

# Where Will the Catalog Go?

By

David W. Lewis

Dean of the University Library

Indiana University Purdue University Indianapolis

Lewis, David W. "Where Will the Catalog Go?" *Library Computing* 18(4):263-268  
1999. (Actually published in the spring of 2001.)

## **Abstract:**

For over a century the catalog has been the core component of most libraries' service strategies. As we move from a world based on print documents to one where most information is digital and networked, libraries must reconstruct their service strategies and the role the catalog plays in them. In doing so it is important to understand what types of navigational tools will best serve the users. We need to create these tools rather than attempting to fit digital and networked documents into the print-based forms of bibliographic control.

This is an interesting time to reflect on the future of the library catalog, because, despite revolutionary changes in technology, libraries are still just tinkering with century old bibliographic practice. We are trying to stretch old structures and practices to fit the new world we live in. This strategy is not adequate to the task and is doomed to fail.

To understand where we will be required to go it is important to understand what our aspirations for the catalog have been, and where these aspirations fell short. We must also understand how the catalog fits into the rest of what libraries do, and what they will do. We need to understand how a library "collection" will be structured, and how users will view and approach it. With this background we can begin to imagine what forms of bibliographic control will, or will not, make sense in the world where most documents will be networked and digital.

### **Back to Cutter**

As Cutter defined it in his *Rules for a Dictionary Catalog*, the object of the library catalog was:

1. To enable a person to find a book of which is either (a) author, (b) title, or (c) subject is known.
2. To show what the library has (d) by a given author, (e) on a given subject, or (f) in a general kind of literature.
3. To assist in the choice of a book (g) as to its edition (bibliographically) or (h) as to its character (literary or topical). (Cutter, 1891, p. 8.)

As it turned out, Cutter was too ambitious. His first objective quickly passed to the realm of national bibliography and to reference publishers who created bibliographies of all sorts. The actual practice has been for the catalog to describe and index the holdings of a particular library. A century after Cutter, the 1985 edition of the *Introduction to Cataloging and Classification* put it this way, "In order to provide access to the holdings of a library, an index or list of the materials in the collection must be maintained.... Its [the catalog's] prime purpose is to record, describe, and index the holdings of a specific collection." (Wynar, 1985, p. 3-4.) This objective, as it turned out, was also too ambitious. In the cataloging

textbook's long list of items that might be contained in a collection, including coins and stamps, one item is conspicuously absent — journal articles. Journal articles, the most important means of scholarly communication in many fields and the source of choice for students and the general public in many instances, are left out of most library catalogs. For well over a century bibliographic control for journal articles been managed through indexes and abstracts produced by commercial publishers and scholarly societies. The catalog has only managed to keep track of the holdings of journals themselves.

In the world of print on paper documents and library collections based on them, this limited practice made sense and was an appropriate adaptation to economic and technological constraints. For well over a hundred years the library catalog has been a locally produced index for a local physical collection. It has focused on monographs (mostly books) and serials at the title, not the article, level. Efficiencies have been introduced as the catalog moved from handwritten to typed to computer produced cards and then to computer databases. These same technological developments allowed to the increasingly effective sharing of records beginning with the Library of Congress card distribution service and ending with the international databases of OCLC and RLG. But the fundamentals have not changed — a catalog describes and indexes a unique local physical library collection, and its construction and maintenance is primarily the result of work done by and in that local library. Despite all the efforts to make the process efficient, it remains expensive.

### **Why a Catalog?**

"To provide access," seems, at least to librarians, an obviously valuable purpose, but we must take a deeper look. Why do we provide access? Why do we have collections? Libraries have large collections of information because humankind has accumulated a very large amount of knowledge and if this knowledge is not kept in organized collections, it is useless. Local library collections exist because they have been the best way to make knowledge available to the organizations and communities in which they exist. Library catalogs exist to make these local library collections useable. It is critical to understand that libraries and catalogs are not the goal. Organized knowledge is the goal. Library collections and library catalogs have been an effective way of reaching this goal, but they are not the only way, and they may not, given the changes in technology, be the best way in the future.

But even in the print world catalogs were not the whole story. As noted above, economic and technological constraints have limited what the catalog can accomplish. The catalog has been central for libraries because the library's

investment in its local collection, particularly its book collection, has been its most expensive and useful asset. But this is not all that libraries do. They collect bibliographies and periodical indexes because the catalog could not fulfill Cutter's first objective, and because the local cataloging of journal articles is inefficient and unaffordable. Libraries also collect reference works because they are the easiest ways to find facts. Finally, libraries hire people to help the people who use libraries because the tools we provide are complex and, despite our best efforts, difficult to use.

