Library of Congress

Study of the North American MARC Records Marketplace

October 2009

R2 Consulting LLC

Contents

I. Introduction

In January 2009, the Library of Congress (LC) contracted with R2 Consulting LLC (R2) to investigate and describe current approaches to the creation and distribution of MARC records in US and Canadian libraries. The primary focus is on the economics of existing practice, in effect mapping the "marketplace" for cataloging records, including incentives for and barriers to production. The underlying question is whether sufficient cataloging capacity exists in North America, and how that capacity is distributed. This project was designed to be descriptive rather than prescriptive, seeking to understand in detail the ways in which cataloging records are produced and distributed, as well as who bears the costs and who realizes the value. We are not attempting to offer solutions or suggest changes, though some have become obvious as we've looked at the data. One especially critical aspect of the project has been to assess the degree to which sources other than LC create records in significant quantities, and to determine the extent to which "all roads lead to DLC/DLC."

The goal is to achieve the best possible understanding of current circumstances and practices:

- What is the overall cataloging capacity in North America?
- Where does it reside?
- What are the primary distribution pathways and channels for sharing records?
- How much redundancy is there?
- What can we predict about cataloging capacity over the next 5-10 years?
- What is the estimated need/demand? How does this compare with capacity?
- What is the relative importance of authority control to libraries?
- What is the current reliance by North American Libraries on LC cataloging?

Over the course of six months, R2 employed a number of information-gathering techniques. First, we developed a social network called Bibliographic Record Production: www.bibrecordproduction.ning.com which ultimately attracted more than 800 members. This forum was used to develop and refine surveys, to assure that we were asking the right questions, and to enlist proportionate representation from all market segments. We performed a literature search as highlighted in the bibliography. We developed two extensive surveys, one for libraries and one for vendors, and worked diligently to assure the participation of school, public, academic and specialized libraries, and of Canadian as well as US libraries. We took special care with the school and small public library markets, as they are often under-represented in such studies, and rely almost exclusively on records produced by LC, even if those records reach them through other channels. We also interviewed key people by phone, and made a site visit to the Library of Congress.

The surveys were released in April and completed in May 2009. There are a handful of areas where gaps exist, but the response was proportionate to the size of the respective markets, a factor that gives us confidence in the results.   Overall, survey responses were strong, with 972 libraries and 70 vendors participating. Results are summarized in sections II and III of the report; Library and Distributor responses respectively.   Note that the survey questions themselves can be found online at:

- www.r2test.net/pdfs/Survey Questions - Libraries.pdf

- www.r2test.net/pdfs/Survey Questions - MARC Systems, Distributors, and Service Providers.pdf

Despite many revisions and our best efforts to achieve clarity in the survey questions, it is apparent that a common understanding does not apply across all market segments. There is, in fact, not really a shared understanding of what constitutes a MARC record, since it can serve purposes other than cataloging. In addition, the distinction between creating a record (which ideally occurs once for each title) and distributing a record (where the same record may be provided to multiple customers) proved confusing to some respondents. This has made quantitative comparisons unreliable, and we have introduced them only in cases where the data are relatively unambiguous.

Our primary observations and conclusions are described in the two subsequent sections of the report:

      III. The Conflicted Market
      IV. Economics of Cataloging

Conclusions are based on careful consideration of survey results; interviews and conversations with practicing librarians and vendors; discussion among members of the Bibliographic Record Production social network; extensive reading; participation in the OCLC/NISO Metadata Symposium (April 2009); and our own direct experience with cataloging production and distribution. Primary conclusions include:

1. **Library of Congress cataloging continues to be widely valued**: Libraries, vendors, and cooperatives speak with their actions. There is heavy reliance on LC's output throughout all segments of the profession and industry. This is demonstrated by 500,000 searches per day against LC's Z39.50 servers and WebOPAC; by extensive re-sale and re-use of records distributed by the MARC Distribution Service (MDS); and by the variety and scale of use across all library sizes and types, and all vendor sizes and types. LC records are the cornerstone of the entire market. School and public libraries are especially reliant on them, but all market segments have built services on the foundation of inexpensive and easily obtainable LC records.

2. **The Library of Congress subsidizes portions of the market:** LC catalogs many titles that ultimately are not retained in its collections. As a result, LC bears significant costs from which it receives no direct benefit, for activity that is not explicitly in support of its core users. The 1902 law that governs distribution of its records deliberately excludes the cost of production from the pricing for those records. There is no revenue to offset those costs, other than the value of the free copies of the CIP books provided by publishers. The market relies to a surprising degree on LC's willingness to bear these costs and forgo this revenue. If LC were to redirect its catalogers' efforts solely to materials deemed necessary by its users, CIP production would diminish significantly. Other organizations would need to assume those costs. At present, libraries and vendors enjoy the largely unrecognized benefits of an LC subsidy.

3. **LC records are significantly underpriced:** Not only does LC bear a disproportionate share of the costs associated with producing records for titles it may not retain, the law governing its sale of those records allows only the cost of distribution (plus 10%) to be recouped. The cost of production is assumed to be part of LC's ongoing operations. Such low prices contribute to the impression that cataloging should cost less than it actually does.

4. **Cataloging backlogs continue to grow in many areas and market segments:** As outlined in the library survey responses, non-Roman languages, maps, and DVDs pose particular problems. But

to our surprise, many libraries are also losing ground on mainstream materials such as English-language monographs.

5. **There is adequate cataloging capacity in North America to meet the collective need:** This finding surprised us, especially given the aging of the profession and imminent retirements. However, a conservative interpretation of survey data shown on pages 9-10 strongly suggests that there are more than enough catalogers to handle everything. In the academic market alone, for instance, the survey indicates that more than 8,000 original catalogers are employed. If each original cataloger produced on average one record per work day (or 200 per year), that would indicate capacity for 1.6 million original records annually. Unfortunately, that capacity is not well distributed, disciplined, or coordinated, despite decades of experience with cooperative cataloging.

6. **Cooperative cataloging has not realized its full potential:** Shared catalogs, bibliographic utilities and other tools make cooperative efforts more convenient and effective. The Program for Cooperative Cataloging (PCC)'s BIBCO initiative contributed more than 76,000 records in 2008. But BIBCO and CONSER each have fewer than 50 members, and ten BIBCO libraries contributed nearly 2/3 of all records produced in 2008.

7. **The market for cataloging records is conflicted:** The library market must accommodate both community values and commercial values. In most elements of the market, this works well, but cataloging brings the two into conflict. Libraries operate within a "community" value system that prizes openness, accessibility, and free access to information. Vendors operate within a "commercial" system that creates sustainability, growth and profit.

8. **The market provides insufficient incentives to stimulate additional original cataloging:** Since backlogs continue to grow in many areas, it is curious that so many libraries choose not to participate in cooperative cataloging programs. It is equally curious that vendors do not see opportunities to serve as cataloging agencies. For some reason, neither commercial nor community incentives have much effect. Is it possible this is because the cost (and therefore the price) of cataloging has been understated? Or does the profession as a whole no longer believe original cataloging is worth what it actually costs?

9. **80% of libraries edit records for English-language monographs in their local catalog:** Most editing is performed "to meet local needs"; such as re-Cuttering, adding workmarks, or removing unwanted subject headings. There is still widespread resistance to the idea of simply accepting the work of another library.

10. **78% of libraries are unaware of any restrictions on MARC record use or redistribution:** In part this stems from the perception that cataloging records are free, or perhaps community-owned. While this is true for LC records, it becomes problematic for those records where the cost of production has not yet been recovered. This erodes the market.

These findings are fully developed in the pages that follow, along with many other more modest observations and analysis.

II. Library MARC Record Survey Response Summary

The total number of responses to the Library MARC Record Use Survey was 972. While the library respondents were self-selected (not random), we are happy to have achieved proportionate representation from all segments of the library community. The chart on the left indicates the approximate number of North American libraries by type (estimated in the original project plan), and the chart on the right indicates the actual breakdown of library responses to our survey.
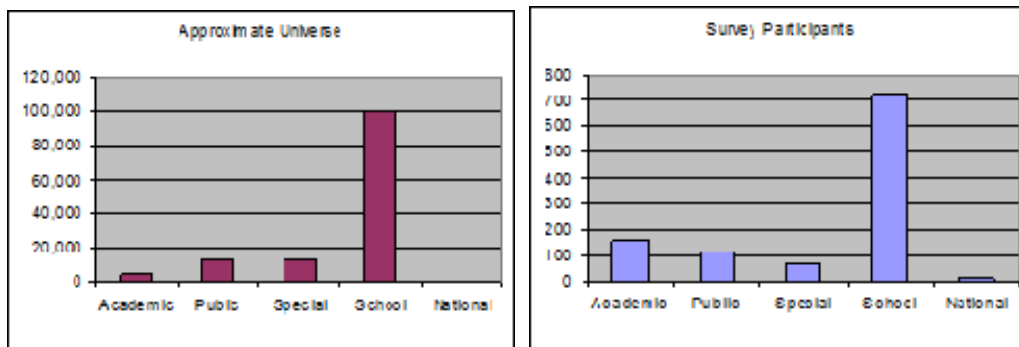
As we expected, academic libraries are somewhat overrepresented. In general, though, our analysis of library data is type specific so we don't believe this to be problematic.

The breakdown of academic library participants is as follows:
- 5 post- high school technical or trade school libraries
- 18 community college libraries
- 32 four-year college libraries
- 56 mid-sized university libraries
- 47 large university libraries (ARL)

With regard to public libraries, small, mid-sized, and large libraries participated as follows:
- 19 small (population served less than 10,000)
- 73 mid-sized (population served between 10,000 and 250,000)
- 22 large (population served more than 250,000)

School libraries participated as follows:
- 233 elementary school libraries
- 193 middle school libraries
- 299 high school libraries

65 specialized libraries completed the survey, including law, music, art, museum, health sciences, business, and corporate, and departmental libraries, etc.

5 national libraries participated, including the National Agricultural Library, the Law Library of Congress, the National Library of Medicine, Library and Archives Canada, and the Library of Congress itself.

One disappointment regarding our library data is that Canadian libraries are under-represented. Despite significant effort to invite their participation, only 33 Canadian libraries (3.5% of the sample) completed the survey. 11-12% of the whole would have been appropriate. It is not possible, therefore, to draw

conclusions differentiating their practices from those of U.S. libraries. We do, however, want to draw readers' attention to a 2002-2003 phased study of "The Use of NLC MARC Records in Canadian Libraries". [1 & 2] There are very many differences between that project and this, but there are certainly parallels. Beheshti et al gathered useful information regarding cataloging practice in Canadian libraries vis-à-vis Amicus (Library and Archives Canada) records which, in conjunction with this study, offers some insight into the similarities and differences between Canadian and US practices.

One particularly soothing similarity is that the Canadian researchers encountered some of the same trouble we did in relation to librarians and vendors that can sometimes be uninformed consumers. For example:

> Some respondents were not aware of the ultimate sources of the records they receive, they do not know about the twists and turns in the MARC record distribution and redistribution system. For example, a library that uses AG-Canada as a source claimed little use of NLC records, but AG-Canada loads the full NLC source file, or libraries claim they first check larger public libraries or university libraries, who themselves may well be deriving NLC records. The following quote is from the cover note of a library that responded that they would not complete the questionnaire as it did not apply to them:
>
> "We use Amicus records to catalogue our items. I don't think that they are NLC sources."
>
> Some libraries heavily use vendor outsourcing or batch searching, but do not know from where the service ultimately obtains those records. For example, many Manitoba libraries that use MAPLIN think they do not use any NLC records, even though MAPLIN is a Z39.50 distributed search which includes Amicus in its list of targets, and thus is quite likely to include NLC records in those retrieved. Other libraries that use vendors (for example, the Library Services Centre which was frequently mentioned in Ontario) do not know where those records originate. However, many of these vendors use copy cataloguing as the basis for their product. Some vendors add a great deal to the records (such as SDM which makes sure the record is suitable for a French-language catalogue), but others mainly add only acquisitions and holdings information.[2]

Even though R2 encountered similar problems, we believe our library survey responses to be largely valid. For the most part, data collected confirmed previously held perceptions about cataloging practices, and quantified previously identified problems within the profession. Because the differences between academic, public, and school libraries are so marked, most of our analysis is type specific. However, the collective response, undifferentiated by library type, is sometimes worth noting because certain practices and problems are shared across all types. With regard to copy cataloging, we present the collective response because it speaks so clearly to the issue of redundancy and the failure to cooperate.

Copy Cataloging Redundancy
Survey results show that 80% of libraries edit records for English language monographs in their local catalog. Only 50% of those that do also upload local edits to their bibliographic utility or consortial catalog. Those that do not share their edits report the following reasons (they were invited to flag more than one):

- 50% say their edits are only appropriate to the local catalog
- 32% say they don't have enough time

- 24% say they are not authorized or trained
- 7% say they lack confidence
- 5% say that there are technical obstacles

These data suggest an enormously high degree of redundancy with regard to copy cataloging those resources for which copy is most widely available. R2 observations in academic libraries over the last ten years suggest that this level of scrutiny is often applied even when the source of the copy is LC.

This information becomes even more worrisome when we consider that in some libraries, copy cataloging is performed by MLS and other professional librarians. As stated by the Primary Research Group in 2008, 52% of their academic library survey respondents use professional librarians for copy cataloging. This percentage varies by library size, as follows:

> … 39% of libraries in colleges with more than 10,000 students use professional librarians routinely for copy cataloging, as did 53% of community colleges and … 11.1% of research level universities. [18]

As R2 has learned through careful observation in libraries, and as the PRG research confirms, there is a profound lack of consistency regarding the use of professional and paraprofessional staff for various cataloging-related tasks. While professional librarians often perform routine copy cataloging, paraprofessionals are just as likely to perform original cataloging. This too varies by size of library.

> About 27% of survey participants routinely use paraprofessional staff for original cataloging. Public colleges were more than three times more likely than private colleges to use paraprofessionals for original cataloging, and larger colleges were more than twice as likely as smaller ones to do so. More than two-thirds of research universities use paraprofessionals for original cataloging.[18]

We raise this issue not to judge the behavior, but to suggest that there is little uniformity on which to base staffing/cost estimates.

