Online Catalogs: What Users and Librarians Want

An OCLC Report

Synopsis





Online Catalogs: What Users and Librarians Want

An OCLC Report

Synopsis

Principal contributors

Karen Calhoun, Vice President, WorldCat and Metadata Services

Joanne Cantrell, Marketing Analyst

Peggy Gallagher, Market Analysis Manager

Janet Hawk, Director, Market Analysis and Sales Programs

Graphics, layout and editing

Brad Gauder, Creative Services Writer **Rick Limes**, Art Director **Sam Smith**, Art Director

Contributor

Diane Cellentani, Market Research Consultant to OCLC

Copyright © 2009, OCLC Online Computer Library Center, Inc. 6565 Kilgour Place Dublin, Ohio 43017-3395

ALL RIGHTS RESERVED. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, mechanical, photocopying or otherwise, without prior written permission of the copyright holder.

The following are trademarks and/or service marks of OCLC: Connexion, FirstSearch, OCLC, the OCLC logo, WorldCat, WorldCat.org, WorldCat Resource Sharing and "The World's Libraries. Connected."

Third-party product, service, business and other proprietary names are trademarks and/or service marks of their respective owners.

Printed in the United States of America

Cataloged in WorldCat on August 4, 2009 OCLC Control Number: 429057067

ISBN-10: 1-55653-415-9 ISBN: 978-1-55653-415-7

Abstract

In 2008, OCLC conducted focus groups, administered a pop-up survey on WorldCat.org— OCLC's freely available end-user interface on the Web—and conducted a Web-based survey of librarians worldwide.

The Online Catalogs report presents findings from these research efforts in order to understand:

- The metadata elements that are most important to end users in determining if an item will meet his or her needs
- The enhancements end users would like to see made in online library catalogs to assist them in consistently identifying appropriate materials
- The enhancements librarians would recommend for online library catalogs to better assist them in their work.

The findings indicate, among other things, that although library catalogs are often thought of as discovery tools, the catalog's delivery-related information is just as important to end users.

Introduction

"A persistent shortcoming in the decision-making process [about library database quality] that needs to be addressed is the lack of serious research into user needs and benefits, and the actual impact on users of database quality decisions."

What constitutes "quality" in catalog data has been reasonably well-understood by library professionals. The informed librarian's definition of catalog quality can be traced directly to Charles Cutter's statement in 1876 of the objectives of a library catalog; these objectives have guided librarians preferences for catalog design for over a hundred years. Thanks to Cutter and the theorists who followed,³ today's library catalogs are founded on predictable and consistent record and heading structures, which facilitate serendipitous discovery, efficient known-item retrieval and many ways to browse.

A study conducted by the OCLC Online Data Quality Control Section in 1989 confirmed that librarians' consensus about quality in their own library catalogs carried over to their expectations of WorldCat as a source of shared catalog records.4 Carol Davis, then head of the Online Data Quality Control Section, found that librarians' top three quality concerns with WorldCat at the time were duplicate records (more than one record describing the same edition), incorrect (unauthorized) forms of name headings and incorrect (unauthorized) forms of subject headings. Today, OCLC's WorldCat quality program continues to center on managing these top three database quality priorities.

Many writers have affirmed that Cutter designed his objectives of the catalog with the convenience of the user in mind. A similar motivation (the convenience of the user) underlies the "Functional Requirements for Bibliographic Records" (FRBR), a conceptual model based on the user tasks of finding, identifying, selecting and obtaining wanted information. However, an examination of the literature turns up little evidence that Cutter, the distinguished theorists who followed him, or those who framed FRBR and "Resource Description and Access" (RDA)6 rigorously tested their conceptual frameworks with information users.

Information-seeking Behavior

In 2003, the OCLC Environmental Scan identified self-service, satisfaction and seamlessness as definitive of information seekers' expectations. That report documented ease of use, convenience and availability as equally important to information seekers as information quality and trustworthiness. In 2005, the report Perceptions of Libraries and Information Resources looked further into people's information-seeking behaviors and preferences with respect to libraries, most notably revealing the trend of information seekers to begin a search for information with a search engine (84%) rather than on a library Web site (1%).8 Social science researchers have investigated many paradigms in information-seeking research?; the "Principle of Least Effort," attributed to philologist George Zipf, is probably the best known in libraries.

The recent library literature contains numerous articles on the need for change in online catalogs to better satisfy the expectations of information seekers who are accustomed to Web search engines, online bookstores and seamless linking to full text. Increasingly it is understood that an end user's expectations and work practices on the Web matter a good deal to whether he or she will use or revisit a library online catalog. In his August 2005 paper for the International Federation of Library Associations and Institutions (IFLA), John Byrum, Library of Congress, wrote of the need for library catalogs to provide access to more content and to offer significantly enhanced functionality based on the features of popular search engines. 10

At the same time, many of the users of library online catalogs are librarians and staff—these individuals form an important catalog user community themselves. Therefore, just as catalog end users (e.g., citizens, students and faculty) have information needs, preferences and expectations that need to be supported by catalog data, so do librarians who get their work done using the data underpinning the catalog.

