

OCLC: A Review

Jeffrey Beall

Introduction

One of the suggested titles I was given for this chapter was “OCLC sucks.” But I rejected that title, for, despite the temptation to use it, I realized it was unfair, untrue, and perhaps even unwarranted. It’s easier to write a chapter that trashes an organization than it is to write one that presents an objective analysis of it. I’ve resisted the temptation to wholeheartedly trash OCLC, easy as that would have been. Instead, I aspire to the high road: objective analysis, keeping in mind that the word *radical* is in this book’s title. That’s not to say that OCLC is wonderful; it certainly is not. Indeed it is a malevolent organization at times, in the way that all large, rapacious, transnational conglomerates are.

There were two things that led me to reject that title for this chapter. One of them was a series of blog entries by Karen Schneider entitled, “How OPACs Suck.”¹ Her title is specious, of course. OPACs do not suck; they have quietly and efficiently been linking researchers and others with desired information for about twenty years. They are a wonderful tool and represent the best implementation of metadata in the history of mankind. Seeing her use the word “suck” in a title in a supposedly reputable publication (the ALA TechSource Blog) looked really stupid to me. The use of that term in that context seemed wholly inappropriate, and I did not want to be guilty of such poor judgment. Who is the spineless editor that allowed such rubbish to be published?

The other thing that influenced me was an interview of Barbara Tillett that appeared in the blog *Library Juice*.² The interview dealt partly with gadfly Sandy Berman’s harassment of the Library of Congress regarding its choice of terms for the Library of Congress Subject headings. In this interview she said:

Most of our correspondence contains helpful and constructive suggestions—what criticism we receive is simply not as he characterizes it. There is no onslaught of letters and emails and faxes from outraged librarians or researchers. For the most part, public criticism comes from Mr. Berman or other individuals he has urged to write to us. We’re more inclined to react favorably to constructive suggestions than to coercive techniques such as petitions, hostile articles in the library literature, emotional attacks, or letters of complaint to members of Congress. Methods such as these are almost always counterproductive, whereas more cooperative and positive approaches usually produce good results.³

That’s the approach I want to take in this article: the cooperative and positive one. If I have trouble sticking to that approach, it’s understandable, for OCLC claims to be a “member cooperative” when it is really a profit-hungry leech on libraries.

About OCLC

OCLC was born of cataloging, chiefly cataloging in academic libraries. It’s safe to say that the foundation of OCLC rests on the labor of over thirty years of catalogers’ work. OCLC’s fortunes have been built on the backs of catalogers, librarians who are most often underpaid, over-qualified, and under appreciated.

But OCLC has turned its back on catalogers and on MLS librarians as well. This year (2006), I attended the Innovative Users Group meeting held in Denver and heard a talk by Jerry Kline, the president of Innovative Interfaces, Inc. He remarked in his speech that, after the Los Angeles County Public Library, Innovative is the second largest employer of MLS librarians in the state of California. His company values the profession, it values the credential we agree on as the standard for professional librarianship, and it values the people who have earned that credential.

OCLC, Inc. does not share those values. OCLC does not sufficiently value librarianship, the MLS degree, or librarians. This antipathy towards librarians can be measured by observing OCLC employment advertisements. OCLC rarely seeks to hire MLS librarians. One rarely sees a job opening advertised at OCLC that requires an MLS degree. What OCLC values more than anything else is business entrepreneurship. It wants people who are going to come up with the means and strategies for making money for the organization and for destroying the organization's competition. That's what OCLC values.