Staffing Levels
The data we gathered with regard to staffing levels must be parsed by library type to be meaningful. With regard to original cataloging talent, library respondents reported having no original catalogers at rates as follows:
- Academic = 5.3%
- Public = 14.4%
- Special = 26%
- School = 77%

With regard to copy cataloging talent, library respondents reported having no copy catalogers at the following rates:
- Academic = 6%
- Public = 6%
- Special = 24%
- School = 63%

An even more granular break-down is helpful as we seek to estimate existing capacity:

	special libraries	4 –year college libraries	mid-sized university libraries	large research libraries (ARL)	mid-sized public libraries	large public libraries	high-school libraries	elementary school libraries
sample size	65	32	56	47	73	22	199	233
average # of original catalogers	1	1	2.6	6.8	1.3	4.5	.2	.2
median # of original catalogers	1	1	2	6.75	1	2	0	0
lowest reported # of original catalogers	0	0	0	0	0	.6	0	0
highest reported # of original catalogers	8	4	11	32	4	15	5	3
average # of copy catalogers	1.5	1.8	4.4	10.7	1.9	4.3	.5	.3
median # of copy catalogers	1	1.75	3	10	2	2.5	0	0
lowest reported # of copy catalogers	0	0	0	0	0	1	0	0
highest reported # of copy catalogers	5	6	22	50	9	11	8	3

Please note that when less than full-time equivalents were reported, they were counted as 0. For example, a library reporting 1.5 original cataloging FTEs was counted as having 1; .5 FTE was counted as 0.

Cataloging Capacity
It's a bit of a leap, but we suggest that our median numbers, multiplied by the estimated universe of libraries (by type) could offer a very rough estimate as to the current cataloging capacity in North American libraries.   In an effort not to overestimate, we'll take the most conservative view possible. For purposes of this exercise let's assume that:

- every academic library employs 2 original catalogers and 2 copy catalogers
- every public library employs 1 original cataloger and 1 copy cataloger
- every specialized library employs 1 original cataloger and 1 copy cataloger
- there are no original catalogers and no copy catalogers in school libraries

The data indicates that there are, in fact, many more catalogers than this, but again, the point is to offer a very conservative estimate of overall capacity.

Here is the math, in its simplest form:
- 4182 academic libraries in NA x 2 original catalogers = 8364
- 12,442 public libraries in NA x 1 original cataloger = 12,442
- 13,349 specialized libraries in NA x 1 original cataloger = 13,349

    Note that the estimated numbers of North American libraries by type were drawn from OCLC's WorldMap Beta v2.0 on 2/10/09: http://www.oclc.org/research/researchworks/worldmap/prototype.htm

Total original catalogers, then, in North American libraries would be well over 34,155 and we'd estimate the same number of copy catalogers. If each original cataloger were to create just one new record each work day (or 200 per year), 6,831,000 original records could be created annually. Given the frequency with which copy catalogers create original records, that number could be even higher.

Obviously, these numbers reflect the immediate circumstance.   For many years, dire predictions about expected cataloger retirements have caused widespread concern about the future of the profession. While our study did not specifically address this issue, several others have:

- In 2002, Stanley Wilder predicted that one-third of the ARL cataloging population would retire by 2010.[23]

- In 2005, Leyson and Boydston found that one-third of professional ARL catalogers were projected to retire by 2015.[10]

- In 2008, the Primary Research Group reported that the mean number of academic librarians in mostly cataloging functions that are likely to retire over the next five years was a mean of 27%.[18]

It is nearly 2010, and while we have no doubt lost a great number of professional catalogers since 2002, the impact of those retirements is somewhat unclear. Perhaps cataloging backlogs are bigger now? Perhaps more paraprofessionals are picking up the slack? Perhaps the level of description is dropping off? R2 suggests that more research in this area may be appropriate - with an eye towards the impact of all these retirements on patron access; but we suggest that in the immediate, cataloging capacity is not the biggest concern.

Cataloging Backlogs
Apparently adequate capacity, however, does not necessarily translate into current or timely cataloging.
Survey results show that backlogs are continuing to grow in many libraries and in the most common
material types. Again, a granular view, by library type is appropriate. While cataloging backlogs for
mainstream materials such as English print monographs, videos, and DVDs, and sound recordings exist
across all library types, there are variations in percentage among types as well as variations in the kinds
of materials backlogged. Reportedly, cataloging backlogs are increasing in size in the following areas by
the highest percentages (>20% for at least one group):

	Academic	Public	School
Eng print monographs	**22.7%**	**51.6%**	**54.9%**
Non-Eng print monographs	**22.7%**	12.9%	3.9%
Non-Eng, Non-Roman	**29.3%**	16.1%	2.6%
eBooks	**21.3%**	11.3%	17.1%
Videos & DVDs	**26.7%**	**56.5%**	**67.1%**
Audio Books	2.7%	**33.9%**	**34.5%**
Maps	**22.7%**	11.3%	8.9%
Sound recordings	18.7%	**30.6%**	18.1%
Rare Books	**34.7%**	9.7%	2.3%

56% of all libraries report having backlogs that are increasing in size, with videos and DVDs being the
most problematic, followed closely by English language print monographs. Combined survey responses,
regardless of library type report growing backlogs in the following percentages:
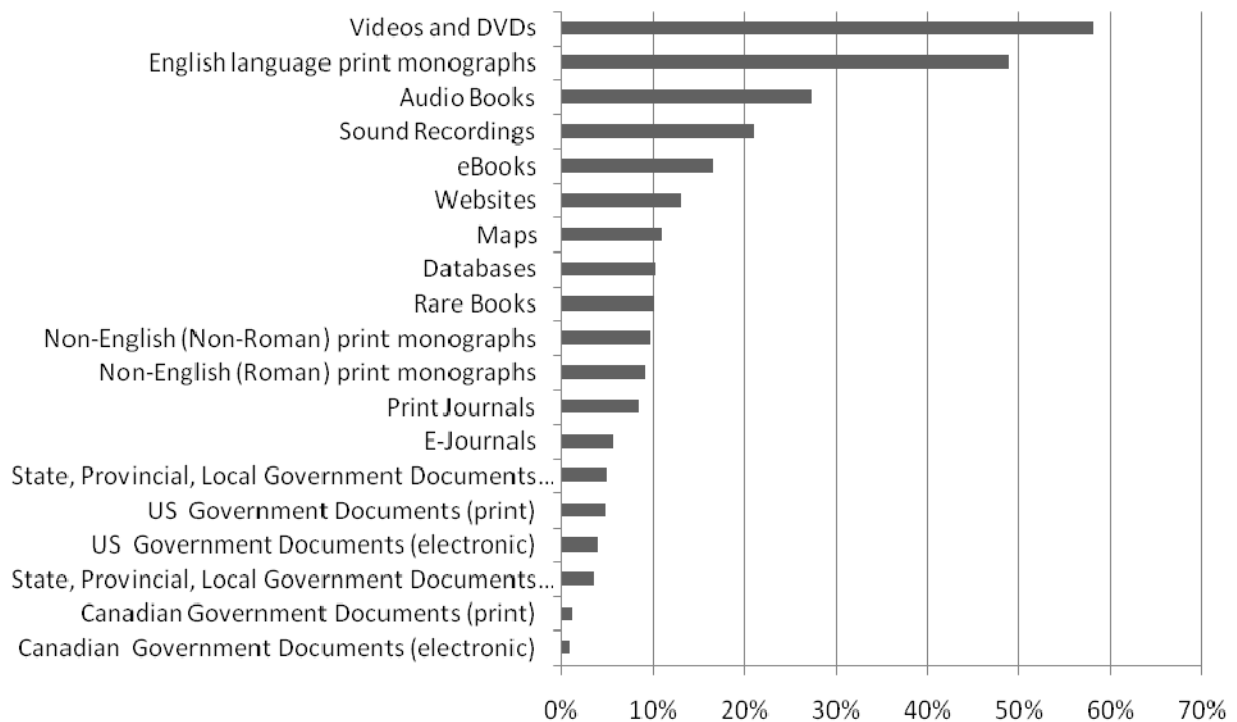
These data lead us to ask what catalogers are doing. Bob Wolven and others suggest that catalogers are
being called upon to apply their knowledge of cataloging principles to new initiatives; and specifically to

creating metadata for digital and archival collections.[24] As well, it is very common for high performing catalogers to be moved into supervisory, training, and management positions, effectively reducing or even eliminating their individual production. Survey responses generally support these theories.

The majority of libraries with original catalogers report their primary focus as follows:

- Academics = specific (non-book and non-English) formats followed by special collections
        and non-MARC resource description
- Publics = special collections followed by specific (non-book) formats
- Schools = subjects by a wide margin

The majority of libraries with copy catalogers report their primary focus as follows:

- Academics = specific (non-book and non-English) formats, followed closely
        by catalog maintenance
- Publics = catalog maintenance
- Schools = catalog maintenance

For the most part, we interpret catalog maintenance to mean holdings, URLs, headings, reclamation projects, batch record loads, and routine withdrawals and transfers, etc. in the LOCAL OPAC.

More about Copy Cataloging

A significant amount of routine copy cataloging that traditionally occurred within libraries has been outsourced to material vendors and commercial bibliographic utilities. The sheer number of WorldCat Cataloging Partners and Commercial Bibliographic Utilities (see the next section of the report), suggest a high degree of reliance on third parties for copy cataloging. Nevertheless, a great deal of entirely manual copy cataloging still occurs in all types of libraries. Survey respondents reported the following with regard to copy cataloging practices in their own institutions.

- The most reliable sources for copy cataloging are reportedly OCLC, LC, and Amicus. School Libraries rely primarily on Z39.50 searches (of LC) and MARC records supplied by vendors.

- Although most edits are described as addressing local needs, the following are the specific record edits most frequently cited:

  - Adding pagination
  - Changing headings or removing headings for Juvenile and non-English materials
  - Adding contents notes
  - Adding Dewey numbers and Sears headings
  - Correcting the date of publication
  - Adding or editing URLs

- In response to the question, "How long might you wait for copy to be available from elsewhere? -- answers varied from, "We don't wait" to "years and years." The time span cited most frequently was three to six months.

- The most important data elements sought in third party records are standard numbers, author, title, publication information, and extensive subject description – full records with as many access points as possible. LC records are the most highly sought, period.

Collections

Without question, different library types vary dramatically regarding the level of cataloging needed. These differences are driven in part by characteristics of their user communities, but even more so by the size and complexity of the local collection. Obviously, the bigger and more permanent the collection, the more granularity and depth is needed with regard to resource description. In seeking to understand these differences, we thought it might be helpful to compare collecting patterns by library type and resource format. In the end, these patterns don't apparently offer us any new insights, but we include the following summary with the hope that someone else will see something of significance:

- The majority of all types of libraries, *except* those in schools, collect print journals. The same is true for Non-English, Roman Alphabet materials.

- E-journals are collected by most academic and national libraries, by 40% of specialized and public libraries, but seldom by school libraries.

- Audio Books are collected by most public libraries, the majority of school libraries but by few libraries in other groups.

- Print Maps, Micro Formats, and Non-Roman Print Monographs are collected mainly by academic and national libraries as well as by 20-30% of public and specialized libraries

- Databases are collected by 44.1% of academic, 83.7 % of public, 27.7 % of schools and 44.4% of specialized libraries.

- Websites are cataloged by a slight majority of the academic group, 30% of specialized, and by few others.

- Rare books are collected by 17.8% of respondents, mostly academic libraries

- All library groups report collecting eBooks with the highest concentration in academic, followed by public.

- State, Provincial and/or local print documents are collected by most academic and public libraries. E-versions are collected to some degree by all library groups except schools. Both electronic and print U.S. Government Documents are collected by the majority of national and academic, but by few others. Collections of electronic and print Canadian documents are collected by few libraries other than National and Canadian libraries.

Authority Control

As we would expect, the data we gathered with regard to authority control practices suggest that those responsible for managing bigger more complex collections pay considerably more attention to authority control than those managing smaller more ephemeral collections. Here again, there are no real surprises. Libraries reported:

	Academic	Public	School
Access to a current file:	73.8%	55.3%	42.4%
Access to a non-current file:	30.2%	38.1%	23.4%
No Access:	.8%	9.5%	35%

Some libraries reported _no_ headings verification activity:

- Academics = 2.3%
- Publics = 7.7%
- Schools = 28%

But a higher number report that their authority control has been outsourced:

- Academics = 36.2%
- Publics = 17.3%
- Schools = 6.7%.

71.1% of all respondents report that authority practices are determined by the specific functionality of local systems.

Loose Ends

- 80.6% of library respondents indicated that end users are not able to add information to the records in their database/service primarily due to ILS limitations or to policy restrictions in school libraries. Adding tagging features is a topic of discussion for all non-school libraries planning ILS replacement or enhancement.

- In response to the query about what else we should know about libraries' MARC records environments, there was great appreciation for LC's work as well as pleas that it continue. There is general dismay about the quality of non-LC vendor provided records. There were several negative remarks about LC's cessation of Series Authority work. Many libraries stated that free records are essential to their operations. And many reported difficulty managing vendor supplied record sets for eBook collections and serials packages.

III. MARC Record Distributors and Service Providers

There are approximately 200 organizations that create, sell, and/or distribute MARC records and/or MARC services to North American libraries. Many are commercial and many others are public or non-profit cooperatives and service providers.   Additionally, there are more than a dozen library system vendors that compile MARC record databases that they embed in, or sell in conjunction with their hardware/software packages. Some of the largest publishers distribute MARC records together with their content, and various consortia and large libraries also make their records available, often free of charge. OCLC is the biggest player in this arena, and in most ways, is in a league of its own.

At the outset of the study, R2 contacted as many of these organizations as we could, to inform them of the project, and to invite participation in our online survey. A list of entities contacted can be found in Appendix A.   Ten or twelve offered useful suggestions for how to improve the survey and 70 completed it, including nearly all of the biggest most influential players. Aside from OCLC and LC itself, individual distributor responses are masked, but their high level of participation offers an extremely reliable representation of this segment of the market.

This is how they chose to describe themselves (they were invited to choose more than one descriptor if appropriate):

Descriptor	Response Percent	Response Count
Commercial bibliographic utility, cataloging service or other MARC records service provider	32.3%	21
Material vendor	27.7%	18
Aggregator	20.0%	13
Publisher	16.9%	11
Other (book binderies, national libraries, research organizations, individual contract catalogers, etc.)	16.9%	11
System Vendor	13.8%	9
Non-profit consortium/cooperative	12.3%	8
Open database provider (Z39.50 compliant)	6.2%	4
Public or school library hub (central Tech Services unit for affiliate libraries)	4.6%	3

One difficulty worth noting here is that non-profits, consortial entities, and centralized technical services operations have significant roles in the distribution of bibliographic and authorities records to their member libraries. Many are named as the primary record source by their affiliate libraries but they do not define themselves as distributors and were therefore reluctant to participate in this survey.   Neither do they define themselves as libraries. These kinds of entities, then, are underrepresented in both surveys. Nevertheless, we recognize them to be important in the overall scheme of things, and suggest that further research may be necessary to scope their full breadth and depth of MARC record activities.

