Published in Technicalities, Nov./Dec. 2006, v. 26, no. 6

Back to the Future

By Diane I. Hillmann
September 2006

The recent passing of Henriette Avram and Fred Kilgour reminds us yet again that as librarians we stand on the shoulders of giants (the irony, of course, is that Henriette was a tiny woman in the physical sense). To a great extent, Avram and Kilgour were the two most critical pioneers in creating the ability for libraries to share metadata easily and efficiently: Avram through the development of MARC, and Kilgour as the founding director of OCLC. Between them, they set the stage for an incredible revolution in libraries, as the MARC-based network became the vital underpinnings of a nation-wide collaboration that informed the last forty years of library development. I feel my age as I say this, remembering quite clearly the telephone couplers that were a "feature" of the first OCLC "terminals" available to me as toiled away creating MARC records in the early seventies. And who working then can forget the boxes of printed cards that arrived inexorably as a result of the hours on the terminal (we were so excited that they were interfiled!). Those were the days when catalogers typed their worksheets on manual typewriters, and they were then "input" by specialized staff—online time was precious in those days.

But enough of Memory Lane. Some of the most important changes that were happening as a result of this initial "wiring together" of libraries were far more social and cultural than technical. And like most social change there were resisters and complainers and those who never did figure out why "local needs" had to give way to standardization. None of the revolution in library cooperation we witnessed then happened overnight, and to a great extent, the "culture of sharing" that resulted is still evolving, even as it is clearly beginning to devolve in some significant ways.

There was a great deal written about the early days of OCLC during the late seventies and eighties, when the technological whiz-bang of what suddenly became possible was still new enough to be deemed interesting and worth documenting. Although a bit was written about the social changes beginning to happen as a result of the introduction of technology, there were only a few who understood how far libraries would move away from the notion of each library as an independent entity to one where each was a node in a much larger picture of information access. It was not so long ago that considerable time and energy could be spent discussing whether libraries should include in their catalogs information on materials that they did not "own"—needless to say, this topic of discussion evaporated on its own not too long after the advent of the Internet, bringing with it ubiquitous electronic accessibility of information. I was one of those early adopters of the idea of "cataloging the Internet" and still have some great slides on how to do it.

If we believe the stereotypes (and of course, they're often true), most catalogers have a much better grasp of the trees than they do of the forest. Sometimes this means that even when concepts are well accepted in one realm, the expansion of those same ideas into other realms meets great resistance. I was reminded of much of this a few years ago as I worked with serials colleagues on an extension of the sharing of bibliographic information to serial holdings. In that case, of course, it wasn't just resistance by catalogers. OCLC, the avatar of the sharing culture, persisted for many years in asserting that holdings was a local concern, not one that extended across libraries (except in the case

of union lists of periodical holdings that they maintained for some consortia). Thus, it took OCLC twenty years to implement the MARC Holdings format, and to accept the notion that the detail of what was published in a serial run was a much a part of the bibliographic description as the title changes. The fact that it was carried in coded form in a MARC Holdings record did not necessarily mean that the benefits of cooperative creation of holdings patterns for check-in systems and the sharing of cooperatively maintained detail on published holdings were not relevant. It took years of effort by a handful of pushy serialists (and the dawning interest of the CONSER Project) to make those points, even after thirty years of acceptance of shared bibliographic records.

But like shared bibliographic records, shared holdings information is based on the fact that most libraries either hold physically or maintain access to some subset of the information published for scholarly consumption, and this made shared cataloging ultimately a no-brainer. Of course it made little sense for dozens of catalogers to create descriptions of the exact same book or serial, using the same rules, conventions and syntax. But does this rationale extend into the metadata world, where most of the resources are not physical but digital, metadata descriptions seldom overlap (much less use the same schema), and nobody's quite sure where the boundaries of what a library should be cataloging actually extend?

There's an increasing sense that we're either on the verge or in the middle of the next big revolution in libraries. I was just a lowly catalog assistant for much of the last one, and though I remember things like the telephone couplers I wasn't really aware of what was going on at higher levels, so I decided to do some research. Probably the most accessible historical account of those times is Kathleen Maciuszko's book on OCLC's first decade, written from the perspective of only a few years past that decade's end. [1] I hadn't the time or inclination to read the whole thing, but what I did read certainly made me think about a number of interesting differences between those times and ours.