OCLC fills most of its professional job openings locally. It relies heavily on the Ohio State University in nearby Columbus as a source of workers. The organization is filled with employees who have received either advanced business or computer science degrees from the University of Ohio. Many of the so-called researchers who work in "OCLC Research" (the research department's official name) are computer science students or computer science graduates of the University. Computer science is probably one of the professions that exhibits the most antipathy and condescension towards library science and towards librarians. Computer scientists think they can invent a way to make library catalogs and librarians obsolete (take the so-called Semantic Web, for example), even though despite many years of trying, they have been unable to do so. The Semantic Web, is, according to Wikipedia, "a project to create a universal medium for information exchange by putting documents with computer-processable meaning (semantics) on the World Wide Web".⁴ The Semantic Web is so far ahead of its time, about one thousand years, that it doesn't work yet. Now computer scientists are working on romantic notions like automatic indexing of images, in which a computer recognizes a house in an image and applies the subject heading "house" to the image automatically. It's as if computer scientists even refuse to consider the value of human-created metadata and think that even crude metadata created by a computer is better than rich metadata created by a human.

Many other professionals at OCLC are hired with bachelor's degrees in various fields and then encouraged to earn an MBA, and many do. Note that the organization does not encourage its employees to study library science; it encourages them to study business. One example of this is Erik Jul. Remember him? He was extremely visible for a while at ALA conferences and on popular email lists. He was ALA's toast of Mayfair. According to the LITA web site⁵, he had a master's degree in classics when he was hired at OCLC as a technical writer. He was one of the OCLC employees who were behind the "InterCat" project, when we all started cataloging websites for the first time in WorldCat. Later Jul became the director of the now essentially-defunct OCLC Institute, and disappeared from the library landscape. OCLC is very much an Ohio organization—instead of a national organization—when you look at the origins of its workers. It's rumored that Jul now plays a role in OCLC's WebJunction, which is supposed to be a training and knowledge sharing resource for librarians, but in reality it serves mainly as a means of funneling Gates Foundation money into OCLC.

For its non-professional (staff) employees, OCLC has a questionable and perhaps exploitative recruiting method. These staff members are generally non-permanent, low-level technicians. For example, it hires people referred from employment agencies in Columbus to do its contract cataloging work. This is work that in most libraries is done by professionals, but at OCLC, where bibliographic data quality doesn't count for much, they basically take people off the street and teach them to catalog in a day or so, often in languages they can barely understand. This part of OCLC used to be called TechPro, but I can't find any reference to it on the OCLC website any more. Instead they now call the service Contract Cataloging, which I actually think is a better name than TechPro, which sounded like it was made up by a clueless person in OCLC's marketing department. While they have improved the name, they haven't improved the working conditions there. OCLC Contract Cataloging is basically an information sweatshop, where employees are forced to produce such a high quota of metadata that quality suffers greatly. When employees fail to meet the quota, they are let go, and a call is placed to the employment agency, which sends out another employee to

replace the one that is let go, much the same way that workers are replaced in the maquiladoras along the U.S.-Mexican border.

OCLC's Mission

The mission of OCLC is to separate libraries from their money. There's no other way to say it. Never mind that OCLC is technically classified as a non-profit, presents itself as a library cooperative, and controls the world's largest bibliographic database. There are dozens of MBAs at OCLC whose job it is to think up new ways to generate revenue from libraries.

One of these ways that they've come up with is monopolization. OCLC buys up small, promising companies, companies that promise to have the potential to generate a lot of revenue, especially from libraries. In the past few years OCLC has acquired the assets of the only two competing bibliographic utilities—WLN and RLIN—and is now positioned as the only bibliographic utility in the country. It has no competition. Long ago, under the leadership of robber baron K. Wayne Smith, OCLC began its practice of gobbling up companies that supply libraries with goods or services. For example, OCLC owns Forest Press, which publishes the Dewey Decimal Classification (DDC). Forest Press comes out with a new edition of DDC every few years so that libraries have to shell out capital in order to buy the latest edition and keep up, in much the same way that Mary Baker Eddy, the founder of Christian Science, would periodically update her *Science & Health: With A Key To The Scriptures* and require her followers to purchase the updated editions just to generate revenue.