For those vendors and distributors that did participate in our survey, the following chart offers a summary of the kinds of records and services provided. Each participant was invited to flag all that applied. We were initially surprised that there are more entities that distribute records for eBooks than for print monographs, but a closer look reveals that traditional book vendors now distribute records for both, but eBook sellers and other aggregators of digital content, provide records for electronic editions only.
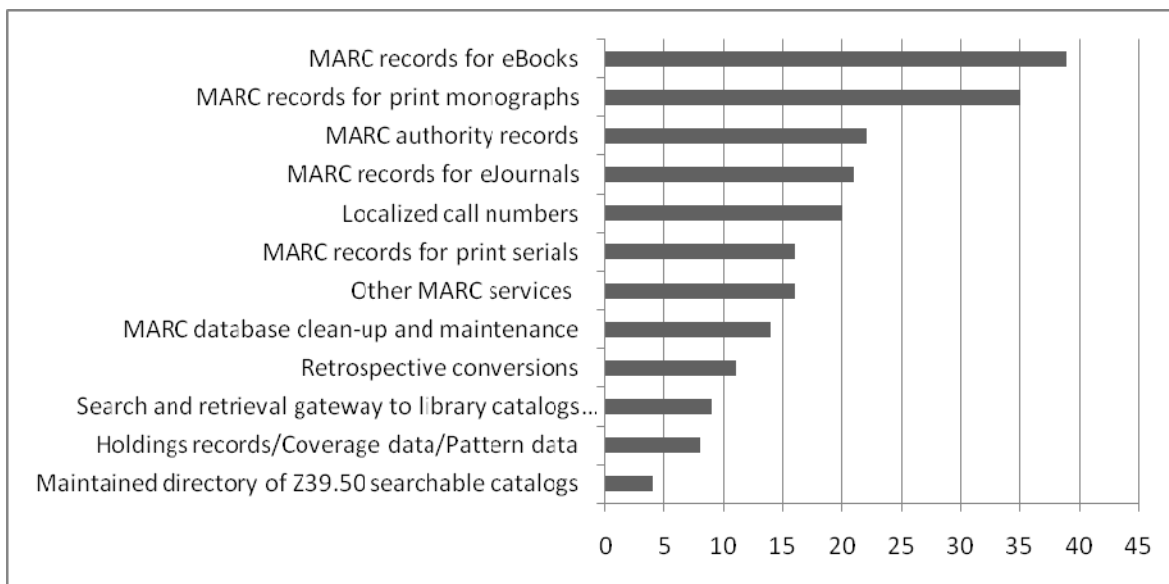
Note that "Other MARC services" (above) include the delivery of brief MARC records for efficient acquisitions process; MARC records for monographic standing orders (not serials); Open URL Links from online database to library OPACs via standard numbers (ISBN, ISSN, ISMN); MARC record enrichment including TOCs, Accelerated Reader, Lexile, and Reading Counts; Maintained directory of Z39.50 searchable catalogs; etc.

Most of our distributor respondents do not specialize in specific or subject areas:

Does your organization specialize in specific subject areas?	Number of Respondents
No	53
Yes, Reference Works	4
Yes, Music	3
Yes, Art	2
Yes, Business	2
Yes, Science and Technology	2
Yes, Law	2
Yes, Health Sciences	2

But a somewhat greater number specialize in materials published in specific geographic regions:

Does your organization specialize in materials published in a specific geographic region? Select all that apply.	Number of Respondents
No	43
Yes, Western Europe	13
Yes, Canada	12
Yes, US	10
Yes, Latin and/or South America	5
Yes, Eastern Europe	2
Yes, Australia/Pacifica	2
Yes, Asia	1
Yes, Africa	0
Yes, Middle East	0

Most distributor respondents produce and/or supply MARC records for multiple formats, as demonstrated in the following chart. The x axis represents number of respondents:
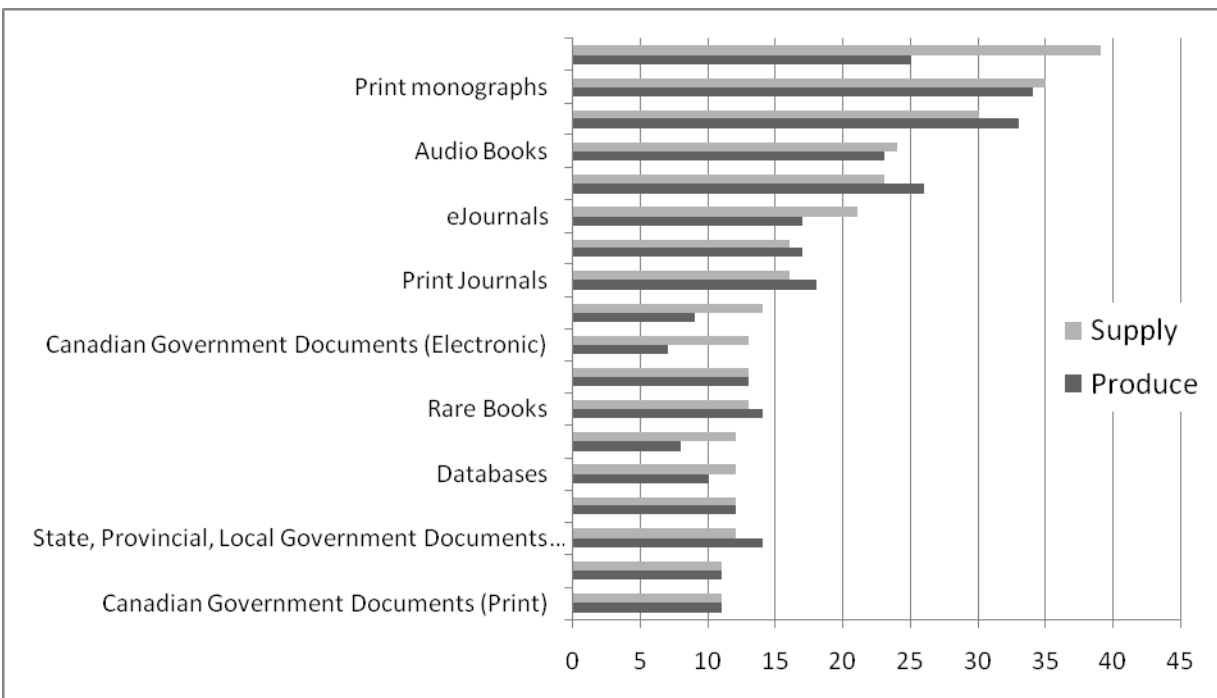

In many cases, the production lines for a particular format are longer than the supply lines, which must be understood in a couple of ways:

o   Some of our survey participants are contract catalogers that work with one library at a time, within that library's catalog. While these people produce records, they do not distribute them.

o   Some of our survey participants are database providers, meaning that their customers and/or member libraries "come and take" records. These entities also, do not distribute per se.

It is also important to remember that these indicators of production and supply activity reflect all levels of record, from machine generated provisional records, to Full MARC21.   Nonetheless, the relative levels of activity are reasonable; generally matching our expectations.

With regard to language, this same pattern exists, and for similar reasons.   Our participants reported the following:
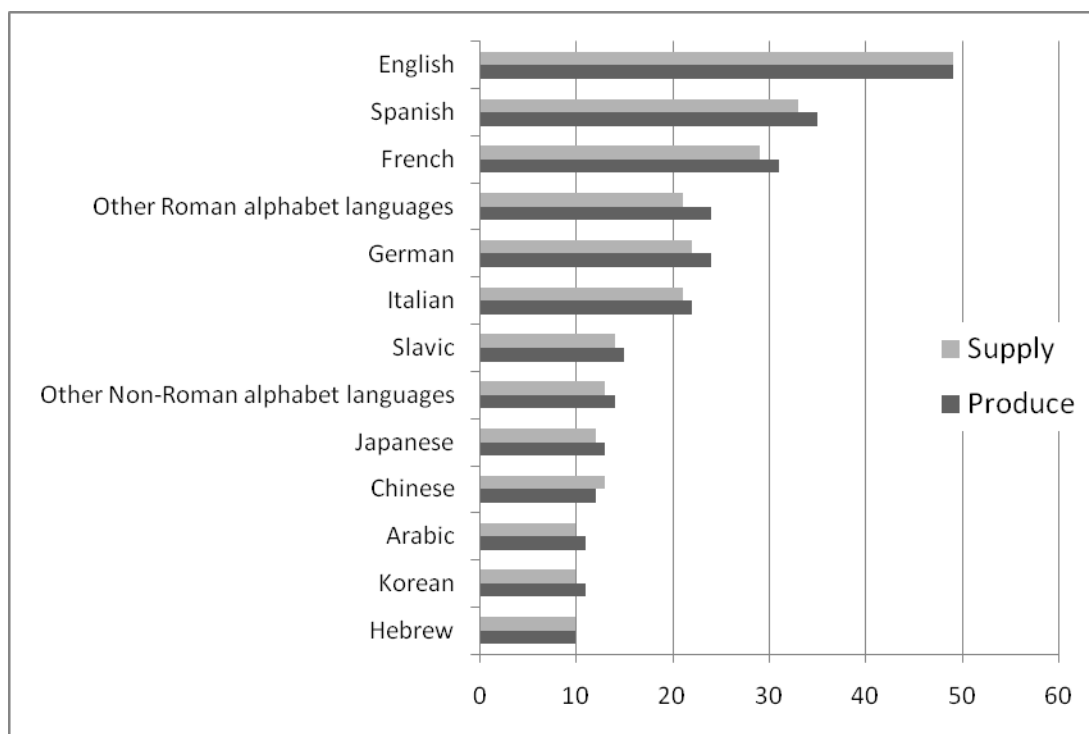
Additional data gathered from the vendor survey can be summarized as follows:

- The MARC databases maintained by our sample of distributors range in size from academic publishers that report hundreds of records, to OCLC, which reports the largest database by far (housing 195,000,000 bibliographic records).   The following is a rough break-down of distributor participants by database size:
    - o   14 report having a MARC database that houses fewer than 20,000 records
    - o   13 report having a MARC database that houses between 20,000 and 300,000 records
    - o   6 report having a MARC database that houses between 300,000 and 1,000,000 records
    - o   14 report having a MARC database that houses between 1 and 10 million records
    - o   12 report having a MARC database that houses between 10 and 41 million records

- LC maintains a MARC database of 22 million bibliographic records. As it turns out, several entities report larger databases. These respondents described themselves as commercial bibliographic utilities, open database providers, and one non-profit library consortium.

- 15 participating distributors purchase records directly from LC. None of them report databases smaller than 300,000 records but beyond that, there is no apparent pattern as to which entities do so.

- 7 participating distributors report that they do not acquire MARC records from external sources, but the rest do. Of those external sources, LC was predominant, followed by OCLC, LC record resellers, Library and Archives Canada, and the British National Library. Approximately 14% of respondents acquire a significant portion of their records via Z39.50 protocols and various web crawlers.

- While only 15 (21%) of distributor respondents purchase records directly from LC, 39 (56% of respondents) report that a significant number of the records they distribute are full LC records.

- 25 (37%) of participating distributors report that they perform CIP upgrades. In total, they reportedly perform 351,000 annually.

- A slightly different 25 (37%) of participating distributors are WorldCat Cataloging Partners.

- As described in an earlier section of the report, we do not believe this number of annualized CIP upgrades to be valid, and suggest that some respondents reported the number of CIP upgrades distributed (copy counts) rather than title counts. Nevertheless, we can state with confidence that there is considerable duplication of effort here, particularly within the sector of domestic book vendors. More on this below.

- When asked whether their organization creates original MARC records:
  - 20 (28%) of participating distributors said no

  - 25 (36%) of participating distributors said yes; with book in hand
  - 13 (19%) of participating distributors said yes; on the basis of metadata or surrogates
  - 12 (17%) of participating distributors said yes; and reported a combination of book in hand cataloging and cataloging based on metadata or surrogates

- Of the 50 participating distributors who claim to create original MARC records:
  - 12 – manually create Provisional Records (limited access points; no call number)
  - 12 – machine generate Provisional Records (limited access points; no call number)
  - 12 – create Brief Records (call number; and limited access points)
  - 31 – create Full or LC Core Records (i.e. AACR2, LCSH; and LC classification)

- The total number of Full MARC Records created last year (i.e. AACR2, LCSH; and LC classification), as reported by our 70 survey participants, was 1,376,688. Here again, some of this must represent redundant effort, but we have no confidence in this estimate as the number of catalogers reported (see below) could not have achieved this number of original records.

- 22 (31%) of participating distributors make their records available via Z39.50 protocols and the same number report that at least some of their records are free of charge. 18 of those entities report that all their records are openly accessible and available without charge.

- 18 (29%) of participating distributors chose not to comment on profitability, but of those that did:
  - 52% report that their MARC record products and services are profitable
  - 40% report that their MARC record products and services break-even
  - 8% report that their MARC record products and services are not profitable

- 55% of distributor participants report that they place no restrictions on the re-use of records they distribute. The other 45% report variable restrictions.

- 9% of participating distributors report that all their records are made available (harvested by) Google/Google Books/Google Scholar, etc.; 8% report that some of their records are made available; and 11% report that plans are being made to make their records available. Still 71% of survey respondents report that their records are not available for this kind of harvesting, and have no plans to make them so.

- Excluding LC, distributor participants report 694 copy and/or original catalogers in their employ; 332 (48%) of whom have a Master's of Library Science or an equivalent degree.

- LC reports 288 professional and 67 paraprofessional catalogers.

So how to interpret this untidy accumulation of data?

Even while protecting the privacy of survey participants, it is possible to categorize them by type, and represent several of their characteristics in graphic form. In general, visualization can be a helpful way to organize and communicate an otherwise overwhelming mass of details, like those collected via our survey.
In this case, the resulting graphic takes the form of a map (thumbnail at right), which can be found on the last two pages of this report; Appendix B.

