Challenges of Electronic Resources: State of the Art and Unresolved Issues

by

John D. Byrum, Jr.
Library of Congress
Washington, D. C. USA

As the first year of the new millennium comes to an end, we are well into an era that has revolutionized the theory and practice of library and information science. Anyone who entered the profession before the early 1990s could not have been prepared for what has followed the birth and proliferation of what we currently call "electronic resources". The overwhelming and every expanding mass of material technologically available on the Internet has been especially daunting to the profession. According to studies conducted by staff at OCLC, by June 30, 2001, the public Web[i] includes more than 3.1 million sites, a six percent increase over last year's total. The Web as a whole grew by 18 percent, reaching a total of nearly 9 million sites. Even though the OCLC study reveals that the rate of expansion has begun to slow down, the sheer magnitude of material already on the Web and being posted there each and every day presents an unprecedented challenge to the profession in terms of traditional responsibility to organize, provide access to, and preserve information. As another measure, the Internet Archive now has more than 10 billion entries, dating from 1996 and growing at a rate of 10 terabytes per month, thereby eclipsing the amount of data contained in every library in the world, including the Library of Congress.[ii]

Clearly, the challenges of remote access electronic resources cannot be denied nor avoided. We are not in a position to adopt the attitude of the Vatican censors who in the mid-16th century faced a similar problem, a surge in printed publications. As they struggled to broaden the Index of Prohibited Books, they found an easy solution: "What we need," they said, "is a halt to printing so that the Church can catch up with this deluge of publications."[iii] But, just as the censors proved unable to stem the tide, we too must face the reality that electronic resources have established themselves as an important and valuable type of library material to users throughout the world.

Beyond their sheer quantity, what other problems and issues do electronic resources pose to the profession? Further consider these questions from Nancy Cline, Librarian of Harvard College at Harvard University:

> ...[A]mid the proliferation of information, are we creating sustainable *systems* of access? Are we building reliable databases and durable objects? In our enthusiasm for access, are we overlooking important issues of reliability, redundancy, the ability to replicate results – important elements for continuity for scholars? While we work to incorporate vast amounts of digital information into our libraries, schools, universities, and colleges, how much should we concern ourselves with 'virtual community'?[iv]

---

[i] A public web site is defined as a distinct location on the Internet offering unrestricted public access to content via web protocols. From 1997 to 2000, the public web increased by about 700,000 sites each year, but the rate of growth is slowing; it increased by only 200,000 sites between 2000 and 2001. For further information regarding the OCLC Web Characterization Project and the results of its investigations, visit: <http://wcp.oclc.org/ > [Nov. 2001]

[ii] The Internet Archive 24 Oct. 2001; <http://www.archive.org/wayback/press_kit/press_release.html> [Nov. 2001]

[iii] Knowles, Jeremy R. "Facing the challenges of tomorrow". Letter dated Jan. 24, 2001 to the Faculty of Arts and Sciences, Harvard University, p. 5; reprinted in *Harvard College Gazette*, March 2001.

[iv] Cline, Nancy M. "Virtual Continuity: The Challenge for Research Libraries Today." *Educause Review*. May/June 2000, p. 22

Last year the Library of Congress celebrated its 200th anniversary with a series of bicentennial symposia including a Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. The ultimate goal of this event was to enable creation of an overall strategy to address the challenges Internet resources entail, particularly those related to discovery and description. Many specific objectives of the Library of Congress Conference are directly related to topics to be addressed here one year later. For example:

- Establishing characteristics of those Internet resources of sufficient value to justify the costs of cataloguing and preserving;

- Investigating changes to cataloguing practices and policies to make them more adaptable to accessing the proliferation and diversity of Web resources;

- Fostering standards to enable metadata communities to meet the discovery and retrieval needs of Internet users

- Promoting wider use of established subject heading thesauri and classification systems for more effective resource discovery and organization for research and reference.

- Exploring the portal interfaces between the online public catalogue and electronic resources related to publications the OPAC lists.