Designing a New Database Quality Program

As next-generation library online catalogs emerge, and as OCLC makes decisions about how to shape its next-generation WorldCat data quality program, it is essential to gain an evidence-based, user-centered understanding of what catalog data "quality" is to the various communities (both end users and librarians) for whom library online catalogs are or can be an important, frequently visited information resource.

To this end, OCLC formed a research team to:

- Identify and compare the data quality expectations of catalog end users and librarians
- Compare the catalog data quality expectations of types of librarians
- Recommend catalog data quality priorities, taking into account the perspectives of both end users and librarians.

While many of the research team's findings relate specifically to defining a WorldCat data quality program, many findings are generalizable to the data supporting the current generation of library online catalogs and integrated library systems (ILSs). This report describes those generalizable findings. Readers who are seeking to define requirements for improved catalog data (exposed in both end-user and staff interfaces) may find this report helpful as a source of ideas.

Methodology

The OCLC research team employed three methods to identify the catalog data quality expectations of catalog end users and library staff: focus groups, a pop-up survey on WorldCat.org and a Web-based survey. WorldCat.org, OCLC's freely available end-user interface on the Web, provided the means for the team to study focus group and pop-up survey participants' reactions to the data elements underlying a recently designed library online catalog. The third method, the Web-based survey, targeted librarians and library staff accessing WorldCat via both end-user and staff interfaces. It was appropriate to consider all interfaces to WorldCat data in the library survey, since the primary research interest centered not on a particular interface but on librarian/ staff preferences around catalog data quality. To the extent that WorldCat and library catalogs are comparable, the methodology's focus on data quality, rather than interface, permits insights gained about WorldCat to be germane to the data quality requirements of library online catalogs and integrated library systems in general.

End-user Focus Groups: OCLC commissioned Blue Bear LLC to facilitate three qualitative focus groups during May 2008 in Columbus, Ohio. Blue Bear's facilitator conducted one session each for the following groups: undergraduates ages 18 to 24, casual searchers ages 25 to 59, and faculty and graduate students (referred to as scholars in this report).

End-user Pop-up Survey: OCLC commissioned ForeSee Results to conduct an online pop-up survey targeting worldwide end users of WorldCat.org, OCLC's freely available interface to WorldCat on the Web. ForeSee collected a total of 11,151 total responses between May 12, 2008 and July 9, 2008, with a 4% response rate based on the number of times the survey displayed to users. In this survey, respondents were asked to identify their role as students, teachers/professors, other general searchers or librarians/library staff. To capture the perceptions of end users only, responses from the librarians/library staff group were excluded from the end-user data. Of the 11,151 total responses, 68% were from end users. Forty-four percent (44%) of the end-user respondents were from outside North America.

Library Survey: OCLC commissioned Marketing Backup, an independent marketing research firm, to conduct a Web-based survey targeting librarians and library staff who use the WorldCat database from a variety of interfaces. A total of 1,397 librarians and library staff responded to the survey between September 10, 2008 and November 30, 2008. Respondents had roles as library director/dean, public services/reference, cataloging, acquisitions and interlibrary loan. Thirty-six percent (36%) of respondents were from outside North America.

Data Quality: What End Users Want

Key Findings

The end user's experience of the delivery of wanted items is as important, if not more important, than his or her discovery experience. Appropriate, accurate and reliable data elements supporting the transitional experience from discovery through delivery are critical.

End users expect a seamless flow from discovery to delivery; end users want to know immediately if the item is available and if so, how to get the item. For online materials, end users want more direct links or easier access to the online content, both text and media.

Discovery-related information elements beyond author and title, such as summaries, excerpts and tables of contents, are essential aspects connecting the stages of an end user's discovery-to-delivery experience.

Search results must be relevant and the relevance must be obvious. End users have expectations of the types of results they should get when they conduct searches. They want the library catalog to return relevant results based on those expectations. Behind the scenes, the catalog needs supporting data elements and ranking algorithms that make it obvious to end users why search results are returned.

Keyword searching is king, but an advanced search option (supporting fielded searching) and facets help end users refine searches, navigate, browse and manage large result sets.

