagaries. Not to be overlooked is that OCLC, the world's largest database of shared cataloging, with its evolving user services and interfaces, largely comprises records in MARC format. Inasmuch as OCLC's products and services remain relevant in information seeking, MARC will likely be a metadata player.

However, in some particulars, MARC in its current form is inadequate as a metadata format that aids in general resource discovery, as through a search engine. One issue is the common complaint that there is much repetition of metadata content in a single record. How computers recognize subfield information is another such issue. It seems to me that these concerns could be addressed with just some tweaking of MARC and not an outright overhaul. For example, several articles have been written on OCLC's Faceted Application of Subject Terminology (FAST). One such article appeared recently in *TechKnow*, by Ione T. Damasco, of the University of Dayton, in which she explains how a team at OCLC has developed a scheme for using Library of Congress subject headings (LCSH) in a way that makes them easier for a computer to read. It is a straightforward approach that any experienced subject cataloger would be able to apply after very little training. It involves deconstructing subject strings, line by line, in a systematic fashion such that subfields are eliminated while the terminology of LCSH is retained.[63,64] There have been calls for faceted classification, beyond this simple technique from others.[65]

Thomas Mann has written that faceting of subject terms may be excellent for computers, but it is not so good for people who need to see subject *strings* to make sense of the relationships the elements bear to one another.[66] The need for the systems to recognize field elements and the opposing need for humans to see full subject strings in displays

could be resolved by putting both types of subject fields in the same rec-
ord. The system algorithms could ignore the fields with the full strings
(MARC 6XX fields), though these fields would display to the public,
while the fields with the faceted terms could be manipulated by the
computer but suppressed from the public view.

   So-called "atomization" of data in the MARC record is a potential
problem with MARC in general. Fallgren mentions in her Library of
Congress Working Group background paper that in MARC authors'
names are input in a composite string, which hobbles rich name search-
ing.[67] Again, a tweaking of the format may be all that would be required
to correct this.

   Some of the tweaking of MARC could come about during conver-
sion, or mapping, from MARC to SGML (standard generalized markup
language) or XML (extensible markup language).[68,69,70,71] These con-
versions may at some point become reliably automated.[72,73] At this
point, whether the conversion/mapping, automated or not, would affect
the day-to-day work of the cataloger in the trenches remains to be seen.
How well evolving systems are able to convert MARC formatted data to
user-friendly interfaces with underlying robust searching capability
will influence how much this issue will affect our work as catalogers.
Within the framework of the discussion of the impetus to change the
ILS infrastructure, Eric Morgan, in "A 'next generation' library cata-
log," describes some of the possible avenues to data conversion.[74]

·NEW HORIZONS IN BIBLIOGRAPHIC METADATA

   In "Framework for a Bibliographic Future," Coyle, Hillman, Roch-
kind, and Weiss propose an outline for a bibliographic record construct
that transcends current methods. In their proposal, they state:

> We are interested in producing metadata that is both highly exten-
> sible and that will promote compatibility between communities
> and applications that extend the metadata. The four components
> that we propose are: a **model** of basic structures and relationships,
> a **schemata** that defines an extensible set of properties, **guidance**
> for application of the properties, and **encoding**. The·model can be
> used to create one or more schema, and any schema can be ex-
> pressed using one or more encodings. The guidance document is a
> key element that provides both direction to creators but also de-
> scribes the semantics of the data elements in a human-understand-

able way. These four components provide a basis for creation of machine-manipulable metadata that has meaning to a community yet it can be defined in a rigorous way to communicate clearly to any users of the data.[75]

Their framework is not MARC-dependent and, in fact, the authors are doubtful that MARC can be sufficiently re-worked to retain relevancy. Similar to the scenario the authors have outlined, perhaps the cataloging community will be able to make a quantum leap in the near future so that we can remain competitive in today's ever more digital world, providing access to both print sources and digital resources with equal success.

CATALOGERS HAVE SOMETHING TO OFFER IN THE FULL-TEXT ERA

In summary, catalogers are able to give added value to full texts, and we are able to provide very well organized metadata for both textual and non-textual digital products that facilitate spot-on resource discovery. The intellectual input that we offer to this process for non-textual digital resources, especially, is going to be superior to any system-supplied metadata due to the lack of text in such resources. The current ILS infrastructures and interfaces require extensive re-working. As for the bibliographic surrogate template we use, whether we adopt an entirely new framework or whether we retain MARC in combination with AACR2/ RDA, the consistency that our metadata schemes offer, coupled with . relevant bibliographic standards, will be key to spot-on resource discovery, that provides not just tremendous recall, but also precision and relevance.