Navigating large collections of information requires assistance, and even small libraries contain large amounts of information. In the first half of the 20<sup>th</sup> century libraries developed the service strategies that are still generally in use today. The combination of the catalog, a reference collection, and reference librarians made libraries manageable in the late print era. The strategy has been effective even when collections exceeded millions of printed volumes, but it assumes a central collection and users who come to that collection. This, of course, is no longer the way things are.

### **What Will the Collection Be? – The Library Perspective**

While no one expects books to completely disappear anytime soon, it is clear that locally held physical collections of printed materials are becoming less important. The resources available to library users are increasingly not physical, but digital and network based. Indexes and reference works that were first put into electronic form as CD-ROMs have migrated to the Web. Scholarly journals and the popular press are increasingly available as web products. Sometimes the titles maintain their bibliographic integrity. Sometimes they are pulled apart and individual articles are amalgamated into collections that combine indexing and full text. Electronic books, while still in the experimental stage, can be expected to replace, at least in some circumstances, their printed equivalents in the very near future.

And then there is all the "stuff" out there on the Web freely available to everyone, unstructured, often unevaluated, with haphazard access provided by a variety of search engines. Much of this stuff is of little value, but much is more current and richer than the printed material libraries have traditionally collected. Importantly, network based information is broadly and easily available. Too often librarians forget how difficult print collections are to use and how much time users must invest in traveling to and using them.

As we think about what the catalog will become, we must first think about what the collection will become, because the catalog will have to adapt to this reality. A library collection in the future will have several components.

1. Tangible physical items that are purchased, held, and loaned will remain for some time. This part of the collection will become smaller and less important for most libraries.
2. Electronic content that is purchased or licensed, will become an increasingly large part of most libraries' collections. In some cases this content will be purchased as familiar bibliographic units — electronic journals or e-books. But the larger and more important part of this collection will be amalgamations. These amalgamations will not be static entities, but rather will be in constant flux as vendors add and subtract content based on their judgments about what is important, their contracts with publishers, and the economics of providing specific materials. It will also be increasingly the case that these purchases will be made by consortiums, not by individual libraries. So individual libraries will have even less control.
3. Some electronic content will not be licensed or purchased in anticipation of use, but rather will be acquired only when an actual use materializes. This can be done with pay-for-view systems or user-driven purchasing. In either case, the "collection" includes all of the material, which is potentially available, even though it has yet to be acquired, and may in fact never be acquired.
4. Finally, much electronic content will be free. Like the on-demand materials, these items are a part of the library's "collection" only to the extent that the library assists users in finding them. Much of this material, the web "stuff" will be in constant flux, though the best will likely become more stable with time.

All of the digital material will be available outside of library buildings in homes, offices, classrooms, coffee shops, and everywhere else. This means the other services and collections that with the catalog made up the library's service strategy will not be there.

### **The User Perspective — Asking the Right Question**

From the user's perspective, the "collection" is everything that is available. Users, as members of communities and organizations will have access to printed collections and rights to purchased or licensed digital material that the library provides as an information subsidy for those affiliated with that organization or community. (Lewis, 1998) But they will also have access to the free material on the Internet. Users do not care that some information was paid for and some was

not. All they care about is getting answers to questions and finding information to help them solve the problems they face.

Cutter's proposed objectives for a catalog, with the appropriate updating, are still a good guide to what users want in a tool to help them navigate the information sea. They will want a tool that will:

1. Provide them with information about known items (items by a particular author or title) and items on a given subject.
2. Show them where this item is located (retrieve the full-text or complete item).
3. Help in the selection of the best items or the best copy of a particular item.

The question then becomes not how do libraries provide access to the collection as the library sees it, but rather, how do we create or acquire the tools that meet the user's needs.

### **Outdated Assumptions**

It is difficult at this juncture to predict with any certainty what the future of the catalog will be, but it is clear to me that there are several assumptions that are at the core of current bibliographic practice that need to be discarded if we are to create the required tools.

1. *The catalog is a handcrafted product.* The current practice is, almost without exception, for people to add items to catalogs one at a time. Even when this is done by clerical staff the cost will be more than we can afford. Whatever our catalogs become, we will want to have data created for it and loaded into it automatically by a variety of vendors. What Marcive does for U.S. government documents we will want done for us by all of our content providers. We will want our serials and book vendors to load data directly to the catalog. OCLC's PromptCat service, or services like it, will need to become the norm.