In the presentation of unresolved problems and some possible responses to them which follows, my theme has been clearly influenced by the content of an Action Plan to guide the Library of Congress in dealing with the bibliographic issues and recommendations that emerged from that Conference, and I would commend to you the published proceedings which are available both in print and on the Web.[v] My remarks are intended to be broad and general and to set the stage for many of the presentations that will follow.

The first and perhaps greatest need is for increased availability of bibliographic records for Web resources. While each institution needs to establish its own priorities, including within its service plan a commitment to creating bibliographic records for electronic resources, individual efforts should coalesce into regional and national efforts to achieve databases where a large number of bibliographic records are supported by standardized access points contributed by a large number of libraries, museums, and archives. These databases ideally would contain high quality records which could be shared extensively with considerable savings to those who re-use them. Therefore, we need to be sure that our bibliographic utilities, already rich in cataloguing data for traditional library materials, find ways to encourage participating members to contribute standard records for the full gamut of electronic resources.

One of the most impressive responses to this challenge is OCLC's Cooperative Online Resource Catalog. This service is a Web-based metadata creation system for bibliographic records and pathfinders describing electronic resources. Since Stuart Hunt will inform us more fully later regarding CORC, I will only mention now that it has proved to be much used, not only because of its user-friendly interfaces in which persons creating catalogue records (in either MARC or Dublin Core format) are assisted by programs but also because of its "Pathfinders" feature which provides electronic subject bibliographies that point to the

[v] *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, D. C. November 15-17, 2000. Sponsored by the Library of Congress Cataloging Directorate; ed. By Ann M. Sandberg-Fox.* Washington, D. C.: Library of Congress Cataloging Distribution Service, 2001. 536 p. (Also available on the Web at <http://www.loc.gov/catdir/bibcontrol/> [Nov. 2001].)

electronic resources represented by the bibliographic records. By October 2001, OCLC's World Cat contained some 500,000 records for electronic resources.[vi] Expansion of data should increase, since in addition to libraries, OCLC plans to seek the contribution of metadata from museums, archives, professional societies, publishers and others, including authors.[vii] CORC expansion is a component of OCLC's plan to transform WorldCat from a bibliographic database and online union catalog to a globally networked information resource of text, graphics, sound and motion. With attention to recruiting members in a host of countries outside the United States, this initiative should prove important to European librarians.

The Research Libraries Group (RLG) hosts another large bibliographic database in which a large number of records for networked resources is available. At this time, there are about 250,000 records that include "electronic access" links to E-journals, finding aids, table of contents, electronic texts, and Web sites. In addition, an important RLG undertaking, with support from the Ford Foundation, the RLG Cultural Materials Initiative[viii] has focused on primary sources and cultural materials—the rare and often unique works held largely by institutions that are so important for education and research, including include published and unpublished texts, images, objects, and artifacts of many types. Thus, this initiative is targeted at enhancing access to materials not only of interest to libraries but to archives, museums, and other cultural repositories. Access is provided through Web browsers connecting to an information retrieval interface developed specifically for this new resource. Common access points will be provided through mapping, while at the same time preserving description particular to the originating discipline. Also noteworthy, RLG has taken a leadership role, at least within the U. S., to support increased production and sharing of archival collection guides online.[ix]

Collaboration among librarians has also enabled development of INFOMINE , a database of about 25,000 records for scholarly and educational Internet resources, that have been selected and described by librarians from various institutions. Intended for an academic audience, INFOMINE covers most major disciplines and includes both free and fee-based Internet resources such as databases, image-bases, bibliographies, software archives, e-journals, e-texts, digital collections, and other finding tools and search engines. One of the goals of INFOMINE is to create an Internet finding tool that is freely available and provides an alternative to more general search engines for locating scholarly or educational Web resources, by providing more focused and appropriate search results focused on significant core and/or reference resources of interest to serious academic researchers. In many ways, INFOMINE's goals sound are similar to those of CORC – a database of librarian-selected and described resources, using standard subject headings and Dublin Core, built cooperatively thus hopefully leveraging time and money as an alternative to each institution building multiple redundant tools. Unlike CORC, INFOMINE is created using open source software (GNU GPL), and can be searched by anyone with a Web browser.[x]

CORC, the RLG initiatives, and INFOMINE exemplify different responses to the challenge of providing increased availability of records for and access to digital resources. I am sure that there are many more, and they too form part of the profession's response to this need.