One of the types of library suppliers that OCLC currently does not have its hand in is library supplies. Companies like Gaylord and DEMCO supply libraries with things like labels, book trucks, archival materials, etc. I predict that OCLC will fill this product gap and soon acquire one of these companies. The MBA entrepreneurs at OCLC are possibly planning this type of acquisition, for it represents a big gap in library services and supplies that OCLC does not currently cover, and judging from its recent acquisitions, OCLC surely seeks to completely control the all of the library services and supplies market.

Another way that OCLC separates libraries from their money is by means of its being a so-called library cooperative. OCLC uses its status as a library cooperative and as an organization that serves libraries to ennoble itself. It puts forth the attitude: "We serve libraries; we are noble, so everything we do is noble." Librarians are sickened by this attitude because they know how OCLC exploits libraries and librarians. OCLC uses bibliographic data created by librarians and then re-sells it at inflated rates to other libraries! Also, OCLC loads very poor quality bibliographic records into its WorldCat database and then charges libraries to access and use this data, charging them the same as it charges for higher quality records, such as those that come from the Library of Congress.

A more honest way of looking at OCLC is this: it buys bibliographic data from libraries at exploitatively low rates, and then sells that data back to libraries at exorbitantly high rates. Also, a lot of the bibliographic data it gets is basically free, such as vendor records (brief, low-quality bibliographic records supplied by book vendors) and records from the Library of Congress, which OCLC gets for free or for a very low cost.

Bibliographic Record Quality

One of the reasons that low-quality bibliographic records remain in WorldCat is that OCLC does not provide sufficient incentive for libraries to upgrade the records. Many member libraries have the ability to enhance vendor records or to at least correct them, but few libraries do this because the monetary incentives for enhancing and correcting records that OCLC offers are insufficient. That is

to say, the credit a library gets for fixing up a poor-quality record in WorldCat is not enough to motivate libraries to do it. The result is that bad data stays in WorldCat for years and years, and possibly forever. Libraries choose instead to download the records directly into their online catalogs and fix up the records there, where the editing process is easier and generally quicker. If OCLC were to sufficiently increase the incentives it offers for enhancing bibliographic records then data quality in WorldCat would be greatly improved.

Regarding the vendor records—some of which are really, really bad—the question has been asked, “Is it better to have a bad record than to have no record at all?” I think this is a valid question, and the answer you get is going to differ depending on whom you ask. If you were to ask a reference librarian, he would probably say that any record is better than none, because in the course of helping someone at a reference desk, it is more helpful to find at least a minimal record than none at all. Finding even a stub record can help confirm a citation and confirms that a book or other information resource does exist.

On the other hand, I have seen many vendor records that are so bad that they have multiple errors, including typographical and transcription errors, that would render the record virtually unfindable in any database. By this, I mean that if the reference librarian mentioned above were to conduct a search for a book for the patron he’s helping, many of the vendor records in OCLC are so bad that he’s not going to find them—not because they aren’t there, but because the title or other fields contains so many errors that it won’t be retrieved in a search. In this case there is no difference between inputting bad records and not inputting a record because the result is the same: no record found!

For several years, OCLC has been in the business of loading large sets of bibliographic records into WorldCat that it obtains from, for example, foreign libraries. These records are minimal level, have name and title headings that do not match any authority scheme, and are often missing much of the data required to meet even minimal-level record standards. There are literally millions of these records that OCLC has incorporated into WorldCat that are sub-minimal level, even according to OCLC’s standards. OCLC has a webpage entitled “WorldCat Principles of Cooperation.” One of the principles enumerated on this page directs libraries to make a commitment to “Create bibliographic records and related data at the fullest possible level, consistent with the standards and guidelines adopted by OCLC”.⁶ So, because OCLC dumps so much dirty data from around the world into WorldCat, this shows that the organization has no desire to follow its own principles of cooperation.