On the next few pages, various elements, relationships, and activity patterns are described; along with observations made as a result of studying the map. We encourage the reader to have the full-sized map in hand as it is described. We are hopeful that additional observations will be made by others.
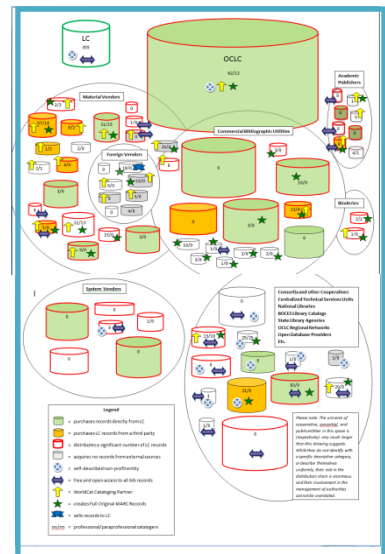
Detailed Map of MARC Record Distributors and Service Providers
Again, R2 believes that participation in the vendor/distributor survey was very strong, and except for cooperatives and consortia, the information gathered is fully representative of this segment of the market.   A handful of important players chose not to participate, but based on what we know about

their business models and market share, etc, we do not believe their presence on the map would alter its basic shape.

Seven black circles encompass entities categorically:

- Material Vendors and Aggregators
  - A subset of foreign material vendors
- Commercial bibliographic utilities (CBUs)
- System Vendors
- Academic Publishers
- Binderies
- Noncommercial Entities - including consortia, cooperatives, national libraries, open database providers, regional networks, etc. (this universe is vastly under represented)

There is overlap between various categories, insofar as some entities identify with more than one. The biggest area of overlap is between Material Vendors and Commercial Bibliographic Utilities. For some material vendors and aggregators, for example, cataloging services have become a critical component of their business. There is also overlap between CBUs and System Vendors, between System Vendors and Material Vendors, and beween System Vendors and Non-commercial entities. Still, the categories are distinct enough to be meaningful, and we have located each entity as aptly as possible.

Within the universe of CBUs, primary markets cited include an almost even mix of school, public, and academic libraries. Bibliographic and authority service providers are represented; and both US and Canadian libraries are served by various of these companies. It is interesting to note that several of these employ no catalogers.

Within the universe of Material Vendors and Aggregators, print and digital formats are well represented, as are monographs and serials. Primary library markets again include an even mix of school, public , and academic libraries. As would be expected, the foreign vendors are primarily oriented towards academic libraries in the US and Canada.

Academic Publishers in the upper right quadrant are relative newcomers to the arena of MARC distribution. Their drive to sell directly to libraries, has, over the last 5-10 years required that they support their material sales with associated MARC records. There is surprising variability regarding their inidividual practices, which may demonstrate a less developed sector, and/or a lack of especially successful solutions.

The Library of Congress and OCLC are the only two named organizations on the map and are located at the top. These two organizations are represented by a cylinder, as is every survey respondent. Each cylinder roughly represents the relative size of the database reported. As previously mentioned, a handful of organizations maintain databases of MARC records that are bigger than LC's. For the most part, these very large databases support the work of commercial and/or cooperative bibliographic utilities, and there are a few systems vendors that reported databases that are only slightly smaller. Obviously, none rival OCLC with regard to size. We are unable to evaluate the quality or level of records in each database, but suggest that shear size is a valid comparative factor with regard to prominence in the market. The only exception to this is probably LC itself. LC's relatively small database almost defies its singular importance.

Please note that the smallest cylinders on the map are not nearly small enough, relative to the others. In the case of individual contract catalogers in the universe of CBUs, for example, there may be no database at all, but are represented by cylinders nonetheless. It is important that these entities remain visible, as they create a considerable number of full original MARC records.

The green cylinders represent those entities that puchase records directly from CDS, LC's Catalog Distribution Service. Green cylinders appear almost randomly on the map. Most of the very largest organizations are direct customers of LC, but there is an equal number of medium and small entities across the market, including system vendors, material vendors, commercial bibliographic utilities, non-commercial entities, and open database providers, that also buy CDS products directly.

Gold cylinders represent organizations that puchase LC records from third-party suppliers. These too, appear randomly throughout the market. Grey cylinders represent those few entities that do not acquire records from any external sources. White cylinders represent organizations that acquire most of their records via Z39.50 search and retrieval protocols; have no permanent database; acquire records from non-US sources; or chose not to share information about where they get their records.

Cylinders outlined in red represent entities that report distribution of significant numbers of Library of Congress records. LC records are, of course, the "coin of the realm" here, and touted whenever possible. Organizations that do not purchase LC records but distribute them, have several potential sources from which to acquire them. One is LC itself, but there are also hundreds and hundreds of academic and research libraries (not represented on this map) that provide open access to their MARC databases as well.

WorldCat Cataloging Partners are marked with ⬆ and are heavily clustered in the universe of Material Vendors and Aggregators. Three of our eight Publishers are WC Partners, and a few more WC Partners are CBUs or Non-Commercial entities. WCP (formerly PromptCat) is primarily a workflow solution for libraries who prefer this batch approach to copy cataloging for mainstream monographs. It is understood to be cost effective for OCLC member libraries, holdings can be set in batch, and the libraries' catalogers can focus on non-mainstream resources. There is a general assumption of a better hit rate (member copy; UKMARC; etc.) than any individual vendor could provide so libraries have less need for exception processing. The uptake of this workflow solution, at least among US Academic Libraries, cannot be overstated.

Most WC Partners have some cataloging capacity within their organizations, and most report that they routinely perform a CIP upgrades. The combined number of CIP upgrades reported is astronomically high, and again, we do not trust that this survey question was interpreted consistently. Nonetheless, the degree of redundancy with regard to CIP upgrades and provisional records created within the universe of Material Vendors must be very high as a consequence of their strictly competitive relationships.

A ⭐ denotes an organization that creates Full Original MARC Records. It is fascinating to see that except for the universe of systems vendors, professional level cataloging takes place in every segment of the market, and with significant volume. The numbers within each cylinder are those reported for professional/and paraprofessional catalogers. 4/6, for example, indicates that the organization employs 4 professional and 6 paraprofessional catalogers. While we do not have confidence in the production numbers reported, we do believe these staffing numbers to be valid and take us a long way towards estimating cataloging capacity in these sectors.

Several survey respondents describe their organizations as non-profits, which are marked with . Not surprisingly, these appear predominantly in the lower right quadrant of Non-Commercial entities.

A indicates that access to the database is entirely open (mostly via Z39.50) and records are free of charge. These organizations are primarily non-commercial entities or content providers seeking to support the sale of their primary products, quite often books or eBooks. It should not go unnoticed that LC itself provides open access to its MARC records via multiple channels. The prevalence of open databases is a key factor in the economic confusion that plagues the MARC Record Market, and is discussed at more length in the next section of the report.

IV. A Conflicted Market

Librarians have always been somewhat uneasy with the idea that they constitute and participate in a "market" at all. With few exceptions, libraries exist to serve the communities that fund them. They do not receive direct payment for such services. Their mission includes making information as broadly, freely, and conveniently accessible to their users as is possible. The motivations and values of libraries are not commercial, but rather are directed toward the good of the community they serve. In general, libraries do not produce revenue; from an accounting standpoint, they exist primarily as cost centers. But libraries do consume resources of all kinds, from buildings, network services, and staff to published content, computers, and a vast array of third-party services. These necessities draw them into a marketplace of entities willing and eager to supply those resources.

Most of the vendors that serve the library market operate on a different set of premises; those that constitute the basic principles of capitalism. These companies invest resources to develop products and services which they attempt to sell for more than it costs to produce them. Those products and services that provide sufficient value to libraries typically succeed in the market, and therefore can be sustained. Those that do not provide sufficient value are modified to do so, or eventually disappear. In most cases, multiple entities provide competing products or services, and libraries have an alternative if they are dissatisfied with a particular provider.

Library services, however, face no such competition. A dissatisfied user of one library rarely has the option to patronize another. Each library is free to provide value to users according to its own judgment of what is wanted or needed, within the limits of its budget. That budget rises or falls in proportion to the health of the parent institution and the perceived value of the library to that institution. But in general, libraries do not disappear. They remain in operation no matter what, supported as part of the infrastructure by their funding body.

Both libraries and vendors (at least the good ones) rely on "service" to their respective clienteles to distinguish themselves, but there are important distinctions in their respective definitions of the term. In the commercial world, service must exist within a context of profitability, in which all costs are covered and some additional increment is contributed to the company's continued growth and as a return on the capital initially invested. The library service ethic is much more open-ended and less directly constrained by costs. In public libraries, for instance, the missions of creating an informed citizenry and assuring equality of access to information resources lead to different criteria for success. Public universities (and many private universities and colleges) typically permit local residents who are not affiliated with the university access to at least some of their resources. While the associated costs are not ignored, neither are they a central element in the decision to be "open" in this way. In part, this is because such openness is consonant with the library mission, and in part it's because libraries are not required to generate a financial return to assure their continued existence.

To put it another way, the return on investment for libraries comes in the form of a better-educated (or at least better-entertained) community. Return on investment for a vendor comes in the form of sustainability and growth for its enterprise, and compensation for the use and risk of the capital underpinning its development.

These two value systems—let's call them community and commercial --- fit together comfortably in many respects, provided that everyone understands and accepts the differing roles and rules. The long and successful history of library-vendor interaction demonstrates just how well these disparate value systems can co-exist in a single market. But there are instances where conflicts and fundamental

disagreements arise. A current example is the role of Google in the mass digitization of books held by some major academic libraries. While there are many benefits that might accrue from such a community-commercial partnership, the underlying motivations of Google and its library partners (and eventually, its library customers) are different. Although Google's mission "to organize the world's information and make it universally accessible and useful" resonates very clearly with the mission of libraries, its commercial values dictate that this will first and foremost be accomplished profitably. Libraries, operating from a community-oriented set of values, assert that this mission is inherently valuable, and should be pursued regardless of whether it can be done profitably—that is, it should be subsidized as a public good. And therein lies the rub.

This tension -- between community values and commercial values, between idealism and pragmatism, between social responsibility and private benefit – has deeply affected some aspects of the library market. Cataloging, regarded by many as the heart of librarianship, is one of those areas. The conflict and the underlying issues affecting the cataloging market may be best illustrated by an example: the long-running "pilot project" involving the Library of Congress and the Italian library supplier Casalini Libri. It is a case in which the commercial and community value schemes have come into direct conflict, with questionable results for all sides.

In the fall of 2004, LC approached Casalini Libri to propose a pilot project in which Casalini would supply PCC core-level records and accompanying authority records for the approximately 4,000 Italian books acquired each year by LC. Over the ensuing year, LC trained Casalini staff in the necessary procedures. While LC agreed to pay a significant premium to license these original records, that price accounted for only 1/3 of Casalini's production costs. The intent was that other libraries and the bibliographic utilities (OCLC and RLIN at the time) would also license or purchase these records, assuring that Casalini's investments in space, equipment, staff and training would be covered. Ideally, the firm might ultimately realize a modest profit, while adding original cataloging capacity to the market. In other words, once the program had been jump-started by LC's support and Casalini's willingness to take a significant risk, it was expected that the market would provide a commercial incentive to produce original cataloging records.

For a number of reasons, this has yet to occur, nearly five years into the experiment. First, the bibliographic utilities initially declined to license these records. To protect its interests, Casalini therefore added a "non-distribution" clause to their agreements with participating libraries, including LC. This enjoined those libraries from distributing Casalini records to the utilities or sharing them with other libraries, since that distribution would eliminate the opportunity for any subsequent sale of those records. Absent the non-distribution clause, Casalini would have been left in a position of continually losing money on their cataloging operation. From a commercial viewpoint, there is no incentive to produce records under these conditions—and in fact, there are major disincentives, financial and otherwise. While some aspects of the situation have since improved, "opportunity" of this sort does not look very attractive to other vendors. The impulse to share records for which the costs have not been fully recovered may make sense as a form of community good, but is not sustainable without some form of subsidy or exchange. From the commercial viewpoint, it's simply bad business.

In many respects, this seems like a simple problem to solve, assuming that everyone is willing to be realistic about costs and to recognize that two different value systems are in place. Unfortunately, in our experience, most libraries do not fully understand their own production costs. There are constant assertions that "we could do it cheaper in-house" when confronted with contract cataloging price schedules. While that may be true in some cases, the library's calculations (in those instances where they are actually performed) often exclude overhead costs (employee benefits, equipment, subscriptions, software, phones, space, etc.) and opportunity costs (what else that cataloger could be

doing). In general, libraries understate (or simply don't recognize) the full costs associated with cataloging. This renders questionable any comparison with stated prices from vendors, who typically do have a good handle on costs, since their continued operation depends upon it.

Given that the work is based on standards, original cataloging should require essentially the same amount of effort, no matter who does the work. Each title requires authority work, subject analysis, physical description, and classification. Unless there are training issues, cataloging an individual title should take the same amount of time, again no matter who is doing the work. The cost of production, then, would be similar, though probably not identical. Pay scales can vary significantly; some formats require more time or skill than others; some operations may be better designed; and some individuals more productive than others. But these factors would exert only marginal effect. For any title cataloged to the same standard in two operations, it's reasonable to assume that the task requires a similarly high level of investment.

But let's ignore both the variances and the issue of exactly what the cost per title is or should be, and focus on principles. For ease of calculation, let's posit that, on average, it costs $100 per title to create an original cataloging record. From the commercial vantage point, it is only viable to engage in that activity if $100 in offsetting revenue can be realized, and it's only *attractive* if $110 or $120 can be realized. From the community vantage point, the calculus is entirely different. In that instance, an individual library bears the $100 cost of production. That library may receive some offsetting compensation, in the form of an OCLC credit, but beyond that it does not expect to directly recoup its production costs. Instead, it expects to be relieved of the comparable cost of producing some other record or records that it also needs, by relying on another library to produce them. If a library is especially well funded, or views net contribution of records to the community as a part of its mission, it may not even expect to recoup all of its costs. The prevalence of Z39.50-compliant catalogs, which enable other libraries to use freely the records produced locally (or purchased elsewhere) by other libraries is a characteristic expression of these community-oriented values.

LC's below-cost distribution of its own records is another. For decades now, LC has in effect subsidized the production of cataloging records, by diverting significant staff time from its own collection and users to provide authoritative records for the CIP program. In effect, LC contributes extensively to the common good, producing records upon which much of the library and vendor world depend. At the same time, Title 2, Chapter 5, Section 150 of the US Code states that:

> *"The Librarian of Congress is authorized to furnish to such institutions or individuals as may desire to buy them, such copies of the card indexes or other publications of the Library as may not be required for its ordinary transactions, and charge for the same a price which will cover their cost and ten per centum added, and all moneys received by him shall be deposited in the Treasury and shall be credited to the appropriation for necessary expenses for the preparation and distribution of catalog cards and other publications of the Library."*

This law dates back to 1902, and is clearly based on the premise that all such distribution of records will occur in printed form, i.e., a single copy to a single buyer. By design, its cost recovery provision relates solely to the cost of distributing the records and is not intended to recoup the cost of production. The assumption appears to be that LC is creating these records for its own collections, and that distribution is simply a by-product of the Library's internal activity.