Essential to the success of our efforts to join users and patrons with the resources they need is our ability to create a strategy for identifying those worthy of the cost of

[vi] Source < http://www.oclc.org/news/>. See also: <http://www.oclc.org/corc/> [Nov. 2001].

[vii] As reported on its home page OCLC's Global Strategy at <http://www.oclc.org/strategy/> [Nov. 2001].

[viii] For further information, visit: <http://www.rlg.org/culturalres> [Nov. 2001]

[ix] For further information, see: <http://www.rlg.org/arr/index.html> [Nov. 2001]

[x] For further information, see: <http://infomine.ucr.edu/> [Nov. 2001].

cataloguing them. The OCLC Web Characterization Project provides data regarding the distribution of public Web site providers across types of economic activity, and their statistics reveal that Web resources are available for the full range of topics of interest to researchers worldwide. This means that selection criteria which libraries, archives, and museums will have established for printed publications and other non-book materials would be applicable to evaluation of Web resources as well. These include authorship, content, provenance, accuracy, relevance to institutional mission, and subject matter. As Timothy Jewell has concluded in a recent survey of selection policy documents: "Perhaps the most commom thread running through ... these documents is that although electronic resources raise new questions, the value system brought to bear on selecting more traditional resources is still valid."[xi] But, in addition to these traditional selection criteria, remote access resources entail other characteristics that need to be considered in determining which of them justify cataloguing. Such additional features include design of the resource, ease of use, timeliness of content, permanence, quality of links to other sites, value-added utility beyond print versions, scholarly reputation of the originating domain, and uniqueness. And, of course, persistence of the resource itself and the URL by which it is accessed are obviously very important considerations in selecting Web resources. In addition, for those resources available on the Internet which are commercially available, their cost must be factored into the selection formula. Michael Gorman and other speakers will further address this topic.

In most institutions, generally staff other than cataloguers are responsible for making selection decisions and deciding processing priorities whether these pertain to traditional materials or electronic resources. In her paper on "Redesign of Library Workflows: Experimental Model for Electronic Resource Description,"[xii] Karen Calhoun argues that the highly centralized model for cataloguing library materials so characteristic of most libraries needs to give way to an "iterative, collaborative, and broadly distributed model". This concept values a team-based work organization, bringing together selectors, public services librarians, and cataloguers into the record creation process. At the Library of Congress, we implemented in the mid-1990s such an approach,[xiii] whereby reference staff who were building pathfinders would create initial bibliographic records for cataloguing purposes at the same time using software that guided them through the process of identifying and recording many of the data elements needed by the cataloguers who then reviewed the records and revised them as needed in accordance with standards. Communications between reference and technical services staff are in place via a "traffic manager" so that each member of the virtual team is up to date on the status of the records being created. As a result of this workflow, the traditional division of labor whereby the selector chooses resources, the cataloguer describes them, and the reference staff service them to the public is discarded in favor of a workflow that embodies cross-functional collaboration among the selectors, catalogers, and reference staff. Web resources lend themselves to this distributed model much more effectively than do physical objects being acquired and processed which must be handed from person to person.

Beyond increasing the number of standard records for digital resources, the profession is challenged to enhance access to and display of these records and to do so across multiple systems. This brings us to some of the issues related to standards – standards for creating bibliographic records, standards for access points of all kinds, and standards for communicating data across systems.

[xi] Jewell, Timothy D. *Selection and Presentation of Commercially Available Electronic Resources: Issues and Practices*. Washington, D. C.: Digital Library Federation, July 2001. Available at: <http://www.wlir.org/pubs/reports/pub99/oyb99pdf> [Nov. 2001].

[xii] *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, D. C. November 15-17, 2000*. pp. 357-376.