Some libraries have tried to find a way to circumvent the high cost (and low return) of OCLC’s exploitative membership by seeking other sources of bibliographic records. MARC data is everywhere on the web, and much of it is of better quality than what’s found in WorldCat because you can access the poor-quality records that have been downloaded from OCLC and then enhanced, as described above, locally (that is, in their own OPACs only) by OCLC member libraries. It’s also possible to access records that originated in OCLC in the public catalogs of thousands of libraries around the world, a fact that annoys OCLC. Many of the MARC records that OCLC resells to libraries originated from the Library of Congress and are therefore in the public domain, but this doesn’t stop OCLC from copyrighting the data and going after people who use this data that has once resided in the WorldCat database.

Several years ago, OCLC invented a term for using their records second hand; they call it “record nabbing”.⁷ Coining this term was very clever on OCLC’s part because it stigmatized and gave a negative connotation to a practice that, from a different perspective is a creative and innovative way of sharing bibliographic data. When you want to demonize something, you give it a negative name. But OCLC now uses the term “resource sharing” to refer to interlibrary loan, and I really don’t see how the practice of using OCLC records found in online catalogs isn’t really just the same thing:

resource sharing. OCLC has pressured many consortia and libraries to “turn off” their Z39.50 ports to stop the practice of this method of resource sharing, thus cutting off a free source of bibliographic data that was being used by small libraries and libraries unable to afford OCLC’s predatory pricing policies. In summary, when OCLC makes money on the deal, it’s called “resource sharing;” when OCLC doesn’t make money on the deal, it’s called “record nabbing.” A frank and helpful description of the emergence of the concept of record nabbing appeared in the June 23, 2003 issue of *Technical Services Librarian*, written by Michael Maben:

As I was preparing this column in late April, a discussion arose on Autocat concerning usage of OCLC-derived records by nonmember libraries. There was a presentation and discussion about this issue at the February 2003 Members Council meeting. The presentation was done by Gary Houk, OCLC Vice President of Cataloging and Metadata Services. The slides of his presentation are available at:

<http://www.oclc.org/oclc/uc/feb03/ppt/GaryHouk_NonMemberUseofRecords_files/frame.htm>. OCLC's concern is for "record nabbing," which Houk defined as "the unauthorized downloading of OCLC member cataloging records from library OPAC's by all types of libraries from all corners of the world." So the idea is not the unauthorized use of OCLC records from WorldCat, but rather the unauthorized "nabbing" of OCLC-derived records in member libraries OPACs.

One characteristic of virtually any discussion of OCLC on Autocat is one of extreme hostility towards OCLC by some of the participants. This discussion was no exception. People took issue with OCLC's pricing as forcing libraries to engage in record nabbing. Many people questioned who really owns the records, if the records are under copyright protection, contract issues between OCLC and the library, and the like. Others slammed OCLC's non-profit status and how it evidently took a special act of the Ohio Legislature to maintain that status (having seen OCLC's facilities in Dublin, Ohio, I can attest to the fact that they are bringing in a lot of money). One individual even accused OCLC of being a Ponzi scheme⁸.

OCLC even changed the name of its puppet governing board from “Users’ Council” to “Members Council” a few years ago, a public relations move aimed at selling the notion that OCLC is actually a cooperative and not a rapacious conglomerate.

OCLC, as many people know, is headquartered on a “campus” in suburban Columbus, Ohio. It’s the same place that is headquarters to the artery-clogging company Wendy’s Old Fashioned Hamburgers. On OCLC’s campus, there are several buildings named after past OCLC presidents. There’s the Frederick G. Kilgour Building, the K. Wayne Smith building, and the Rowland C.W. Brown Building. This last one is named after a now forgotten former president of OCLC. Note that the K. Wayne Smith building was named after K. Wayne Smith *while he was still president of OCLC*. What incredible hubris! I am certain that they will construct a new building soon and it will be named—you guessed it—the Jay Jordan building.

