

To: Library of Congress Working Group on the Future of Bibliographic Control
From: Cataloging and Classification Standing Committee, American Association of Law Libraries
August 7, 2007

The following testimony was written on behalf of the Cataloging and Classification Standing Committee of the American Association of Law Libraries (AALL). It does not represent any official position of AALL itself.

As the de facto national library for the United States, the Library of Congress has had tremendous influence over the development and use of bibliographic standards and structures used throughout the library community.

Furthermore, standards and procedures adopted by the Library of Congress, as the provider of cataloging to the Law Library of Congress, have significant impact upon the law library community. As we saw with last year's change in policy regarding the creation of series authority records and series headings, there was a cascade of decisions made by libraries everywhere, with some conforming to the new Library of Congress practice and others continuing to apply previous policy. For libraries continuing to control series entries (including the libraries participating in the Project for Cooperative Cataloging) there are significant challenges. A significant stream of records coming from the Library of Congress are now not standards-compliant and need intervention at the end of the bibliographic control chain, significantly raising expenses and lowering quality for everyone, and often resulting in duplicative work at many libraries.

1) Users and Uses of Bibliographic Data

Bibliographic metadata is inherently complex because of the myriad variations present in the materials being described. The current MARC standards have evolved over the years to accommodate new variations and formats. However, current ILS systems do a poor job of searching the data included in the bibliographic record and of making use of the inherent relationships between the various data elements. In other words, we need much better data-mining capabilities. It is not enough to rely simply on "word search" capabilities. We should not discard current practices of content designation just because our current systems don't do enough with the data. Systems come and go, but data is forever.

Users may only be interested in the Google search box to begin with, but probably want much better ways to filter the results to get at what they need. North Carolina State University's Endeca catalog suggests what could be accomplished by using facets to limit search results, but facets depend on the presence of accurate metadata in catalog records.

The "one size fits all" model of information retrieval does not work. There needs to be a variety of ways to search the data, and at different levels. Keyword searching might be adequate if you have very specific words to search. Even then, as has been pointed out by Thomas Mann, in a multilingual catalog, word searching can be problematic and

limiting. If we are waiting for the semantic web, we are going to be waiting for a long time.

Classification systems are the most underutilized mechanism we have for finding like materials in an ILS. The only thing most systems do is to provide a call number index that just puts the records in a shelf order. Users of catalogs have no idea what the numbers mean because there is no apparent underlying data structure, i.e. the MARC classification record, to support it. This could be extremely useful in the field of law. Wouldn't it be nice to search by terms in the classification hierarchy and retrieve results that way? Dr. Jolande Goldberg's conception for the KIA-KIX classification for the law of indigenous peoples is a step in this direction, i.e. the integration of web resources into the class schedule. The next step would be to integrate it into the catalog.

Better integration of catalog-based bibliographic data with external resources is needed, e.g. searching for catalogued and uncatalogued resources at the same time.

2) *Structures and Standards for Bibliographic Data*

MARC may be old, but it still works and can be translated to a variety of different machine formats. It isn't perfect, but it's the content designation that is important. Machines don't know a statute from a loose-leaf service.

We need better standards of interoperability between MARC records and other formats. We also need to separate the data itself from the presentation aspects of the data. Machines can do a lot if we tell them what we want them to do and exactly how to do it. We need to leverage the huge investment we have in ILS systems and bibliographic records to get these systems to do more.

3) *Economics and Organization of Bibliographic Data*

The greater library community faces ongoing economic challenges including:

Recruitment, training, and retention of staff and catalogers

The skills that are desirable in catalogers and cataloging staff are in demand not just in libraries but throughout the job market.

Shared cataloging and maintenance

Being at the end of the chain of bibliographic record creation and use, all actions that we take locally become that much more expensive when not shared in meaningful ways.

Dissimilar information systems

The information systems that we use to do our work are often difficult to use, proprietary, or non-standards-compliant. The tools that we use daily are often in different systems, implemented in different ways, meaning metadata and catalog records are created at differing levels and standards. This poses a significant ongoing expense throughout the entire cycle of record creation and maintenance.

Enrichment of cataloging data

Enrichment of cataloging data has many barriers. Many of these are economic, including staffing shortages and disparate systems which make integration expensive. Another barrier is the fact that enriched bibliographic records purchased from vendors are often, by contract, not allowed to be shared on a bibliographic utility. Other barriers, such as OCLC's enhance process, appear to be policy-driven.

Controlled access points

The creation and maintenance of controlled access points and their associated authority files are part of the essential core of what is considered cataloging. These are expensive operations because they require human intervention at many points, but significant value is added. Just because authority control is expensive does not mean it is not a worthwhile investment of time and effort. The collocation of materials under controlled access points enables users to more readily find related works. If this concept is extended to a "work," as it is in the FRBR model, then many variations and access points can be applied to a single work and controlled through a single record. It would also provide for more consistent application of access points.

Cost of maintenance

The cost of bibliographic maintenance is one inordinately placed upon local institutions. There is a lack of structure and standards for the shared updating and enhancement of records and their subsequent re-consumption and use by all cooperating institutions.

If the Library of Congress is to abandon or radically change its historic role in maintaining MARC21, the Library of Congress Classification System, Library of Congress Subject Headings, and other standards and practices, the impacts upon the greater library community have the potential to be monumental. These standards and practices, developed over the last century, would shift the greater library community from a unified environment to a system of scattered standards and practices, much like what is currently plaguing the Internet. We view the abandonment of established bibliographic structures and practices as a step backward, not forward.

Much attention is paid to the idea of libraries as **cost** centers – certain processes and policies are deemed too expensive and therefore "unsustainable." Perhaps this focus on cost has come about because it is much easier to measure cost than to measure **value**. We cannot know precisely what good has come about in the world because of libraries; we must satisfy ourselves with subjective quality assessment surveys as gauges of value. Like Lord Darlington's cynic, we risk becoming people who know the price of everything and the value of nothing.

We do not resent or fear change. If anything, libraries have been on the forefront of change, especially when it comes to technology. Catalogers have been implementing the "wisdom of crowds" since the advent of OCLC and shared cataloging. What concerns us is change justified solely by short-term gains in productivity or other limited goals.

Respectfully submitted by:

Jean M. Pajerek

Cornell Law Library

Ithaca, NY 14853

jmp8@cornell.edu

Chair, AALL Cataloging and Classification Standing Committee