[xiii] For a detailed description of the workflow described visit: <http://lcweb.loc.gov/rr/business/beonline/workflo2.html> [Nov. 2001]

Much work is ongoing to update the International Standard Bibliographic Descriptions as well as most national and multinational cataloguing rules to provide guidance in dealing with the new and perplexing problems that electronic resources present – especially those which form the category called "Integrating Resources." Although now underway as our speakers later will demonstrate, progress in updating our cataloguing tools has been slow. As a result, our most basic need in this area is that the ISBDs and cataloguing codes provide standards covering the full array of electronic resources on a more timely basis and in harmony with each other. IFLA's ISBD Review Group is ready to serve as a focal point to which the authors of cataloguing codes can turn for guidance in achieving compatibility in deriving solutions to problems of bibliographic description. This kind of cooperation and coordination is essential to the continuation of the long-standing effort to maximize shared cataloguing at the global level.

Since others will explore various specific cataloging issues raised by the need to establish bibliographic control of electronic resources, I will only refer to one topic of general and increasing interest.

Many electronic resources raise multiple-version issues, by which I mean that some are exact or related digital expressions of works in other formats. In addition, different digital manifestations of the same work are not uncommon. As a practical matter, how shall these electronic versions be represented in our catalogues? This is not a new question, of course, but it is pervasive in the world of electronic resources, affecting both hand-held or remote. Should each version be given its own separate bibliographic description – at the cost of convenience to the user, who normally would prefer to find all representations of the manifestations within a single display? Or, should the bibliographic data for all versions be combined in a single display – usually at the cost of diminished identification of bibliographic features of the separate versions which in turn decreases potential for re-use of cataloguing records? Since the ISBDs and cataloguing codes normally deal with manifestations, cataloguers would find value in guidelines issued by national bibliographic agencies to help them decide in a more consistent manner when to create separate bibliographic records and when to create a single record for resources available in two or more versions. In addition or alternatively, these agencies, or other appropriate organizations, should encourage research for vendors of integrated bibliographic systems to enable local systems, when desired, to consolidate for display the separate records for the various manifestations called for by the ISBDs and cataloguing codes.

Related to improved discovery of digital resources is the need for mechanisms to promote greater efficiency in sharing authority data for persons, corporate bodies, and geographic places used as access points. Developments at the international level to foster more effective use of authority records would, of course, benefit access to the full range of library materials, not just electronic resources. As Barbara Tillett has so clearly pointed out: "Authority control enables 'precision and recall', which are lacking from today's Web searches."[xiv] What makes an international approach to authority control so challenging are the linguistic characteristics of the entity names to be controlled, not to mention lack of agreement among national codes as to the treatment of forms of headings.

These considerations have been fully appreciated by the IFLA Working Group on Functional Requirements and Numbering of Authority Records (FRANAR), which is seeking to establish a model by which records within the already existing national authority files can be linked through their record control numbers to variant forms for the same entities. Marie-France Plassard will have more to say about this important undertaking. Because

[xiv] Tillett, Barbara B., "Authority Control on the Web" In *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, D. C. November 15-17, 2000*, pp. 207.

authority control is such an expensive part of the cataloguing operation, the FRANAR project, which provides a more realistic approach to the problem than was realized by past efforts, must move forward quickly in order to lay the foundation for all the work that must follow.

However, given the model under development by FRANAR and the current state of technology, it is in fact already possible for two or more bibliographic agencies to implement a Virtual International Authority File, starting with personal name authority records. The concept is that participating institutions would use Open Archives Initiative (OAI) protocols to establish one or more servers to harvest essential metadata from their linked retrospective personal name authority files. The OAI protocols would continue future harvesting of metadata to take account of changed status resulting from deletions and updating. Such a project will soon be proposed to involve the Library of Congress, Die Deutsche Bibliothek, and OCLC.