Much has changed since this law was formulated. First, LC took on a community-oriented role by underwriting the CIP program, which accounted for 53,000 new titles in 2008. Second, for the past 25

years or so, LC records have been distributed electronically. This has not only lowered the cost of distribution, but has made the records easily transferable from one institution to another, often without payment. One result is that LC records are significantly underpriced, since the cost of production is not included. Another is that an entire industry has developed around free (or at least very cheap) MARC records. Consider that an LC record for a single title might appear in thousands of library catalogs, while its MARC Distribution Service lists only 74 customers, 30 of them foreign. Most copies of LC records are obtained either free (via its Z39.50 servers and WebOPAC) or purchased from OCLC or vendors who supply those records in conjunction with the materials they sell. In short, many libraries and vendors benefit from a product for which production costs are not recovered.

This can only be described as a subsidy. That subsidy is sustainable only as long as the original production is funded. In exchange for its efforts with CIP, LC does receive a free copy of each of the 53,000 titles submitted to the program. Many of these are not appropriate for LC's own collections, but are used in its exchange programs with foreign libraries, to obtain titles of greater interest. Nonetheless, it seems highly unlikely that these relatively modest offsets (some of which carry additional costs of their own) compensate fully for LC's investment in creating records. At bottom, some portion of LC's cataloging staff is engaged in work that, while beneficial to many other libraries, vendors, and cooperatives, is not essential to LC's own operations.

Over the past five years, LC has absorbed significant budget cuts, and faces continuing pressure. It has undertaken major staff reductions, especially in its cataloging operations. CIP, as a program that is not directly related to LC's mission, and for which the costs of production divert staff resources from other programs, must obviously be considered for adjustment. Given the existing level of dependency, such a change would affect the entire profession and the industry based on that profession.

Perhaps more insidiously, LC's subsidization of a major portion of the cataloging stream reinforces the perception that cataloging costs less than it actually does. Combined with the library community ethos that data should be shared freely, and a limited understanding of cataloging costs within individual libraries, the result is a distorted market. That is the state of things today. The market for cataloging records is in some important respects dysfunctional. In our view, the biggest issue is that the market lacks sufficient incentives to stimulate the production of new cataloging records. Obviously, many books, journals, electronic resources and other items are being cataloged, so some elements of the market are working. But structurally, it seems clear that something is amiss.

OCLC, union catalogs, and cooperative cataloging are some of the primary tools for sharing cataloging workloads, and these do provide mechanisms for exchanging records and labor more efficiently. But they don't directly stimulate production. Backlogs continue to grow in many libraries, despite the fact that adequate capacity exists in the library community as a whole. Our survey results also confirm our direct observation of many "aging" backlogs in libraries. Because of their own staffing constraints, or unwillingness to bear the cost of original record creation, many libraries simply wait for another library to catalog an item they have already received. On average those items are held for three to six months, with periodic searches of OCLC to determine whether another library has blinked. While this makes sense as a way of controlling costs, it does not provide optimal service for users.

Cooperative cataloging programs, with few exceptions, exhibit much of the same reluctance to contribute. The Program for Cooperative Cataloging (PCC) is perhaps the most highly-developed and successful of these, and PCC libraries do contribute significant numbers of records through the BIBCO, CONSER, NACO, and SACO programs. In 2008, for example, BIBCO libraries produced 76,572 new bibliographic records (which were included as part of 316,000 new records distributed in the "LC

Complete" product.) While this is clearly a substantial contribution, it's far from what it could be. For one thing, BIBCO includes fewer than 50 libraries. (To put that in perspective, there are more than 4,000 academic libraries in North America, almost all of which employ at least one professional cataloger.)

In addition, BIBCO contributions appear to have leveled off at about 77,000, after peaking at more than 80,000 in 2002:
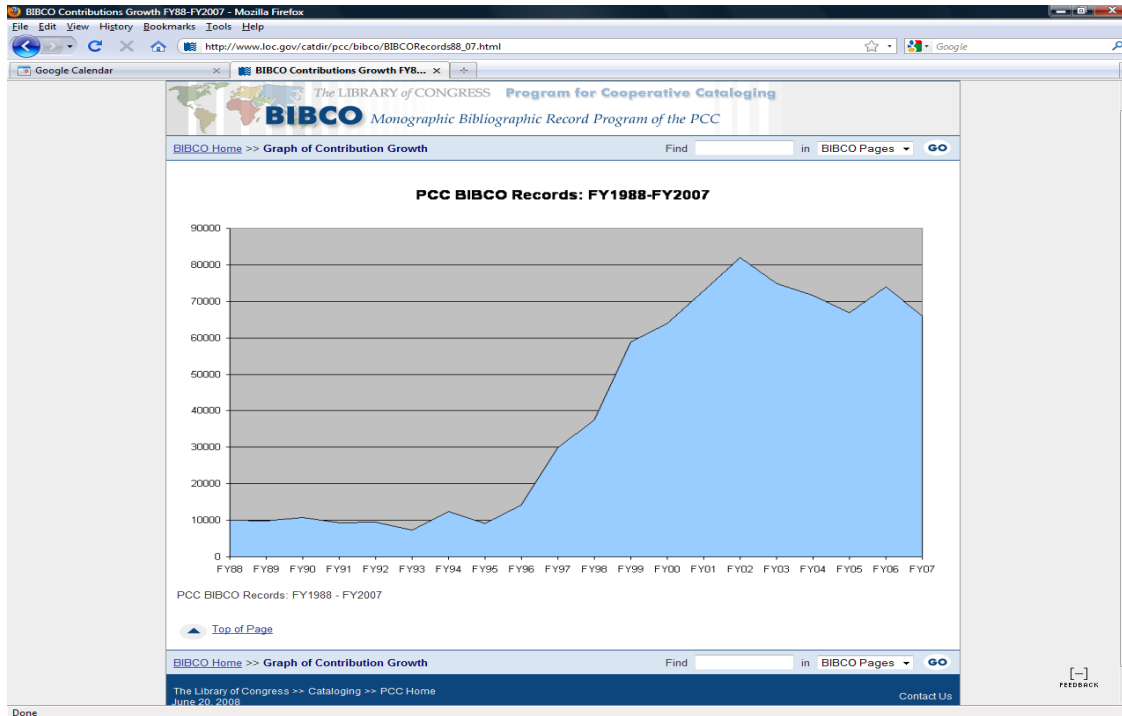
Finally, a mere 10 of those member libraries accounted for 2/3 of all the BIBCO records produced in 2008.

Top 10 BIBCO Contributing Libraries: 2008

Library Name	BIBCO Records Created
Stanford	6,581
Chicago	6,367
National Library of Medicine	6,357
Indiana	6,288
Columbia	5,734
Government Printing Office	4,960
Princeton	4,915
Cornell	3,965
Wisconsin-Madison	3,142
Yale	2,702
TOTAL:	51,011

It would be interesting to explore more fully some of these dynamics. Why do so few libraries join BIBCO or CONSER (which also relies on fewer than 50 members)? And what factors determine which records are contributed? It seems likely that titles most needed and valued locally by the contributing library would take priority. Because most of the big PCC contributors are research libraries, it's also likely that many titles are specialized, and therefore may not be widely held. Therefore, fewer libraries would benefit from that contribution, as opposed to, say, DVDs, which are cited as a major problem by most libraries in the survey. Regardless, these libraries are committed to contributing to the community. Which titles they catalog would matter less if participation were broader. A cooperative system only works well if everyone participates.

Vendors and contract cataloging services also create a substantial number of records. For those whose primary business is selling materials, however, those records are provided as adjuncts to their main products. In most cases, these are either LC-created records or something less than a full cataloging record. Contract cataloging services can and do create full original records, and charge accordingly. This is perhaps an example of where a viable incentive to create records does exist.

Some factor or combination of factors is preventing libraries from participating fully in producing and sharing original records. While it might be attributed to staffing shortages in some locales, our survey indicates that enough capacity exists in the overall community to handle the volume of new material. It may be allocated unevenly, but some capacity is simply used poorly, as indicated by the extent of redundant work measured by our survey.

In a market that is functioning optimally, capacity would be attracted to demand by some type of incentive. That incentive could be community-based or it could be commercial, as long as the costs of production are covered. But somehow the incentives to produce new cataloging records are insufficient, from both the commercial and community viewpoints. Otherwise, there would be greater participation in cooperative programs, and/or more vendors seeking to become cataloging agencies. At bottom, we believe this is because cataloging costs and therefore prices are understated and artificially depressed. An even more sobering possibility is that the profession does not believe that cataloging is worth what it costs to create it; that will be quickly determined once all production costs are factored into the price. In the meantime, this is a market that appears to require adjustment.

As described in the previous section of the report, the North American MARC record marketplace is complex, defined by few mainstream distribution models or consistent pathways. In seeking a simplified "big picture", and to highlight market conflicts, we have found it useful to recognize three discrete tiers of activity. Any single entity (producer or distributor) may operate on more than one level but fundamentally each tier is distinct. The colored diagram on page 32 illustrates these three tiers as concentric circles:

> **The traditional (green) tier** encompasses the oldest, most traditional segment of the market, in which nearly all MARC records originate. The entities in this core create and/or sell, and/or buy MARC records. MARC authority records are also created and distributed here, and valued/maintained in a way that they are not in the other tiers of the market. Nearly 200 businesses are engaged in the sale and distribution of bibliographic and authority records and services, or in the sale of systems or other products that incorporate large bibliographic databases. These commercial entities are more fully described in a previous section of the report.

Those libraries that create and formally contribute original records to OCLC (PCC members and others) reside in this core as do all other OCLC member libraries and consortia. Libraries that purchase records from one or more third parties also inhabit this "traditional tier," which we estimate to include more than two-thirds of North American libraries:

- 97% of academic libraries operate primarily within the traditional green tier
- 63% of public libraries operate primarily within the traditional green tier
- 65% of school libraries operate primarily within the traditional green tier

The opportunistic (blue) tier surrounds the traditional green center and is populated with the remaining third of North American libraries, including non-OCLC libraries and underfunded libraries without adequate cataloging capacity. The "blue tier" is also home to open database providers, and the pervasive Z39.50 protocols used to locate and obtain MARC records free of charge. Library and non-library entities in this tier take full advantage of ubiquitous bibliographic data. Both in Canada and in the US, historically "green libraries" are adopting "blue tier" practices and expectations, as library budgets are cut and as Z39.50 targets proliferate. Nearly all libraries, regardless of size or type are strategically patient, periodically re-searching the "blue tier" for certain records to become "available"; but for "blue-tier" libraries, this is the primary approach to cataloging. Many libraries in this tier have partial, outdated, or no authority control over their bibliographic data. Open Access and Open Archives Initiatives reside in the blue tier, strongly supported by the basic philosophical stance that access to information should be free.

The non-library (purple) tier operates to a large extent without appreciation for or experience with MARC records, and without much regard for the library market in general.   It is important to remain aware of activity in this segment, of course, because developments here pose the most significant competitive threats to the traditional values and economic structures of the "traditional green tier," and even the "opportunistic blue tier."   This is the place where newer technologies and non-MARC data formats are used and developed.   Mass digitization projects are occurring here, and many initiatives are underway to solve authority control issues with non-MARC tools. Entities in this "purple tier" are taking considerable advantage of bibliographic data that flows so freely in the "blue", and are re-inventing primary research and access tools that have traditionally been the domain of libraries. As we all know, both "blue tier" and "green tier" libraries are pointing patrons to "purple tier" links and services like Google Book, Search Snippets, Google Scholar and Amazon's cover scans, reader reviews, and "Look Inside the Book" features.

LC itself clearly resides in the "traditional green tier" as it creates, sells and distributes so many records. A large and perhaps greater portion of its activity, however, falls into the "blue opportunistic tier" insofar as its cataloging records are also available free of charge via the open web and via Z39.50 download.

More than 900 "green tier" libraries and consortia have Z39.50 compliant databases, so in our diagram, they reside in both the "green and blue tiers." Some of the biggest, most entrepreneurial libraries operate in all three tiers of the market, but their number is relatively small.

With its very large "green core membership," its multiple business models and commercial products, its inventive development agenda, and its broad range of market relationships, OCLC pro-actively operates within the "traditional green tier" and within the "purple non-library tier."   OCLC member libraries, however, are also very active in the "opportunistic blue tier," sharing records in ways that may conflict with OCLC's proprietary intent. Their highly controversial record use policy was intended to address this
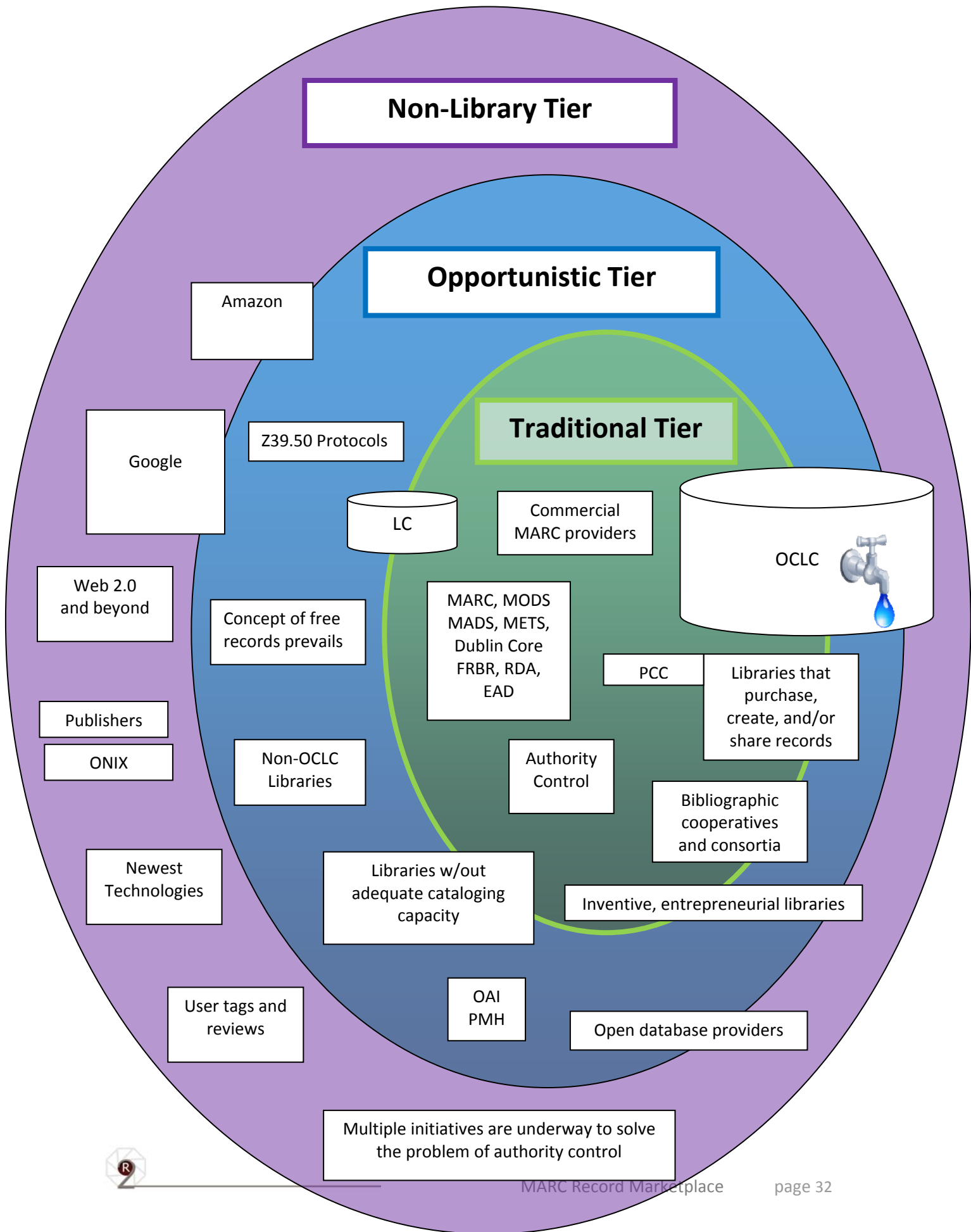
conflict, but the size and strength of the "blue tier" and the tradition of cooperative cataloging could make the controversy difficult to resolve.[17]