Connexion

In March, 2004, OCLC released version 1.1 of Connexion, its new cataloging application for cataloging in WorldCat. The interface is available in a browser version and a client version. I am most familiar with the client version: i.e., a software program that is loaded (continually) on my computer at work. (I say “continually” because it seems like a new version of the software is released every few months or so.) Connexion replaces Passport, the now-obsolete OCLC cataloging application.

The only way to describe OCLC’s implementation of Connexion is this: *complete total laughingstock disaster*. It was probably one of the worst implementations ever of any new software. Everything that could go wrong has gone wrong. I use the perfect tense because the implementation, now two and a

half years later, is still going on. The “Known Problems” web page regularly grows after each release of a new version. The OCLC programmers are the Three Stooges of the coding world. They surely must be yet more clueless computer science types from Ohio State University who have never darkened the doors of a library, much less dealt with metadata of any kind.

One of the worst parts of the implementation of Connexion has been the downtime; it often crashes. You’ll be cataloging away—or, should I say “creating metadata”—and all of a sudden you get a strange, vague error message in a little dialog box that says something like “Unable to access remote server.” This usually means the system has gone down, and it often goes down, especially on Mondays. I think it goes down a lot on Mondays because the beginning coders OCLC hired from the Columbus Coders Discount Employment Agency spend the weekend trying to fix the numerous bugs in the system. They do fix the bugs, but in doing so, they create new ones, bugs that cause the system to crash when heavy usage starts again on Monday.

OCLC is always “very sorry” for the inconveniences these outages cause. These outages mean that literally thousands of library workers cannot do their jobs while the system is down. OCLC has not hired sufficiently skilled programmers to create Connexion. Clearly, they’ve taken the cheap route and have hired people who are not excellent but only satisfactory. To put it another way, the OCLC programmers are about as good as the data in a vendor record from Europe, loaded into WorldCat, with typos and missing and incorrect data.

OCLC, in its Connexion email list, encourages users to either contact their regional service provider or to contact customer service when the system goes down. There are several problems with this. My experience is that when I contact my regional service provider and tell them (by means of an email) that I cannot connect to OCLC via Connection, they send me an email the next day saying that the problem has been fixed. In other words, contacting one’s regional service provider is a waste of time. If you try to call OCLC, one of their other suggested options when the system goes down, you are put on hold and forced to listen to music that only people in Columbus, Ohio could ever tolerate. After about ten minutes of this insipid music, a customer service representative finally answers. Then you explain the problem and hear silence on the other end. It’s clear the person on the other end of the phone has no clue what you are talking about. You get the feeling that you are talking with someone who took her GED instead of finishing high school and took any job she could find. Eventually, however, usually sometime after lunch Mountain Time, the system comes back up and then you are able to begin cataloging again.

The implementation of OCLC Connexion has been an embarrassing disaster from the beginning. It has been a case study of how not to implement a new service. Now that OCLC is a monopoly, the organization will be less mindful of its users’ needs. This has already become evident in email postings by OCLC employees to the OCLC Cataloging email list. For example, David Whitehair, the Connexion product manager, now maintains a very dictatorial attitude on the list, refusing, in one instance, to extend the “end-of-service” date of one of the versions of Connexion client, despite the please of libraries across the country who were having major troubles implementing the next version.

OCLC Propaganda

Am I the only one who cannot understand a word that Lorcan Dempsey says or writes? Have you ever tried reading one of his articles? They are some of the most incoherent and desultory articles in the history of information. Nevertheless, having Depmsey, OCLC Vice President and Chief Strategist, is another clever move on OCLC’s part. His flamboyance makes him a popular attraction at ALA conferences. Also, at these conferences, there is a tendency among the attendees to believe anything that is said with a British accent (unless it’s Michael Gorman, who because he is pro-cataloging, is sometimes stigmatized). Dempsey appeared out of nowhere, purports to “know a lot

about emerging technologies,” and, of course, has a blog. He lives for new technology. It seems as if he thinks that any new library or information technology is automatically better than the technology that predates it and it must be implemented immediately, especially if OCLC has a hand in it. Dempsey’s mission is to make OCLC the center of libraries’ technology strategies. The striking thing about Lorcan Dempsey is that he is all form and very little content. You can listen to him ramble for an hour and you leave the conference hall with less understanding of information technology than when you entered.