As a further challenge to improved access to networked resources, librarians and information specialists need to pursue efforts to achieve semantic interoperability of controlled subject terminology and classification data. This topic too will be discussed later in this Conference and in greater depth. As Lois Mai Chan has noted: "Experimentation conducted on subject access systems in surrogate-based WebPACs and metadata processed systems demonstrate the potential benefit of structured approaches to description and organization of Web resources." [xv] Suffice it to say here that an approach to the task entails encouraging use of established subject heading schemes and thesauri at a general level, recognizing that more local or specific schemes may also be necessary to detailed indexing. In addition, we need to encourage the linking of established schemes to the fullest extent possible. There already is considerable evidence, based on actual correlations of Library of Congress Subject Headings with terminology in French, Spanish, and other languages, that this is a doable undertaking. Similarly, we need to promote use of established classification schemes and to link them not only to each other but to equivalent concepts in the subject heading vocabularies in wide use. However, our success will depend on our ability to overcome the problems of applying subject heading and classification schemes when extending them to the proliferation of networked resources. Chan informs us that these problems include "the need of trained catalogers for their proper application according to current [and I should add often complex] policies and procedures, the cost of maintenance, and their incompatibility with most tools now used on the Web."

As we all can appreciate, a proliferation of structures for metadata has emerged in recent years, perhaps the best known being the Dublin Core which has finally emerged as a NISO standard.[xvi] What is characteristic to the various schemes is that they provide a structure for housing information about resources but, alas, they offer little direction for recording data. The challenges I've described earlier about improving our cataloguing codes, authority apparatus, and controlled vocabularies all apply to metadata element sets as well. However, the metadata providers clearly do not share this view, because most of them do not want to burden their element sets with the full and detailed rules librarians normally apply in generating standard catalogue records. The developers of metadata formats have focused on creators and distributors of electronic documents as their users, believing quite rightly that authors and publishers are not likely to be much interested in the complexities of bibliographic description, standardized access, and subject analysis.

[xv] Chan, Lois Mai, "Exploiting LCSH, LCC, and DDC to Retrieve Networked Resources: Issues and Challenges". In *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium: Confronting the Challenges of Networked Resources and the Web. Washington, D. C. November 15-17, 2000*, pp. 159.

[xvi] *The Dublin Core Metadata Element Set* . Document Number: ANSI/NISO Z39.85-2001. 01-Oct-2001 , ISBN: 188012453X ,16 p. Available at no cost for PDF downloading at: <http://www.techstreet.com/cgi-bin/detail?product_id=926135> [Nov. 2001]

Therefore, the challenge in relation to metadata schema is to communicate and promote the values of standards, especially in relation to discovery, retrieval and display of information. At the Library of Congress, we believe a first step in encouraging the metadata community to give greater attention to content standardization would be to develop and disseminate a wide a statement of basic principles to explain clearly and convincingly why there is cost-benefit from the work we do. To make this statement as universally acceptable among librarians as possible, national or international conferences or workshops should be convened to formulate the principles. The next step would be for our professional associations to publicize them, specifically bringing them to the attention of metadata creators. The purpose of this effort is not to "train" creators and producers of electronic resources to become cataloguers, but to encourage them to supply metadata in a more standardized fashion that would enable better interoperability among different schemes.

I will mention two other challenges related to the metadata community. The first is the need to establish national or international registries that not only identify existing and new metadata schemes but also provides details as to what fields each includes in its element set and how these fields are labeled. A registry with this information could facilitate work to map elements across schemes wherever there is correspondence. National libraries or an international body might appoint a maintenance agency to develop and update the registry and to undertake the needed mapping. These mappings, of course, would include mappings to fields in our established communications formats, such as UNIMARC and MARC21. The other challenge in this area is the need to pursue efforts to develop persistent identifiers or naming systems to help us keep track of "mobile" resources. The OCLC Web Characterization Project established that of the Internet Protocol addresses identifying a Web site in 1998, only 25% also identify a Web site in 2001. This means that 75% either moved or ceased, and according to the authors of the study, the majority of these changes were "moves". (OCLC staff guesstimate that about 20% of Web sites disappear within one year, a statistic which reinforces the importance of capturing and preserving those that are valuable while they are still available.) Obviously with this kind of instability, the long felt need for persistent identifiers is well-founded.