It is interesting to note that many generations and varieties of MARC and non-MARC records have proliferated within the "traditional green tier", including USMARC, CAN/MARC, MARC 21, MODS, MADS, METS, Dublin Core, FRBR, EAD, and RDA. While the "blue tier libraries" utilize MARC21 records (local ILSs require MARC) few have sufficient cataloging capacity to create the new ones they need, and even fewer have embraced any additional record types. None of these record types has been adopted as a primary data storage format by any "purple tier entities."

It is also interesting to note that the non-library tier expects patrons and other end users to be working around the perimeter. The traditional and opportunistic tiers may expect their users to be in the center of the circle—i.e. assuming that MARC records offer the highest quality of access.

ONIX to MARC record translations and fully operable MARC to non-MARC metadata crosswalks could dramatically alter this three-tiered landscape.   To date, major players in the blue and purple tiers have failed to buy into the concept of shared bibliographic and authority data. While some efforts to encourage cross-market cooperation are underway (notably the OCLC/NISO forum), fierce competition flourishes within and between each tier of the market. Even more problematic, each tier has distinctly different needs and incentives, making it difficult to establish an adequate degree of shared urgency and/or investment in new solutions.[19]

Non-Library Tier

Opportunistic Tier

Amazon

Z39.50 Protocols

Traditional Tier

Google

LC

Commercial MARC providers

OCLC

Web 2.0 and beyond

Concept of free records prevails

MARC, MODS MADS, METS, Dublin Core FRBR, RDA, EAD

PCC

Libraries that purchase, create, and/or share records

Publishers

ONIX

Non-OCLC Libraries

Authority Control

Bibliographic cooperatives and consortia

Newest Technologies

Libraries w/out adequate cataloging capacity

Inventive, entrepreneurial libraries

User tags and reviews

OAI PMH

Open database providers

Multiple initiatives are underway to solve the problem of authority control

V. The Economics of Cataloging

The practice of cataloging has never before faced the level of scrutiny it now enjoys … or endures. Two types of question predominate. First, are traditional cataloging and the MARC record—even after modernization by RDA and FRBR—still necessary in an era of full-text indexing, OpenURL linking, and other discovery options? While this is a worthy question, it is fortunately not within the purview of this report. As described below, it is clear from the survey results that MARC records remain a basic requirement of library—and therefore vendor-- operations. While it is vital to attend to the evolution of discovery options and non-MARC metadata, our working assumption is that the MARC cataloging record will remain important for the next five to ten years.

The second type of question, which is directly relevant to this project, relates to the economics of cataloging. How do we as a profession understand and explain the costs and benefits of producing and distributing cataloging records? Where and by whom are most original records produced? What incentives exist to stimulate production? What are the barriers that discourage production? How does the library market assign value to the work of cataloging? What is the return on any organization's investment in producing original catalog records? How does shared cataloging and free or low-cost distribution of records affect the market? To what degree is market activity subsidized by LC and by the work of individual libraries? Some of these questions have been addressed in the previous section.

As noted there, the market is in need of adjustment, if it is to create an incentive for producers while retaining the community ethic of free sharing of data. The ethic of the cooperative can only be sustained if the full costs of production are borne by the community.

The advent of the MARC format and cooperative cataloging in the early 1970s introduced the tantalizing possibility of describing and classifying an item once and using the resulting record many times. But it was not until the widespread adoption of online catalogs in the 1980s that those electronic records could be shared easily. Even then, substantial re-keying, awkward tape loads, and retrospective conversion were needed in order to benefit from cataloging work done by LC and by other libraries. The widespread adoption of integrated library systems (ILS), coupled with the availability in electronic form of authoritative LC records, ushered in a period of radical efficiencies and cost reductions for those libraries prepared, both technically and culturally, to take advantage of them. Reliance on LC records grew as they became more easily, inexpensively, and in some cases "freely"available.

OCLC , RLIN and other bibliographic utilities and cooperatives grew by providing a valuable mechanism for exchanging records and effort, supporting and extending the efficiencies of LC records and shared cataloging. They offered libraries a way of finding and obtaining a record which others had created or enhanced, and a corresponding means to contribute new or upgraded records created locally. Although some individual libraries contributed to these databases from the outset, the Program for Cooperative Cataloging (PCC), begun in 1995, formalized and organized those contributions to still greater effect. PCC, enabled by OCLC and RLIN, increased the library community's overall capacity for producing authoritative records, by creating standards and reducing duplicative work. As of March 31, 2009, PCC libraries have created 976,404 bibliographic records, along with millions of name and subject authority records, since the Program's inception.

Through services such as OCLC's WorldCat Cataloging Partners, MARCIVE's authority control, and many others, vendors of all kinds have made enormous contributions to the overall efficiency of distributing MARC records, and to a lesser degree to creating them. While major economies of scale have been realized, troublesome issues remain. First, there are thousands of small libraries that operate outside of

the shared cataloging infrastructure. Most of these lack the capacity to produce MARC records—they have no catalogers. Second, catalogers have an almost unstoppable urge to improve, tweak, customize and "localize" national-level records; redundant work is still widespread. Third, despite efficiencies, cataloging backlogs continue to grow, not only for audio-visual materials, rare books, and non-Roman languages, but even for the most commonly-held materials.

Many distribution improvements have taken place since 1995, when the Web came into prominence. Instead of boxes of cards, 9-track tapes or CD-ROMs, FTP and Z39.50 protocols and large-scale batch processes and matching algorithms came to provide more convenient and selective delivery. Distribution and re-distribution of records became much more convenient and cost-effective, and ultimately attracted new players. Library Thing, Biblios.net, skyriver.com and others have seized on the opportunity provided by the "free" or nearly free availability of MARC records. The interest in these firms indicate that: 1) demand for MARC records persists and has even proved of interest to individuals; and 2) transaction costs for distribution are trivial.

While the library world was seeking these efficiencies, of course, other forces were also at work in the realms of search and discovery, and the data and techniques that make those activities possible. Google proved that it could index the Web. Amazon described books in a way that was much more compelling to users than library OPACs. Tables of contents, user reviews, snippets and "looking inside the book" became routine. OpenURL linking created direct access to the full text of journal articles in a way that made MARC records seem quaint. In short, outside the cataloging world, people and companies began creating better ways to find information. New metadata and discovery "rules" began to challenge those long held by libraries.

We now operate in a context where questions about the efficacy of the MARC record and the centrality of the OPAC are continually posed. We wrestle with keeping libraries relevant, and assuring their participation at the network level. We have also entered an era where questions about cost and return on investment are routinely asked of non-profit entities such as libraries. In the long run, there may be better and cheaper alternatives than MARC. In the short run, there may be ways to reduce the cost of producing MARC records.

Nonetheless, for the moment MARC remains central for libraries. In large measure this is due to the installed base of library systems, which expect and work well with this data exchange format. This will continue to be true until the next generation of discovery and inventory systems are in place. But its limitations are increasingly clear. LC's own ILS cannot, for instance, accept ONIX records directly – they must first be converted to MARC21. This will undoubtedly change over time, but for now, most libraries will continue to need cataloging records delivered in MARC format—it is the only *usable* solution.

There remain strong arguments for use of standard cataloging principles---controlled vocabulary, classification, subject analysis, and authority control—packaged and delivered in a consistent format. While MARC records may need to be extended, embellished (supplemented with full text, flap copy, excerpts, user tags), for now they provide a common standard and a cooperative infrastructure that controls costs. In the long term, there may emerge better solutions. For at least the next 5-10 years, however, continued savings can be realized by improvements to the production and distribution systems for cataloging records.

The market for these cataloging records involves a number of related components and functions:

Production: Incoming streams of material and records to LC in 2008 included CIP titles (53,000), PCC BIBCO records (76,000), CONSER records (30,000 est.), selected titles from LC's Copyright Receipt Office, LC's Prints & Photographs Division and others. NACO libraries contributed 188,000 new name authority records. From these inputs, in 2008 LC produced and distributed 316,000 new bibliographic records and more than 250,000 authority records, along these lines:

Materials Type/Format	Number of Records
Books All	250,000
Music (printed music & sound recordings)	30,000
Serials (LC and CONSER)	25,000
Visual Materials (motion pictures & videos retained for LC's Collections)	5,000
Maps	5,000
Computer files	1,000
Name Authorities	250,000
Subject Authorities	6,000

Distribution: LC records are distributed in a variety of combinations through its MARC Distribution Service. At present, there are approximately 74 customers for the various MDS products, and most of those entities comprise an extensive secondary distribution market through vendors, bibliographic utilities, etc. The map provided as Appendix B details those dependencies.

Another major distribution channel involves direct downloads from LC's Voyager database. At present, LC offers four separate interfaces:
- A Web OPAC for bib records that supports 875 simultaneous users
- A Web OPAC for authority records that supports 500 simultaneous users
- Z39.50 direct access for users with Z39.50 clients, which supports 340 simultaneous users
- Z39.50 gateway interface that supports up to 250 simultaneous users

In total, these search interfaces process about 500,000 searches each business day. While not every search leads to a download, the volume of searches is a clear indication of interest. Major users, to the degree that can be determined, include school libraries and small publics, who may not be OCLC members. In addition, vendors, open database providers, and firms such as Amazon regularly seek these records.

Adding Value: Once records have been distributed to vendors, most seek to add value. In general, this means matching records to a group of titles being shipped, and adding fund, location and electronic invoicing data, updating or adding proxy prefixes to URLs. These value adds are important to workflows, but do not necessarily change the bibliographic data (although some vendors do perform CIP upgrades if they are needed). Once libraries receive the records, via OCLC, the vendor, or another source, many also seek to add value in other ways. As noted in the survey results, 80% of libraries perform some degree of local editing on the records, to customize them for their own constituency. Increasingly, libraries are adding or linking to table of contents, Amazon, and other external sources in order to enrich the bibliographic description.

Capacity: Although shared cataloging, bibliographic utilities and vendor participation in cataloging have all contributed to reductions in redundant effort and increases in efficiency, the increases in backlogs identified in the survey remains a cause for concern. On the surface, their persistence would suggest that there is insufficient cataloging capacity in the system. However, as noted in the Libraries section of this report, this does not appear to stand up to the most basic questions. In looking at the median numbers of original catalogers reported, we estimated that well over 30,000 professional catalogers are at work in North America. In the earlier example, we suggested that if each of those catalogers were to produce one record per work day, that would provide the capacity to create 6.8 million records per year. In addition, there are hundreds of additional catalogers working for vendors—more than 300 were reported by the 70 respondents to our survey. But let's discount those for now. Let's also reduce our estimated output by another 50%--i.e., assume that each cataloger produces a record only every other day. Even under that scenario, there should be capacity to produce more than 3 million original records per year in North America.

While it is true that subject, language and format expertise would limit the ability to distribute responsibility fully, there is clearly adequate capacity within the library community. We should not have backlogs. Even if we reduce the available hours by another 50%, to account for management responsibilities, involvement in non-MARC metadata, professional development, committee work, managing outsourcing contracts and other activities, that still leaves capacity to produce 1.5 million original records per year. Why are we as a group unable to achieve this level of production?

It seems like a fair question. Some answers are obvious, e.g., the capacity is poorly allocated. The necessary hours are not aligned with the materials that need attention. In theory, OCLC, PCC, and other cooperative mechanisms help with this, and as we've seen, they do work well in some respects. But, especially for print and other tangible items, cataloging requires that the object be in hand as it is cataloged. That is clearly one barrier, at least for now. Electronic materials, of course, lend themselves much more readily to remote cataloging, and use of surrogates is becoming somewhat more acceptable. But there still seems to be more needed here.

Another factor affecting shared capacity is the increasing emphasis on making accessible material that is unique locally. This shifts priorities within a given library, and makes it more difficult to share capacity, and to dedicate hours to work that might benefit the community as a whole over one's own institution.

Expertise: As the profession ages and cataloging practitioners retire, there are likely to be fewer traditional catalogers. Already, professional cataloging talent and hours are often distracted from production work, as catalogers are drawn into management, training of paraprofessional staff, non-MARC metadata, outsourcing contracts, taxonomy, content management systems, "localized" vocabularies, and customization of local content for users. While this bodes well for transforming the profession, in the short term it will create gaps in expertise.

Realizing Value: One of the most striking realizations during our work on this project is the full extent to which the work of LC is used and re-used. While it is true that the records produced by LC need to be better supported, it is difficult to imagine the profession and the industry without them. They provide enormous value, to a degree that is difficult to calculate.