NOTES

1. Karen Calhoun, "The Changing Nature of the Catalog and Its Integration with Other Discovery Tools," (2006) http://www.loc.gov/catdir/calhoun-report-final.pdf.
2. Andrew K. Pace, "Dismantling Integrated Library Systems." *Library Journal* 129, no. 2 (2004): 34-36, http://libraryjournal.com/article/CA374953.html.

3. Library of Congress, Cataloging Policy and Support Office, "Series at the Library of Congress: June 1, 2006," http://www.loc.gov/catdir/cpso/series.html.

4. Chris Anderson, "The Long Tail." *Wired* 12.10 (2004), http://www.wired.com/wired/archive/12.10/tail.html.

5. Edwin A. Abbott, *Flatland: A Romance of Many Dimensions*. (London: Seeley, 1884).

6. Michael Wesch, *Web 2.0: the Machine Is Us/ing Us*. (Manhattan, Kan.: Kansas State University, 2007), http://www.youtube.com/watch?v=6gmP4nk0EOE.

7. Google Book Search Beta, "About Google Book Search," (2007), http://books.google.com/googlebooks/about.html.

8. *Anglo-American Cataloguing Rules*, 2nd ed., rev. 2002. (Chicago: American Library Association, 2003), 750.

9. Karen Coyle and Diane Hillman, "Resource Description and Access (RDA): Cataloging Rules for the 20th Century," *D-Lib Magazine* 13, no. 1/2 (2007), http://www.dlib.org/dlib/january07/coyle/01coyle.html, doi: http://dx.doi.org/10.1045/january2007-coyle.

10. Michael Gorman, "AACR3? Not!" in *The Future of Descriptive Cataloging Rules: Papers from the ALCTS Preconference, AACR2000*. (Chicago: American Library Association, 1998).

11. International Federation of Library Associations and Institutions, "Statement of International Cataloguing Principles," (Jan. 2005 draft), http://www.loc.gov/loc/ifla/imeicc/source/statement-draft_jan05.pdf.

12. Calhoun, "The Changing Nature of the Catalog and Its Integration with Other Discovery Tools."

13. Deanna B. Marcum, "The Future of Cataloging: Address to the EBSCO Leadership Seminar, Boston Massachusetts," (Jan. 16, 2005), http://www.loc.gov/library/reports/CatalogingSpeech.pdf.

14. Thomas Mann, "What Is Going on at the Library of Congress?" (June 19, 2006), http://www.guild2910.org/AFSCMEWhatIsGoingOn.pdf.

15. Thomas Mann, "More on What is Going on at the Library of Congress," (Jan. 1, 2007), http://www.guild2910.org/AFSCMEMoreOnWhatIsGoing.pdf.

16. Thomas Mann, "Will Google's Keyword Searching Eliminate the Need for LC Cataloging and Classification?" (Aug. 15, 2005), http://www.guild2910.org/searching.htm.

17. Jackie Byrd, et al., "A White Paper on the Future of Cataloging at Indiana University," (2006) http://www.iub.edu/~libtserv/pub/Future_of_Cataloging_White_Paper.pdf.

18. University of California Libraries, Bibliographic Services Task Force, "Rethinking How We Provide Bibliographic Services for the University of California." (2005), http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf.

19. Karen G. Schneider, "How OPACs Suck, Part 1: Relevance Rank (or the Lack of It)." *ALA TechSource* (Mar. 13, 2006), http://www.techsource.ala.org/blog/2006/03/how-opacs-suck-part-1-relevance-rank-or-the-lack-of-it.html.

20. Karen G. Schneider, "How OPACs Suck, Part 2: The Checklist of Shame." *ALA TechSource* (Apr. 3, 2006), http://www.techsource.ala.org/blog/2006/04/how-opacs-suck-part-2-the-checklist-of-shame.html.

21. Karen G. Schneider, "How OPACs Suck, Part 3: The Big Picture." *ALA TechSource* (May 20, 2006), http://www.techsource.ala.org/blog/2006/05/how-opacs-suck-part-3-the-big-picture.html.