OCLC also has a large propaganda machine. This machine consists of several very articulate individuals who travel the country attending library conferences large and small and promoting OCLC’s mission, which, as you will recall, is to separate libraries from their money. These individuals include people like Alane Wilson, a very eloquent and winsome speaker who travels the country and, like many others at OCLC, attempts to predict the future of the library and information landscape. In their predictions, they always state that there will be certain needs, needs that OCLC just happens to be working on and will be able to fulfill.

OCLC’s Research Department always conducts research that supports the organization’s goals and views of the information world. Most of their research is carried out in Ohio, as if Ohio were a valid statistical information sample for the whole world. An example of this is a research project carried out by Lynn Silipigni Connaway. You remember her, right? She used to be in charge of netLibrary, another start-up that OCLC gobbled up. She was responsible for distributing netLibrary bibliographic records to thousands of libraries, records that contained many serious errors, such as typos, invalid headings, etc. She proudly distributed this dirty data, and OCLC was only happy to make money from the deal. Anyway, her recent research was entitled, “Sense-Making the Information Confluence: The Whys and Hows of College and University User Satisficing of Information Needs.” What a pedantic title. Their use of the non-word “satisficing” shows how they always attempt to be on the cutting edge, in this case by using a word that sounds technologically trendy. Guess where the research was carried out: Ohio. How convenient. As Ohio goes, so goes the nation, I guess. The main goal of the research was to show that college freshmen don’t use library catalogs anymore. For this the federal grant-making agency Institute of Museum and Library Services awarded thousands of dollars? OCLC wants its research to point to the fact that information seeking behavior is changing, and only OCLC can help you provide the means of supplying information in this new information landscape.

Conclusion

I think it will be fascinating to observe the future of OCLC. I think more and more libraries will begin to conclude that they are not getting their money’s worth from the organization, what with the poor data quality and poor computer applications it provides, and with the way it exploits its employees and snubs the library profession. OCLC at some point will be exposed for the profit-seeking multinational corporation it is, rather than the non-profit, cultural heritage organization it presents itself as. When this happens, libraries may be freed from their connections to OCLC and be forced to innovate and will create better, home-grown solutions to library needs than a large, distant, and disconnected organization could ever do. Libraries will foster true resource sharing, sharing that doesn’t involve them sending their scarce capital to Ohio. In this way, libraries will be able to spend more of their budgets on buying resources for their users.

References

1. Karen Schneider, “How OPACs Suck,” *ALA Tech Source Blog* (2006), <http://www.techsource.ala.org/blog/2006/03/how-opacs-suck-part-1-relevance-rank-or-the-lack-of-it.html>.
2. Barbara Tillett, interviewed in *Library Juice*, <http://libraryjuicepress.com/blog/?p=115>.

3. Barbara Tillet interviewed in *Library Juice*.
4. Wikipedia, s.v. "Semantic Web", http://en.wikipedia.org/wiki/Semantic_Web
5. Library & Information Technology Association, "Erik Jul," (2006)
<http://www.ala.org/ala/lita/litaresources/toptechtrends/erikjul.htm>
6. OCLC, "WorldCat Principles of Cooperation," (2006)
<http://www.oclc.org/worldcat/contribute/principles/>
7. Gary R. Houk, "OCLC speaks out on record nabbing," *Library Collections, Acquisitions, & Technical Services* 27, issue 3 (2003):277-279.
8. Michael Maben, "OBS OCLC Committee," *Technical Services Law Librarian* 28.4, 2003. Online:
<http://www.aallnet.org/sis/tssis/tsll/28-04/oclc.htm>