In 1995, however, Paul Kantor of Rutgers University attempted to do just that, by building a sophisticated model to track potential savings to the library community. In his paper, entitled "Savings to the Nation Resulting from the Existence of LC Cataloging Records"[8], he estimated that "savings from specific sectors total $268 million." His summary table looked like this:

Library Type	Materials Type	Estimated Savings
Academic	Monographs	$26,652,373
Public	Monographs	$87,200,822
Special	Monographs	$15,328,483
Government	Monographs	$ 9,175,381
Academic	Serials	$ 6,075,161
Public	Serials	$ 4,794,666
Special	Serials	$ 5,822,526
Government	Serials	$ 1,705,509
Academic	Special	$14,314,388
Public	Special	$12,156,530
Special	Special	$ 3,229,383
Government	Special	$ 3,900,026
Schools	All	$77,845,833
TOTAL		$268,201,083

It's not necessary to understand every aspect of his methodology, or to agree with all of his assumptions, or to accept these exact numbers to recognize the sheer magnitude of LC's impact. This is confirmed by the comments we've received from libraries, especially school libraries, who "could not function" without LC records. It's also confirmed by the fact that an entire industry has been built on this foundation, as a careful look at the map in Appendix B will show; and confirmed again by the 500,000 searches per day against LC's database. It's somewhat surprising that this work was never widely disseminated or acted upon. While much has changed since 1995, the essential patterns remain. And even if the savings to the nation were 10% of this total, LC's overall value to North American Libraries was then, and continues to be profound.

Innovation: Innovation in the cataloging market of necessity follows two tracks. One of those continues to assume the centrality of the MARC record, and to suggest changes to the process and the market that continue to reduce costs, improve the product, and fine tune the market. As an example of this:

- Option 1: Centralize creation of cataloging records to the highest degree possible. In effect, continue and even expand the work that LC now performs on behalf of the library and vendor community. This could be achieved by funding LC adequately to produce all of the necessary records in a timely manner without adversely affecting its own internal operations. At least some of the necessary funding should be provided by a pricing model that assures the cost of

production is covered. Of course, this centralized operation does not necessarily have to be run by LC.

- Option 2: Improve the incentives to contribute original cataloging records. The mechanisms and programs for this sharing already exist. What's needed is a change in the effort/reward ratio, in which those entities creating records are fully compensated for the associated costs. That compensation could be shaped to either community or commercial values, or some combination of them. The key is that the compensation be adequate to the task.

But, as we've alluded, the realms of metadata and discovery are rich with other ideas, and connecting the activities of MARC cataloging with these emerging techniques will be increasingly important. Again, two examples drawn from the PCC Web site illustrate the point:

- Option 1: PCC's strategic plan for 2006-2010 includes two initiatives that presage a more expansive view of the bibliographic universe.[12]

  - PCC's Strategic Direction (SD) 2, headed "Redefine the Common Enterprise" proposes to "explore potential new economic models for cooperation that cut across traditional sector boundaries by identifying and exploiting common metadata needs."

  - SD3, another major priority, states that PCC will "Build on and expand partnerships and collaborations in support of the common enterprise" – including "develop close working relationships with publishing entities to promote arrangements for data sharing." Integration into the larger community in this manner is critical to retaining relevance.

- Option 2: A quote from NYU cataloger Sherman Clarke, suggests a different take on the work of resource description: "We collectively need to have a model that allows us to do some of the building of BIBCO records mechanically or through accretion of metadata from institutional records or other record loads. OCLC already does considerable building of the master record from incoming records; what we need is something more like the metadata that is becoming usual in NewGen environments. If someone adds a tag or review or picture, that becomes available in the master cluster. Not a BIBCO record, but a BIBCO cloud of metadata for a particular manifestation of a work/expression."

VI.   Closing

Studies like this one can have enormous value, but naturally they have limitations. As we knew from the outset, the scope of the project was broad and complex and we found the analytical process challenging. We're certain that we've made mistakes and we welcome scrutiny by those with expertise that exceeds our own. Without doubt, we have oversimplified. As we sought to understand the big picture and to identify patterns that are obscured by too many details, it is quite possible that we buried or ignored critical elements or influences. Nonetheless, we are hopeful that our interpretation of the data and our overall description of the market are fresh enough, and at the same time valid enough to inspire new ways of thinking about the future.

Bibliography of Related and Referenced Works

1. Beheshti, J. (2002). *The use of NLC MARC Records in Canadian libraries, Phase I: University and large urban public libraries, final report, 31 March 2002*. Montreal, McGill University, Graduate School of Library and Information Studies. Executive Summary at: http://www.collectionscanada.gc.ca/cataloguing-standards/040006-2211-e.html

2. Beheshti, J. (2003). The use of NLC MARC Records in Canadian Libraries,. Phase II. Federal government and non-public CALUPS public libraries. Report prepared for NLC. 91 p.

3. British Library. (2003). *Measuring our value: results of an independent economic impact study commissioned by the British Library to measure the Library's direct and indirect value to the UK economy*. Available at: http://bl.uk/aboutus/stratpolprog/increasingvalue/measuring.pdf

4. Calhoun, Karen. (2006). *Changing nature of the catalog and its integration with other discovery tools*: Report prepared for the Library of Congress. Available at: http://www.loc.gov/catdir/calhoun-report-final.pdf

5. Danskin, Alan. (2006). *Tomorrow never knows: the end of cataloging?* Paper presented at IFLA, Seoul, Korea. Available at: http://www.ifla.org/IVifla72/papers/102-Danskin-en.pdf

6. Indiana University, Task Force on the Future of Cataloging. (2006). *White paper on the future of cataloging at Indiana University.* Available at*:* http://www.iub.edu/~libtserv/pub/Future_of_Cataloging_White_Paper.pdf

7. Institute of Museum and Library Services. (2008). *Public libraries survey: fiscal year 2006*. Available at: http://harvester.census.gov/imls/pubs/publications/pls2006.pdf

8. Kantor, Paul. (1995). *Savings to the nation resulting from the existence of LC cataloging records.* A consulting report*.* Library of Congress. Author website: http://www,paulbkantor.com

9. Kemp, Rebecca. (2008). MARC record services: a comparative study of library practices and perceptions. *The Serials Librarian*, *55* (3). Available at: http://library.uncw.edu/web/faculty/kempr/kemp-MARC-record-services-20080328.pdf

10. Leysen, Joan M. & Boydston, Jeanne M.K. (2005). Supply and demand for cataloguers present and future. *Library Resources & Technical Services, 49* (4), *250-265*. Available at: http://www.accessmylibrary.com/article-1G1-138996260/supply-and-demand-catalogers.html

11. Library of Congress. (2008) *On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control.* Available at: http://loc.gov/bibliographic-future/news/lcwg-ontherecord-jan08-final.pdf

12. Library of Congress. (2009). *Strategic directions for the Program for Cooperative Cataloging (PCC) 2007-2010.* Available at: http://www.loc.gov/catdir/pcc/stratdir-2008.html

13. Library of Congress, Cataloging in Publication Division. (2007). *CIP poised for change: survey findings and recommendations for the 2006 CIP Review Group.* Available at: http://cip.loc.gov/CStpcontentsacknow.pdf

14. Luther, Judy. (2009). *Streamlining book metadata workflow: a white paper prepared for the National Information Standards Organization (NISO) and OCLC Online Computer Library Center, Inc.* Available at: http://www.niso.org/publications/white_papers/StreamlineBookMetadataWorkflowWhitePaper.pdf

15. Marcum, Deanna. (2008). *Deanna Marcum's response to On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control.* Available at: http://www.loc.gov/bibliographic-future/news/LCWGResponse-Marcum-Final-061008.pdf

16. OCLC Research. (2009). *What we've learned from the RLG Partners Metadata Creation Workflow Survey.* Available at: http://www.oclc.org/programs/publications/reports/2009-04.pdf

17. *Policy for use and transfer for WorldCat® records.* Available at: http://www.oclc.org/worldcat/catalog/policy/recordusepolicy.pdf

18. Primary Research Group. (2008). Academic library cataloging practices benchmarks. ISBN:1574401068. Research and Markets, Dublin, Ireland. Available at: http://www.researchandmarkets.com/reports/659569

19. Research Information Network. (2009). *Creating catalogues: bibliographic records in a networked world.* Available at: http://www.rin.ac.uk/files/creating_catalogues_REPORT_June09.pdf

20. Schonfeld, Roger C. & Lavoie, Brian F. (2006). Books without boundaries: a brief tour of the system-wide book collection. *Journal of Electronic Publishing*, 9(2). Available at: http://quod.lib.umich.edu/cgi/t/text/text-idx?c=jep;view=text;rgn=main;idno=3336451.0009.208

21. Thomas, Sarah. (1996). Quality in bibliographic control – history and future of cataloging in libraries: perspective on quality in libraries. *Library Trends, Winter 1996*, 491-505. Available at: http://hdl.net/1819/2669

22. University of California Libraries Bibliographic Services Task Force. (2005). *Final report: rethinking how we provide bibliographic services.* Available at*:* http://libraries.universityofcalifornia.edu/sppag/BSTF/Final.pdf

23. Wilder, Stanley J. (2002). Demographic Trends Affecting Professional Technical Services Staffing in ARL Libraries. *Cataloging and Classification Quarterly,* 34: 1/2.

24. Wolven, Robert. (2008). In search of a new model: Columbia University Libraries: Robert Wolven reflects on what's next for cooperative cataloging. *netConnect*, 1/15/2008. Available at: http://www.libraryjournal.com/article/CA6514925.html

Appendix A

MARC Record Distributors and Service Providers invited to participate in the online survey:

Ŧbiblios.net
102 NovaDoc
ABC-CLIO
ACLS Humanities E-Book
Actrace
A-G Canada, Ltd.
ALEPH 500
Alexander Street Press
Alliance Entertainment
Ambassador Books and Media
Amicus
Amigos Library Services
Auto-Graphics
Aux Amateurs De Livres
AV Café
Axiell
Backstage Library Works
Baker and Taylor
BBC Audio Books
Bibliobase
BiblioCenter
BiblioMondo
Blackstone Audio
Blackwell Book Services
Bolinda Publishing
BookLynx
Books on Tape
BookSystems
BookWhere
Bound to Stay Bound
Bowker
Brill Publishing
Brodart
BUSCA, Inc
California Digital Library
CARLI and I-share
Casalini Libri
Cassidy Cataloguing Service, Inc
Catalis
Central Kansas Library System
CLICnet
College Center for Library Automation
Compact Disc Source

COMPanion Corporation
Concourse
Cooperative Computer Services
Coutts Information Services
CRCnetBASE
Credo Reference
Cuadra Associates
CUFTS2MARC
Cybertools for Libraries
D.K. Agencies
DA Information Services
Davidson Titles
Distribution Audio and Video
Documents Data Miner
East View Information Services
Eastern Book Company
Ebook Library
Ebrary
EBSCO
eDuke Books
Elsevier Engineering Information
Emery-Pratt
Endeavor Voyager
EOS.Web
Equinox Software
Erasmus Boekhandel
Evergreen
Ex Libris
Findaway World
Florida Center for Library Automation
Follett/BWI
Gale
Gibson Library Connections
GIS Information Systems
Greenwood Press eBooks
Harrassowitz
Hartman Cataloguing
Howard Karno Books
Humanities E-Book
Iberbook:
Impact/ONLINE CAT
Impact/ONLINE MARCit
India for Everyone

Infor

Ingram Library Services

Inmagic

Innovative Interfaces, Inc.

Insignia Library System

ITS.MARC

Janium

Karger

Kelowna Software, Ltd.

Keystone Systems

KLAS

Knovel

Koha

Landmark Audio

LASERQuest CD-ROM Cataloging System

Leila Books

Liberty Automated Library

Library and Archives Canada

Library Integrated Solutions & Assoc. Inc.

Library Concepts

Library of Congress

Library Services and Systems (LSSI)

Library Services Centre

Library Technologies, Inc (LTI)

Library World

Library.Solution

LibraryCom.com Services

LibraryThing for Libraries

LibriVision

Lindsay and Croft

Listen and Live Audio

Louisiana Library Network (LOUIS)

Mackin Library Media

Majors Scientific Books, Inc.

MARC Link Retrospective Conversion

MARC Record Distribution Service (LAC)

Marcive Inc

MassCat

Matthews Medical and Scientific Books, Inc.

Medialog, Inc.

Midwest Library Service

Midwest Tape

Minaret

Minitex

Missouri Library Network Corporation (MLNC):

Mitinet, Inc.

Multi-Cultural Books and Videos

Music Library Service Company

NACO Canada

National Library of Medicine

Naxos

NetLibrary

NOTEbookS Library Automation

Nylink

OCLC

OCLC Canada

OhioLink

OpenGalaxy Plus

Orbis Cascade Alliance

Ottowa Public Library

Overdrive

Ovid Technologies

Oxford Reference Online

Perma-Bound

Polaris Library Systems

PortFolio

Precision One

Program for Cooperative Cataloging

Project MUSE

ProQuest

Puvill Libros

Quality Books, Inc.

R.R. Bowker

Rand Corporation

Recorded Books

Regent Book Company

Researcher

Rittenhouse Book Distributors, Inc.

Robert A. Schless & Co. Inc

Russian Publishing House

Safari Books Online

Sagebrush Corporation

Savia Library Services

Serials Solutions

SFX Marcit!:

Simon Fraser University

Sirsi

SkyRiver

Special Libraries Cataloguing Inc.

SmartMARC

SmartPORT (SIRSI)

SPACEPOL Biblioblog

STAT!Ref

Sunrise International, Inc.