Enhancing access to and display of cataloguing and metadata for Web resources across multiple systems is yet another unresolved problem before us. Greatly needed is an effort to determine whether a common interface for searching, retrieving, and sorting across discovery tools can be established, beginning with a definition of the requirements for such an interface. The discovery tools to be the subject of this investigation would include at least abstracting and indexing services, other content databases, and library catalogs. The appropriate body to assume responsibility for this project might be a standards making body, such as NISO or ISO or an organization like the Coalition for Networked Information or a combination of such groups as a joint venture.

Thus far, we have considered challenges of increasing the availability of records for selected resources, as well as challenges related to enhancing their discovery, description, and display. Many of the unresolved problems already mentioned and those which other speakers will address highlight the need for more automated tools, the need for increased research and development, the need for improved training and education for catalogers, and the need to achieve more partnerships with other members of the information community. Let us briefly examine each of these topics in terms of intrinsic unresolved problems to be addressed.

Clearly, the creation of software to assist with the work of creating and maintaining bibliographic records for Web resources would be quite desirable, and the most likely source of this help will come from the vendor community. We already have an exemplary model for the record creation part of this job in CORC where software will extract data from Web resources and manipulate them into an AACR2 or Dublin-Core based display for

cataloger review. This kind of front-end to the Web should become available to cataloguers everywhere and could become a valuable feature of our integrated library systems, if provided.

However, because the Web is such a dynamic environment with so many resources constantly being updated, maintenance of existing bibliographic data is an even greater challenge. Not only are Web sites subject to constant change, but so also are many other types of electronic resources. Aggregator databases are especially useful to the researcher, as we know, but to the cataloguer they pose a difficult challenge due to their fluidity. Librarians need to encourage development of software available on our integrated library systems to detect changes in resource content and to notify staff to modify affected records. Already on the market are such services as Mind-it[xvii] which provide Web masters (for a fee) with a tool to alert subscribed users (at no cost to them) to changes in their Web sites using a program that offers "customizable detection ...either broad (any relevant change is detected), or specific (e.g., minding for specific keywords, phrases, thresholds, etc.)". If built into our cataloguing workflows, this kind of functionality would enable us to improve dramatically the quality of our records for remote electronic resources.

Other programming needed for improving bibliographic control of Internet resources include development of metadata authoring tools to encourage creators of Web material to incorporate usable metadata in their products. Here the target industries are those who produce word processors, HTML editing tools, image creation and manipulation tools, and multimedia production tools. Software that would facilitate resource selection and evaluation is also desirable, resulting in tools that would examine the characteristics of the resources, the extent and nature of their linkages, and use patterns, and report results according to specified criteria.

The wish-list could be extended, but these examples of programming needs suffice to demonstrate how automation could help us better meet the challenges of the Web. The particular role for the librarian is to encourage vendors and other providers to create and market these tools by communicating priority needs for such enhancements and helping to establish appropriate specifications for the software sought.

Beyond software, there are other areas for research and development which entail more direct involvement of cataloguers and other library staff. Let me offer a couple of examples of such R&D challenges.

The first is related to the need already mentioned to achieve improved controlled vocabulary mediating tools and to simplify the policies and practices for subject analysis. Take, for example, the Library of Congress Subject Headings (LCSH), which within the United States is by far the most commonly used and widely accepted subject vocabulary for general application. LCSH has been translated into a variety of languages and adopted or modified by several national bibliographic agencies. It is, therefore, the de facto universal controlled vocabulary. However, research at OCLC to support its efforts to develop an automated and user-friendly interface between the Web and those using its CORC system found that LCSH's complex syntax and rules for constructing headings restrict its application by requiring highly skilled personnel. The syntax and rules also limit LCSH's effectiveness in the area of automated authority control. Therefore, OCLC research staff invented an approach which seeks the goal of creating a variant subject system for metadata which retains the rich LCSH vocabulary while being easier to maintain and apply. Known as "FAST" for Faceted Application of Subject Terminology, this schema takes a post-coordinate approach that separates time, space, and form from data from the subject heading string. According to its creators, FAST maintains upward compatibility with LCSH,