22. Marcum, "The Future of Cataloging: Address to the EBSCO Leadership Seminar, Boston Massachusetts."

23. Calhoun, "The Changing Nature of the Catalog and Its Integration with Other Discovery Tools."

24. University of California Libraries, Bibliographic Services Task Force, "Rethinking How We Provide Bibliographic Services for the University of California."

25. Schneider, "How OPACs Suck, Part 1."

26. Schneider, "How OPACs Suck, Part 2."

27. Schneider, "How OPACs Suck, Part 3."

28. Karen Markey, "The Online Library Catalog: Paradise Lost and Paradise Regained?" *D-Lib Magazine* 13, no. 1/2 (2007), http://www.dlib.org/dlib/january07/markey/01markey.html, doi: http://dx.doi.org/10.1045/january2007-markey.

29. Mann, "What Is Going on at the Library of Congress?"

30. Mann, "More on What Is Going on at the Library of Congress."

31. Mann, "Will Google's Keyword Searching Eliminate the Need for LC Cataloging and Classification?"

32. Byrd et al., "A White Paper on the Future of Cataloging at Indiana University."

33. Mann, "More on What Is Going on at the Library of Congress."

34. Schneider, "How OPACs Suck, Part 1."

35. Schneider, "How OPACs Suck, Part 2."

36. Schneider, "How OPACs Suck, Part 3."

37. Marshall Breeding, "Trends in Library Automation: Meeting the Challenges of a New Generation of Library Users." *OCLC Office of Research, Distinguished Seminar Series* (Nov. 29, 2006), http://www.oclc.org/research/dss/ppt/breeding.ppt.

38. Endeca, "Endeca Information Access Platform," (2007) http://endeca.com/.

39. Kristin Antelman, Emily Lynema and Andrew K. Pace, "Toward a Twenty-First Century Library Catalog." *Information Technology & Libraries* 25, no. 3 (2006): 128-139, http://eprints.rclis.org/archive/00007332/01/antelman_lynema_pace.pdf.

40. Breeding, "Trends in Library Automation: Meeting the Challenges of a New Generation of Library Users."

41. Schneider, "How OPACs Suck, Part 3."

42. University of Rochester, River Campus Libraries, "eXtensible Catalog (XC): About XC," (2006) http://www.extensiblecatalog.info/?page_id=2.

43. Schneider, "How OPACs Suck, Part 3."

44. Lorcan Dempsey, "Lifting Out the Catalog Discovery Experience," *Lorcan Dempsey's Weblog on Libraries, Services and Networks* (May 14, 2006), http://orweblog.oclc.org/archives/001021.html.

45. Jeffrey Toobin, "Google's Moon Shot: The Quest for the Universal Library." *The New Yorker* 82, no. 48 (2007): 30-35, http://www.newyorker.com/reporting/2007/02/05/070205fa_fact_toobin.

46. Eric Morgan, "A 'Next Generation' Library Catalog," (July 7, 2006), http://dewey.library.nd.edu/morgan/ngc/.

47. Nancy J. Fallgren, "Users and Uses of Bibliographic Data: Background Paper for the Working Group on the Future of Bibliographic Control," (Feb. 25, 2007), http://www.loc.gov/bibliographic-future/meetings/docs/UsersandUsesBackgroundPaper.pdf.

48. Markey, "The Online Library Catalog: Paradise Lost and Paradise Regained?"

49. Dempsey, "Lifting Out the Catalog Discovery Experience."

50. Marcum, "The Future of Cataloging: Address to the EBSCO Leadership Seminar, Boston Massachusetts."

51. Markey, "The Online Library Catalog: Paradise Lost and Paradise Regained?"

52. Tina Gross and Arlene G. Taylor, "What Have We Got to Lose? The Effect of Controlled Vocabulary on Keyword Searching Results." *College & Research Libraries* 66, no. 3 (2005): 212-230.

53. Henk J. Voorbij, "Title Keywords and Subject Descriptors: A Comparison of Subject Search Entries of Books in the Humanities and Social Sciences," *Journal of Documentation* 54, no. 4 (1998): 466-476.

54. Clay Shirky, "Ontology is Overrated: Categories, Links and Tags" *Clay Shirky's Writings About the Internet* (2005), http://www.shirky.com/writings/ontology_overrated.html.