Hand crafting subject web pages, has been seen as a way to supplement the catalog and avoid the difficulties of forcing web "stuff" into existing bibliographic structures. This is an expensive and inefficient stopgap measure. This function, like the creation of bibliographies in the print era, will continue, but will have to be taken over by cooperative projects, like OCLC's CORC or by commercial organizations.

2. *The user needs to learn to master the bibliographic complexities of the catalog.* For a long time librarians have argued that the complexities of our bibliographic

practice are required because of the complexity of our collections. This might be true, but users have never understood what we were up to. (Lewis, 1987) We could get away with this arrogant approach because reference librarians held the hands of many users. In a web environment, where users no longer come to the library where they can easily get assistance, this approach cannot continue. Stop for a moment and imagine what the catalog might be like. It could have a spell checker. It could provide results ranked by importance. It might have reviews or comments on items from other users. It might know what I searched the last time I used it, and tell me about new items that might interest me. It could tell me what titles are most used by people like me, and what other books were used by people who used a book I find interesting. It might include cover art or information about the author. If it did all this and offered me an opportunity to purchase the book and have it delivered to my home, it would, of course, be Amazon.com. There is a reasonable debate about whether or not Amazon.com is the appropriate model for a library. (Coffman, 1999; Griffiths, 2000) But there should be no argument that the Amazon.com search engine can deliver more value to the user than the typical OPAC. I want a catalog that does all of the things Amazon.com does, and more. Our users will quickly come to expect systems with this level of power and flexibility. We will have to provide them.

3. *Authority control, Boolean searching, MARC, and AACR2 will live forever.* It is particularly important that we end our dependence on systems based solely on authority control and Boolean searching. The limits of this approach have been clear for some time to researchers, but library practitioners have largely ignored this work. (Frants, et. al., 1999) MARC has proved to be a remarkably robust data format, but we will need to find ways to incorporate other datastructures into the catalogs we build. AACR2, while it represents a century of cataloging practice, is based on print documents, and its governance structures make it unresponsive and slow to adapt as the nature of documents change. In a world where the full document is a couple of clicks away from the bibliographic citation or catalog record, insistence on complete descriptive cataloging seems silly.
4. *Good enough is not good enough.* This is not, as it has been in the past, a question of a simple cost/benefit trade-off. As Marlene Manoff puts it, "No matter how much libraries refine and improve procedures for cataloging Web resources, the mutability of those resources will prevent the achievement of anything comparable to the level of bibliographic control possible for printed objects." (Manoff, 2000, p. 865.) We will have to get used to the fact that the certainty that was the core of bibliographic practice in the print world is no longer possible. This is not a matter of cost; rather it is just the way the world

now is. We will have to learn to accept this uncertainty and develop bibliographic practice based on it.

5. *The catalog is only meant to be used by people.* It will be increasingly important to be able to link together networked resources so that one can go from a reference or citation in one source to the complete item or document in another source. This is a difficult problem, especially where purchased content is involved. Some mechanism will be required to point at copies of content items that individuals in a particular domain have rights to. The most promising approach to this problem is SFX (Van de Sompel and Hochstenbach, 1999, parts 1-3) which uses the OpenURL standard to make this sort of linking possible (Van de Sompel, et. al., 2000). An SFX server is really a catalog for a particular domain which uses metadata, as defined by the OpenURL standard, to link references and citations to the complete item or document. It is thus a library catalog that different information services can use behind the scenes, without direct human intervention in the transaction. I believe we will want to rely on various search mechanisms, whether they are established indexes and abstracts or new web-based services like Yahoo, AltaVista, or Excite, as the way users identify items on a particular subject. The "catalog" then becomes the automatic linking mechanism between these systems.
6. *The public catalog is the core component of a library's service strategy.* As Norman D. Stevens puts it, "Perhaps it is time to recognize that the library catalog, as we once knew it, may be on its way out... What is needed in lieu of the catalog, however it might be augmented and expanded, is the development of a structured library-centered system of access to the wider world of electronic information." (Stevens, 1998, p. 190.) Stevens suggests a "user centered" system in which individuals maintain their own "catalog" of the items they care about. This database would be build by searching a variety of library and other systems, including the Web. Much of this can be done automatically in the background as software agents harvest data from known places and follow leads to unexplored sites. Such a system would understand the overlapping rights that the individual has as a member of multiple communities and organizations. It would then determine where the best copy of a document, based on these rights, can be found. The system would learn and adapt as the information environment or the individual's interests changed. Libraries would provide databases that would be used by such systems, but the single catalog would no longer be the core of the library's service strategy.