[xvii] For further information, visit: <http://netmind.com/> [Nov. 2001]

and any valid set of LC subject headings can be converted to FAST headings. [xviii]

Another example of an area where research and development is needed and in fact is under way is focused on the changing nature of the catalogue to take into account the possibility of integration of other discovery tools. In today's environment, the catalogue is a "tool for access and controlled descriptions of physical and virtual resources" [emphasis added].[xix] To realize its potential as a gateway or portal, user research would help to establish what resources in particular to which the catalogue should enable access – resources, such as abstracting and indexing services and other content databases. In addition, also needed, as already mentioned, is the ability to launch a search from the catalogue that retrieve citations and other information across range of discovery tools.

In the United States, a major investigation to expand the parameters of the catalogue is now underway as the result of the work of the Association for Research Libraries's Scholars Portal Working Group. (Two of our speakers, Sarah Thomas and Olivia Madison, are contributing to this Group's work.) The principal focus of this effort is to encourage development of discovery tools which "operate across both licensed and opening available content in a broad range of fields and delivers high-quality resources."[xx] This portal concept not only envisions enabling "single searches" across a variety of catalogues and databases, but also links the user to other library services including those which focus on document delivery and virtual reference services. The first phase of the Scholars Portal project, now in start-up, is cross-platform search functionality, targeting five categories of information: Library OPACs (those which conform to Z39.50 and are MARC 21 based); subject Web sites (encoded in HTML, SQL, and SGML); public domain A&I services; finding aids for special collections and archives; and resources digitized locally. The Scholars Portal is but one manifestation of a host of similar efforts to make electronic resources more efficiently available to catalogue users. Sarah Thomas will describe more in her presentation tomorrow. Suffice it to say now that we as bibliographic control specialists should seek opportunities to further these initiatives and participate in experimentation as opportunities arise.

Before leaving this topic, however, I should note that no discussion of the future of the catalogue would be complete without acknowledging the vision set out by Carl Lagoze in which he puts forth an "event-aware" model of cataloguing; arguing that a library is not merely a "book museum", he proposes non-traditional roles for the catalogue based on a "new reality" that recognizes digital resources reflecting the inherently dynamic nature of information, how it is delivered, and who takes responsibility for organizing and describing it.[xxi] While his vision is complex, the conceptual model he is developing certainly deserves attention and consideration.

Let me add to this rather long list of challenges by reiterating the importance of improving education and training as a means of addressing the unresolved problems of electronic resources. We must turn to our schools for library and information sciences, to our professional associations, and to our employers for courses, workshops, and in-house training focused on electronic resources. New competencies are needed to deal with the problems already outlined, including knowledge organization structures and systems – topics such as information retrieval, database design, indexing principles. At the Library of Congress Bicentennial conference last year, this topic was discussed at length, leading to

---

[xviii] 18. O'Neill, Edward T. and others. "FAST: Faceted Application of Subject Terminology". Presented at IFLA Pre-conference, Aug.2001. Paper available at: <http://wcp.oclc.org/fast/> [Nov. 2001]

[xix] *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium*; p. 483.

[xx] *ARL Scholars Portal Working Group Report*, May 2001, p. 2.

[xxi] Lagoze, Carl, "Business Unusual: How 'Event-Awareness' May Breathe Life Into the Catalog" in *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium*, pp. 269-286.

recommendations set out in five categories of educational needs[xxii]:

- Re-examining values: for example, refocusing on access instead of "agonizing over description"; demonstrating interest in accepting data from non-traditional sources;

- Determining core competencies: for example, including more flexibility, increased interest in cooperation; ability to participate in virtual teams (such as those responsible for collection development; greater partnering skills; especially knowledge in the use of standards for accessing and describing the full range of resources;

- Developing toolkits; for example to help with finding and harvesting metadata and performing maintenance tasks;

- Managing operations: for example, learning marketing skills and skills to enable practitioners to balance "old" (traditional) and "new" (cutting edge) and to create and adjust workflows;

- Identifying mechanisms to meet educational needs; such as clearing-houses for workshops and Web-based training opportunities.