SUNYConnect

Surpass Software

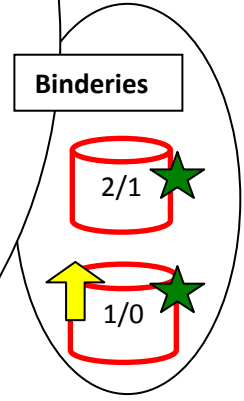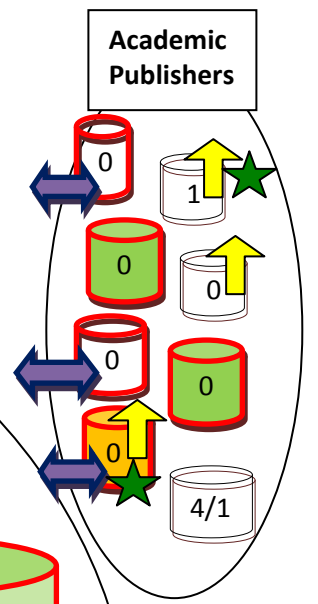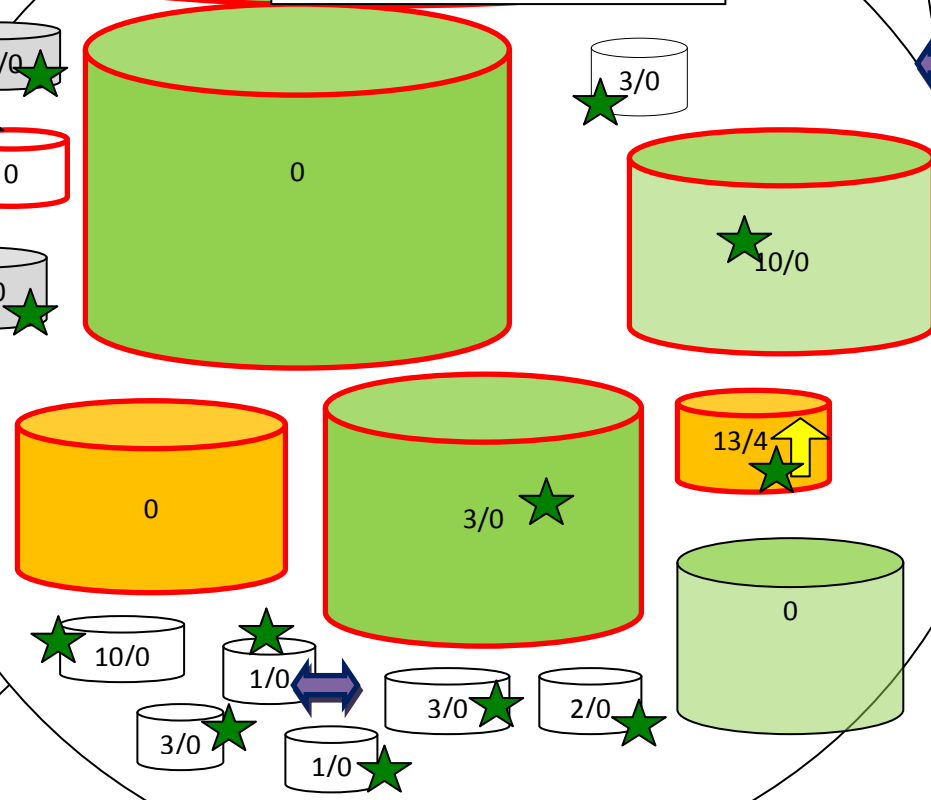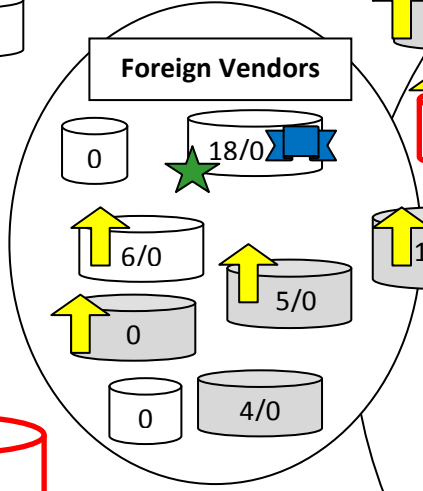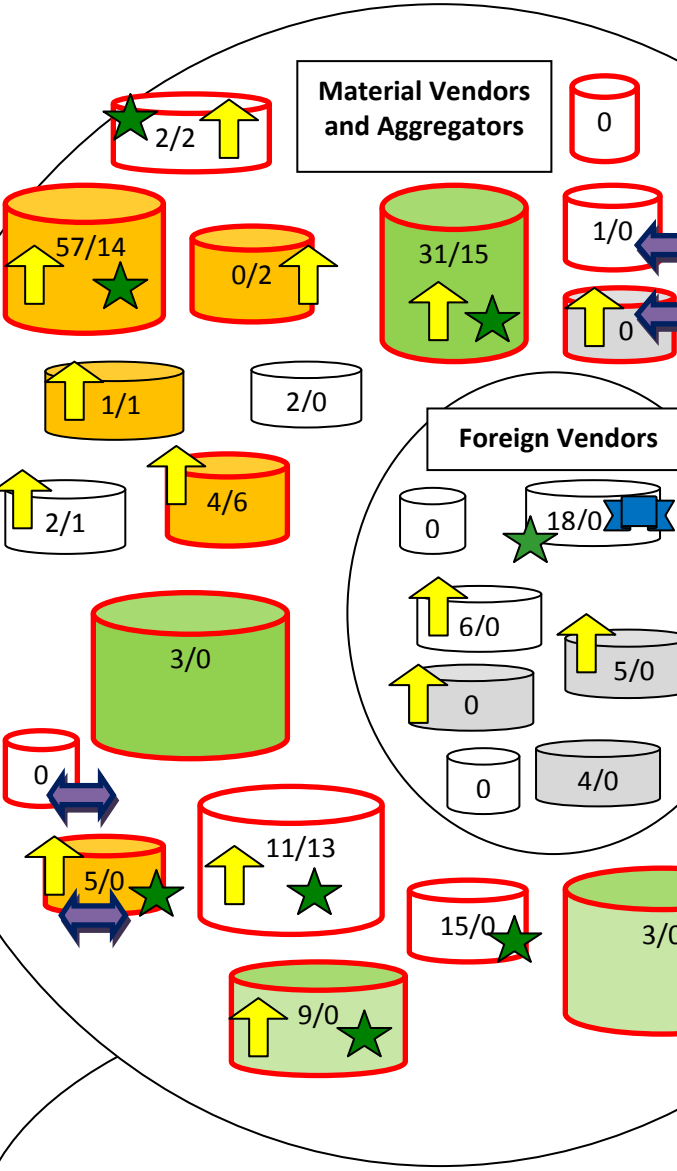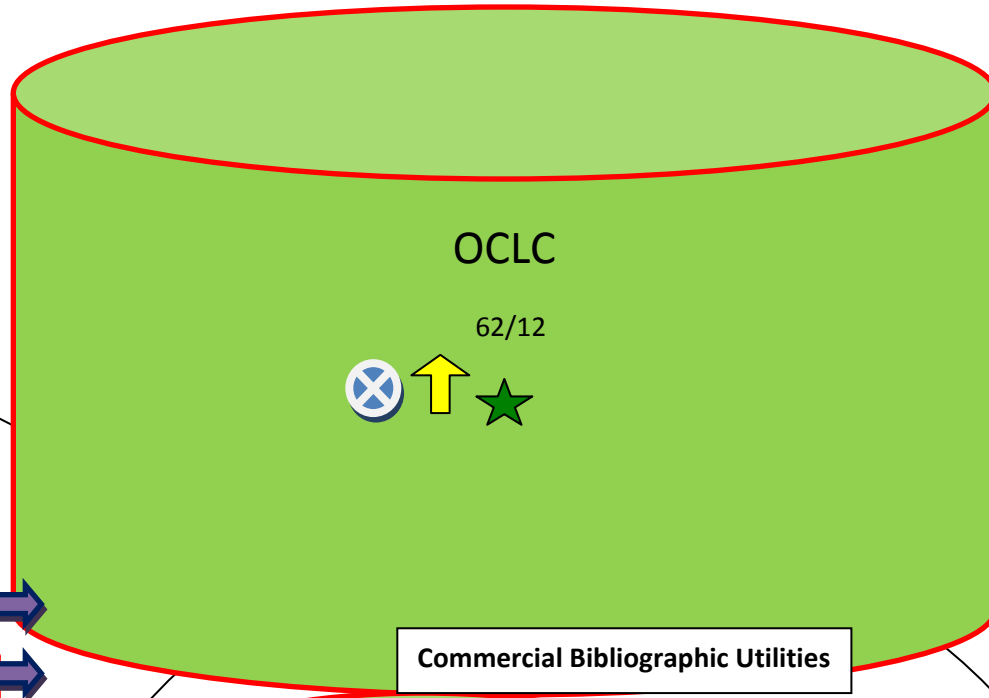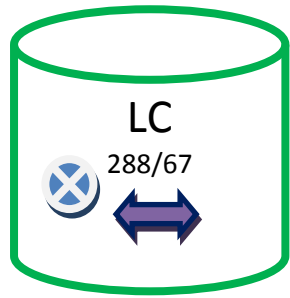Susan Sheldon, Library Consultant

Swets

Swets Canada
SydneyPLUS
Talis Base
Tantor Audio Books
Taped Editions
Taylor & Francis Group,LLC
TDNet Holdings Manager
The AV Cafe
The Book House
The Donahue Group
The Library Corporation
Toronto Public Library
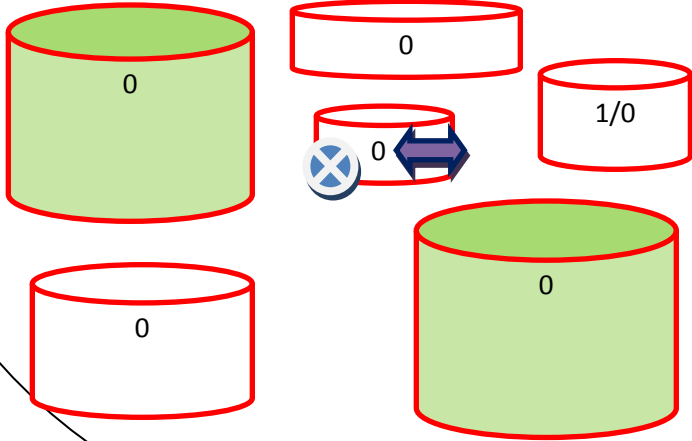Toshokan Ryutsu Center

Touzot Libraire Internat'l
TRC Library Services
UBS Press
United Library Services
Validator TM Subjects and Names Authority Database
Vancouver Public Library
Verso
VNS Group, Inc
VTLS
WebClarity Software Inc.
Wiley Subscription Services
YBP Library Services

LC
288/67

OCLC
62/12

Academic Publishers

Material Vendors and Aggregators

2/2

0

57/14

0/2

31/15

1/0

0

1/1

2/0

24/0

0

4/6

Foreign Vendors

2/1

0

18/0

0

Commercial Bibliographic Utilities

3/0

3/0

6/0

5/0

0

10/0

0

0

10/0

0

0

4/0

13/4

Binderies

0

2/1

11/13

5/0

15/0

3/0

3/0

0

1/0

9/0

0

3/0

10/0

1/0

3/0

2/0

1/0

0

0

0

0

0

4/1

System Vendors

0

0

0

1/0

0

0

0

Non-Commercial Entities, including:

Consortia and other Cooperatives
Centralized Technical Services Units
National Libraries
BOCES Library Catalogs
State Library Agencies
OCLC Regional Networks
Open Database Providers
Etc.

13/10

0

25/25

0

1/0

1/0

0

83/0

20/0

21/0

1

1/0

0
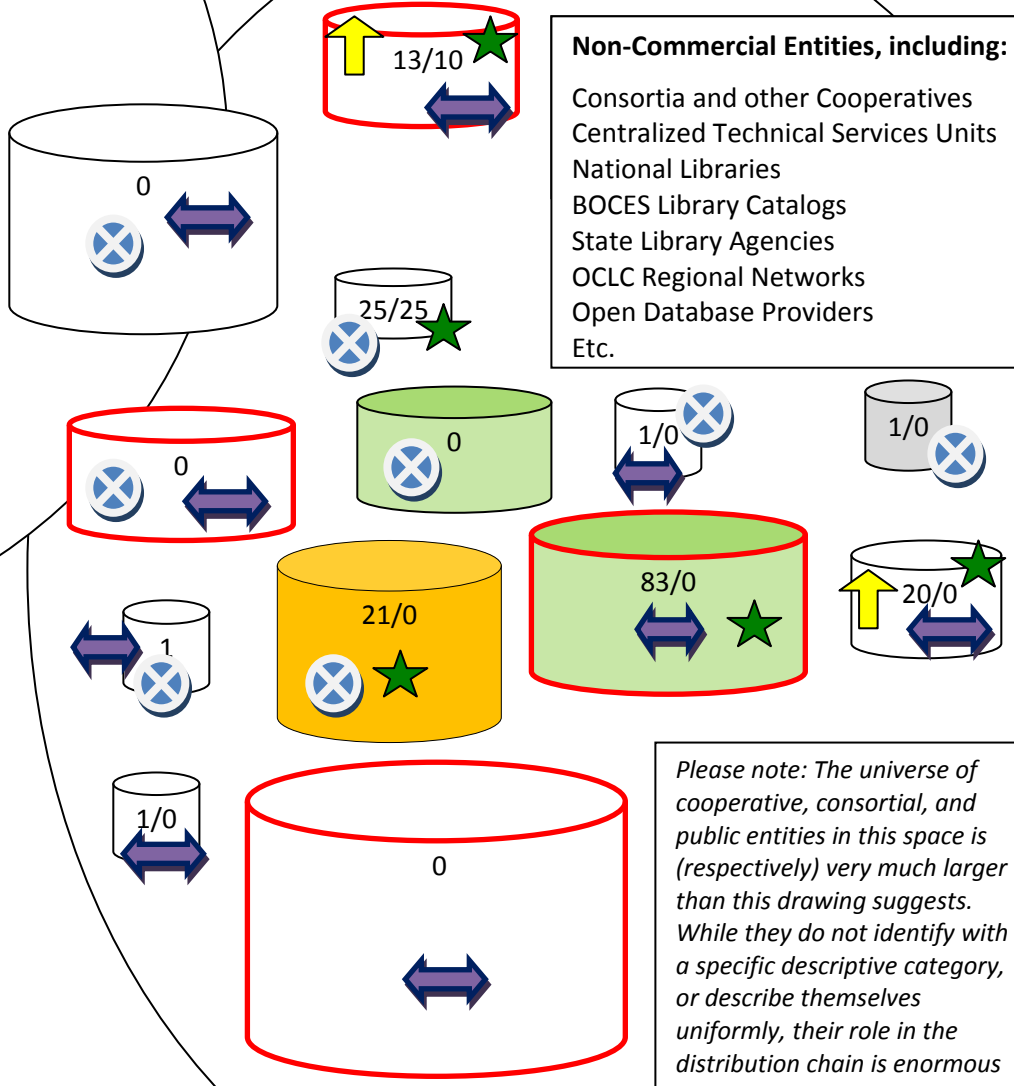
Legend

= purchases records directly from LC

= purchases LC records from a third party

= distributes a significant number of LC records

= acquires no records from external sources

= self-described non-profit entity

= free and open access to all bib records

= WorldCat Cataloging Partner

= creates Full Original MARC Records

= sells records to LC

nn/nn = professional/paraprofessional catalogers

Please note: The universe of cooperative, consortial, and public entities in this space is (respectively) very much larger than this drawing suggests. While they do not identify with a specific descriptive category, or describe themselves uniformly, their role in the distribution chain is enormous and their involvement in the management of authorities cannot be overstated.

MARC Record Distributors and Service Providers 2009