## **Conclusion**

We are at the beginning of a revolutionary period in the way bibliographic control is conceived and delivered. As is often the case, I expect that change will come from the periphery. The guardians of cataloging rules will not lead the way, nor will the leading institutions of the past. Users, who will create for themselves the tools they need to navigate and effectively use the information environments they inhabit, will drive the coming change.

Libraries and librarians who support their users in this effort will remain significant in their communities and organizations. Those libraries that insist on maintaining centralized structures and control will risk becoming irrelevant. We have much to contribute in this effort. But it is our broad understanding, not the details of our current practice, which matter. We will need to use this understanding to create not simply a new form of the catalog, but rather a whole new way of navigating the world of networked information.

**Author Information:**

Communications to the author should be addressed to David W. Lewis,  
University Library, Indiana University Purdue University Indianapolis, 755 West  
Michigan Street, Indianapolis, IN 46202-5195, voice: 317-274-0493, e-mail:  
[dlewis@iupui.edu](mailto:dlewis@iupui.edu)

David W. Lewis is the Dean of the University Library at Indiana University  
Purdue University Indianapolis. He has a 25 year career working in a variety of  
public service and administrative positions in college and university libraries.  
He has published the service, organizational, and economic issues confronting  
libraries.

---

**Cited materials:**

Coffman, Steve. (1999, March). Building the Earth's Largest Library: Driving into the Future. *Searcher* 7(3).

<http://www.infotoday.com/searcher/mar99/coffman.htm> (reviewed October 2, 2000).

Cutter, Charles A. (1891). *Rules for a Dictionary Catalog*, 3<sup>rd</sup> ed., U.S. Bureau of Education, Special Report on Public Libraries — Part II. Washington, D.C.: Government Printing Office.

Frants, Valery I.; Jacob Shapiro; Isak Taksa, and Vladimir G. Voiskunskii. (1999, January). Boolean Search: Current State and Perspectives. *Journal of the American Society for Information Science* 50(1):86-95.

Griffiths, Jose-Marie (2000, August). Deconstructing Earth's Largest Library. *Library Journal* 125(13):44-47.

Lewis, David W. (1987, July). Research on the Use of Online Catalogs and Its Implications for Library Practice. *Journal of Academic Librarianship* 13(3):152-157.

---

Lewis, David W. (1998, December). What If Libraries Are Artifact-Bound Institutions? *Information Technology and Libraries* 17(4):191-197.

Manoff, Marlene. (2000, Summer) Hybridity, Mutability, Multiplicity: Theorizing Electronic Library Collections. *Library Trends* 49(1):857-876.

Stevens, Norman D. (1998, December) Looking Back at Looking Ahead, or "The Catalogs of the Future Revisited" with Additional Speculation. *Information Technology and Libraries* December 1998 17(4):188-190.

Van de Sompel, Herbert and Patrick Hochstenbach (1999, April). Reference Linking in a Hybrid Library Environment: Part 1: Frameworks for Linking. *D-Lib Magazine* (5)4.  
[http://www.dlib.org/dlib/april99/van\\_de\\_sompel/04van\\_de\\_sompel-pt1.html](http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt1.html) (reviewed September 29, 2000).

Van de Sompel, Herbert and Patrick Hochstenbach. (1999, April). Reference Linking in a Hybrid Library Environment: Part 2: SFX, a Generic Linking Solution. *D-Lib Magazine* (5)4.  
[http://www.dlib.org/dlib/april99/van\\_de\\_sompel/04van\\_de\\_sompel-pt2.html](http://www.dlib.org/dlib/april99/van_de_sompel/04van_de_sompel-pt2.html) (reviewed September 29, 2000).

---

Van de Sompel, Herbert and Patrick Hochstenbach. (1999, October). Reference Linking in a Hybrid Library Environment: Part 3: Generalizing the SFX solution in the "SFX@Ghent & SFX@LANL" experiment. *D-Lib Magazine* (5)10.

[http://www.dlib.org/dlib/october99/van\\_de\\_sompel/10van\\_de\\_sompel.html](http://www.dlib.org/dlib/october99/van_de_sompel/10van_de_sompel.html) (reviewed September 29, 2000).

Van de Sompel, Herbert; Patrick Hochstenbach, and Oren Beit-Arie. (2000). "OpenURL Syntax Description," version: OpenURL/1.0f - 2000-05-16. <http://sfx1.exlibris-usa.com/OpenURL/openurl.html> (reviewed September 29, 2000).

Wynar, Bohdan S. (1985). *Introduction to Cataloging and Classification*, Seventh Edition, Edited by Arlene G. Taylor. Littleton, Colorado: Libraries Unlimited.

---