These needs should inform those who are responsible for school curricula and continuing education programs and as a result better prepare both new professionals and current staff to meet the challenges of networked resources as well as improve effectiveness overall, including our ability to service traditional resources.

Finally, a reminder that to succeed in meeting the challenges of electronic resources, we must reaffirm the importance of collaboration and synergistic approaches to our efforts. We need to foster partnerships between libraries and a broad spectrum of other groups: metadata producers; standards developers; systems and software vendors; computing and technology suppliers; scholarly and academic enclaves; publishers; dot.com creators; bibliographic utilities; registration agencies; other information providers; government agencies; other libraries, including national libraries; and museums and archives. Interdisciplinary collaborations are completely necessary to our success in meeting the unresolved problems outlined in this presentation.

In particular, as librarians and information scientists, we should reach out to those who are creating and marketing products which offer us the opportunity to re-purpose the information which they create. Regina Reynolds in her thought-provoking paper "Partnerships to Mine Unexploited Sources of Metadata" pointed out that in this area there is no single blue-print by which we can establish these partnerships. But there basic principles to guide us in identifying and pursuing these opportunities:

- A hierarchy of catalog record levels, with the lower levels based on publisher-supplied metadata, can help bring some Web-based resources under bibliographic control.

- Metadata created for other purposes can be re-proposed for library use.

[xxii] *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium*; p. 465-466. For the latest version of the ABC model, see "The ABC Ontology and Model" (Oct./Nov. 2001), available in PDF format at: <http://www.cs.cornell.edu/lagoze/papers/JODI_Final.pdf> [Nov. 2001]

ï Resources creators and producers can create usable metadata, especially with librarian help and input.

ï While libraries cannot take sole responsibility for description of networked resources, they can help, lead, guide and share expertise.[xxiii]

As Reynolds observes: “Librarians need to collaborate, not replicate. Librarians need to be partners, not competitors. There are more than enough resources to go around!”

Thus we come to the end of our survey of the unresolved problems of electronic resources. Not that we have exhausted them all! To summarize, the most significant challenges before us are:

_ To provide coverage of valuable electronic resources by means of cooperative database building, re-purposing bibliographic information, and adapting our selection criteria to the special features of the digital world.

_ To enhance the discovery, description, and display of the bibliographic and metadata records and to do so across multiple systems.

_ To encourage the development of more automated tools for creating and maintaining bibliographic information and metadata.

_ To foster increased research and development to improve cataloguing tools and to enable the catalogue to search and retrieve information from a variety of sources.

_ To expand educational and training opportunities to prepare catalogers and other library staff to better understand and service electronic resources..

_ And, to seek a wide array of collaborative ventures with partners from throughout the information industry to gain needed resources to make it possible to meet the challenges of electronic resources.

In concluding, I would like simply to echo the words of the Harvard College Librarian Nancy Cline who wrote:

[xxiii] Reynolds, Regina Romano, “Partnerships to Mine Unexploited Sources of Metadata,” *Proceedings of the Bicentennial Conference on Bibliographic Control for the New Millennium*; p. 461-462.

> We cannot ignore the rapid acceleration of digital dependence in all aspects of education and research, nor can we overlook the researcher's need for permanence, reliability, and continuity in this digital age. ...[A]s we look to the new century, we must shape an information environment that has sustainable systems of access to enduring information resources so that users, now and in the future, can rely on them with confidence. Defining this future calls for new combinations of talent and expertise, for short- and long-term collaborations, and for experimentation and risk taking in order to develop the best strategies for managing the rapidly expanding amounts of digital information. [xxiv]

This then is our challenge and our opportunity. As cataloguers, reference librarians, information professionals, I submit that we can and will meet the unresolved problems of electronic resources and thereby ensure effective access to these resources for the patrons to be served in 2025 and beyond.

****

Endnotes

[xxiv] Cline, "Virtual Continuity", p. 28.