I had started out thinking that in a sense, the metadata world is reinventing shared cataloging, but it's clearly not that simple. The original 'revolution' was far more focused on re-articulating what was already being done with LC's catalog card distribution services and automating that process. It could be said that the consequences that followed were largely unintended, or at least, unanticipated, by anyone other than a few visionaries. Library administrators knew they needed to collaborate to survive, but they took many leaps of faith as they began to realize that goal. There's surely a parallel with our current upheaval—how many people that we knew 10 years ago anticipated how integral the Internet would become to our common enterprise? And how many of us believe that the current crop of self-styled visionaries has it right about what's coming?

Maciuszko includes an interesting quote from Henriette Avram about the thinking of the era:

> "The interest of libraries in the computer for library operations was increasing in the 1960's. The availability of cataloging data in machine-readable form supplied by LC, the need to input cataloging data locally (data not within the scope of MARC or titles not cataloged by LC), the possibility of sharing these locally generated records, the potential for using computer programs across organizations to reduce the high cost of designing and writing software, and the need for hardware capable of handling large character sets

> were all factors that put increased emphasis on the establishment and conformity of standards." [2]

The question becomes—are we still in an age of standardization? We've all heard variations on the theme of "The great thing about standards is there are so many to choose from." It's certainly the case that this is truer today than it was yesterday, and the trend shows no sign of abating. I recall at the OCLC FRBR workshop a few years ago a wonderful slide by Godfrey Rush of <indecs> where the little red boxes of standards proliferated on the screen like happy viruses. Clearly this viral proliferation is not the 'top-down' sort of standardization that drove the library revolution of the sixties and seventies, it's far more 'bottom-up'—driven by communities that have been empowered by the notion that they don't have to compromise the needs of their special users to play in the new metadata world. It's more about agreements than standards; a handshake rather than a written contract designed by lawyers with lots of fine print.

So, is conformity to standards still the goal in this new world? If we all do the same thing, in the same way, exchanging data is relatively easy—libraries have been doing that for about 40 years now, and it's been successful beyond the wildest imaginings of the Ohio librarians who got together so many years ago and thought up OCLC. But now Pandora's box is wide open, and anyone who thinks that there's a way to stuff all this activity back into those boundaries lives in a dream world. What we need to dream about instead is something different than conformity: interoperability.

Interoperability is a lot harder than standards conformance. There's no road map, no easy path, no complete understanding yet about what works and what doesn't. It's about communicating with one another about what we are actually doing, and may well include specifics about what standards we're following where. In this world of choices, rather than an on/off switch, communicating with others about the choices made while creating data is a distinct challenge. The rest of the challenge is how to deal with such diversity on the receiving end.

This diverse world affects the development of standards as well. In the AACR2 world, there were indeed options for application of some rules, but the basic rules were developed based on compromise. Choices about main entry for musical performances couldn't vary depending on whether the performance was rock music or classical, even though aficionados for the former tended to think more about the performing group, and for the latter the composer was generally more important. In the compromises made to ensure conformity in rule application, there were winners and losers, and the users generally had to work around these anomalies, because the catalog systems that displayed the data lacked flexibility as well. A catalog record was a catalog record, right? So where's the problem?

In our new world of role uncertainty and concerns about competition from Google for user eyeballs, we can't afford to create losers, much less user-losers. There is not one path for users to take as they look for materials, there is no 'ideal' set of information elements, nor an ideal display, based on what used to fit on a 3 x 5 card with a cute little hole in the bottom. The OCLC databases depended on the notion of standards conformity: there was a "master record" for each book or CD or website, and woe be to those who created a "duplicate" record, no matter how inadvertently. Efficiency, not diversity, was the watchword.

Maciuszko, in her detailed account of the birth and early development of OCLC and the related LC-based shared network for distribution of bibliographic data, records carefully the high-level organizational effort behind the giant steps taken on behalf of libraries in those days. Then the important tasks consisted of enabling a tight, predictable, top-down structure for the development of an efficient, nationwide system of data sharing. They weren't sure it would work, but it did, for a long time.

It still works, sort of. Lots of new records are created every day, they get distributed quickly and efficiently, using this self-same nationwide network. But it's not enough anymore. It's too bounded, too inflexible, and too few people are using its primary vehicle, the Online Public Access Catalog, a creature that acknowledges only a small portion of the resources a library makes available. We have our work cut out for us, as we move from conformity to interoperability, but move we must.

[1] Maciuszko, Kathleen L. OCLC, a decade of development, 1967-1977. Littleton, Colo. : Libraries Unlimited, 1984.

[2] Avram, Henriette D. MARC: its history and implications. Washington, D.C. : Library of Congress, 1975, p. 20.