55. Abby Blachly, "Tagging Meets Subject Headings," *Thingology Blog* (May 14, 2006), http://www.librarything.com/thingology/2006/05/tagging-meets-subject-headings.php.

56. Zhichen Xu et al., "Towards the Semantic Web: Collaborative Tag Suggestions." In: *Collaborative Web Tagging Workshop* (2006). http://www.semanticmetadata.net/hosted/taggingws-www2006-files/13.pdf.

57. Rahmatollah Fattahi, "AACR2 and Catalogue Production Technology: The Relevance of Cataloguing Principles to the Online Environment." In: *The International Conference on the Principles and Future Development of AACR* (Toronto, Canada, Oct. 23-25, 1997), http://eprints.rclis.org/archive/00007715/01/Fattahi-AACR2.pdf.

58. R. Conrad Winke, "Discarding the Main Entry in an Online Cataloging Environment." *Cataloging & Classification Quarterly* 16, no. 1 (1993): 53-70.

59. International Federation of Library Associations and Institutions, Section on Cataloging, "Functional Requirements for Bibliographic Records: Final Report." *UBCIM Publications*, new ser. vol. 19 (1998), http://www.ifla.org/VII/s13/frbr/frbr.pdf.

60. Caroline R. Arms and William Y. Arms, "Mixed Content and Mixed Metadata: Information Discovery in a Messy World: [draft]," (2003), http://www.cs.cornell.edu/wya/papers/ALA-2003.php.

61. Avi Rappoport, "Background Topics: Faceted Metadata Search and Browse." *SearchTools.com* (Feb. 8, 2006), http://www.searchtools.com/info/faceted-metadata.html.

62. Schneider, "How OPACs Suck, Part 2."

63. Rebecca J. Dean, "FAST: Development of Simplified Headings for Metadata" (paper presented at the conference, Authority Control: Definition and International Experiences, Feb. 10-12, 2003: Florence, Italy), http://www.sba.unifi.it/ac/relazioni/dean_eng.pdf.

64. Ione T. Damasco, "FAST: Faceted Application of Subject Terminology–A New Spin on an Old Standard," *TechKnow* 12, issue 2 (2006): 1-8, http://www.library.kent.edu/files/TechKNOW_July_2006.pdf.

65. James D. Anderson and Melissa A. Hofman, "A Fully Faceted Syntax for Library of Congress Subject Headings," *Cataloging & Classification Quarterly* 43, no. 1 (2006): 7-38, http://dx.doi.org/10.1300/J104v43n01_03.

66. Mann, "More on What is Going on at the Library of Congress."

67. Fallgren, "Users and Uses of Bibliographic Data: Background Paper for the Working Group on the Future of Bibliographic Control."

68. Eric Childress, "Non-MARC Cataloging Standards Overview: TEI & EAD, MODS, METS, XML-Based MARC," (presented to OCLC Members Council, Feb. 10, 2003), http://www.oclc.org/research/memberscouncil/2003-02/childress.ppt.

69. Rebecca S. Guenther, "MODS: The Metadata Object Description Schema," *Libraries and the Academy* 3, no. 2 (2003): 137-150.

70. Marty Kurth and David Ruddy, "A Collection-Based Approach to Metadata" (presented at the *SENYLRC Workshop*, June 17, 2003), http://www.library.cornell.edu/alcts/SENYLRC-June03_v2.ppt.

71. Martin Kurth, David Ruddy and Nathan Rupp, "Repurposing MARC Metadata: Using Digital Project Experience to Develop a Metadata Management Design." *Library Hi Tech* 22, no. 2 (2004): 153-165, http://lts.library.cornell.edu/lts/who/pre/upload/p153.pdf.

72. Robin Cover (ed.), "MARC (Machine Readable Cataloging) and SGML/XML," *Cover Pages* (July 23, 2002), http://xml.coverpages.org/marc.html.

73. Edward Gaynor, "From MARC to Markup: SGML and Online Library Systems," *Cover Pages* (May 8, 1996), http://xml.coverpages.org/gaynorMARC96.html.

74. Morgan, "A 'Next Generation' Library Catalog."

75. Karen Coyle, Diane Hillman, Jonathan Rochkind and Paul Weiss, "Framework for a Bibliographic Future," 2007, http://futurelib.pbwiki.com/Framework.