End-user Focus Groups Findings

There were several similar findings across the three focus group sessions such as those mentioned in the key findings above: delivery is as important as discovery; summaries, excerpts and tables of content are essential; and the relevance of search results is vital. Other findings differed across the groups. These differences were most telling when considering social features. The focus group participants offered a mixed reaction to social features, such as the ability to create reviews and share lists. Participants in the undergraduate focus group were more favorable toward user-contributed content, and they were quite discerning in their ability to distinguish authoritative from non-authoritative reviews.

Overall, editorial reviews were considered more valuable than end-user-contributed reviews; this opinion was particularly evident when the focus group participant's information need was for academic or professional purposes.

User-contributed reviews were at least somewhat interesting to some participants in all of the focus groups when they were not conducting academic searches. Scholars were more interested in professional reviews and were less interested in other users' opinions on the materials unless they were "experts."

The participants in the undergraduate group found value in user-created lists, which provide a way for end users to keep track of and share items using WorldCat.org. These participants felt it would be beneficial to be able to see which items others are using and for what purpose; they felt that it would add credibility to an item to know who was using it. Participants in the casual searchers and scholar groups did not perceive user-created lists to be useful. In order to see the lists as useful, these participants would like to be more informed about the author/creator of the lists.

End-user focus group participants used the advanced search feature when searching for known items; the ability to be more precise in the search was more likely to ensure successful results. Supporting an advanced search feature involves indexing bibliographic data elements separately—title, author, subject, format, publication date and so on. The effective operation of advanced search features may also involve controlled vocabularies to ensure, for example, that all of the materials authored by Ernest Hemingway are retrieved by a query on Hemingway's name.

Focus group participants found faceted browsing to be a "very comprehensive way to refine" their search results. Provided the data are appropriately indexed, preexisting fields in a catalog database (such as author or publication date) can be extracted

and presented to permit faceted navigation, also called faceted search or browse. Facets enable searchers to progressively narrow their choices from a large and often unwieldy retrieval set. These findings suggest that facets are another effective application of the controlled terminologies and authorized heading forms that have been built and maintained by librarians.

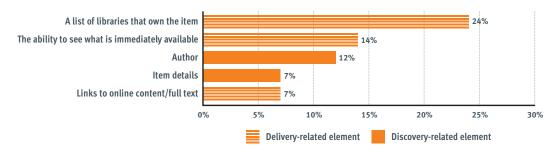
End-user Pop-up Survey Findings

When asked what information in the records was most essential in identifying the items they need, end users responding to the pop-up survey chose a blend of data elements that support the user's delivery experience, as well as some data elements that support the discovery experience.

Most Essential Data Elements

What information is **most** essential in helping you identify the item that you need?





Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (End-user pop-up survey)

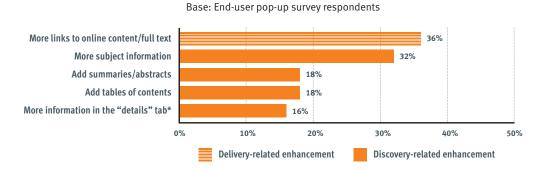
Three of the end users' top five choices included data elements supporting delivery: 24% of the respondents said they want to see the *list of libraries that own the needed item*; 14% said they want to know *what's immediately available*; and 7% want *links to online content/full text*. Two of the end users' top five choices were discovery-related data elements: *author* (12%) and *item details* (7%)*.

^{*}At the time of the pop-up survey, the WorldCat.org "details" tab contained basic bibliographic information plus enriched data such as table of contents and summary/abstract, if available.

When asked what enhancements to a catalog could help them consistently find wanted items, more than one-third (36%) of the pop-up survey respondents answered that having more links to online content/full text would be the enhancement they would want to see.

Top Five Desired Data Quality Enhancements

What changes would be **most** helpful to you in identifying the item that you need?



Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (End-user pop-up survey)

As for the rest of the most frequently chosen enhancements—from more subject information to more information in the details tab*—respondents appear to be expressing the desire for data elements not generally included in a standard catalog description. The desired data elements may be structured or unstructured, but they need to help end users assess the utility of items in a results set and decide which ones merit taking the time to obtain.

End-user Recommendations

Based on the research, end users have the following recommendations concerning library catalogs:

- Improve search relevance
- Add more links to online full text, and make the linking easy
- Add more summaries/abstracts, and make them more prominent
- Add more details such as cover art and summaries in the search results.

^{*}At the time of the pop-up survey, the WorldCat.org "details" tab contained basic bibliographic information plus enriched data such as table of contents and summary/abstract, if available.

Data Quality: What Librarians Want

Key Findings

Librarians and library staff, like end users, approach catalogs and catalog data purposefully. End users generally want to find and obtain needed information; librarians and staff generally have work responsibilities to carry out using catalog data. The data quality preferences of librarians and staff are driven by their work assignments.

Duplicate records (multiple records for the same edition or manifestation) impede the work of librarians and staff. The merging of duplicate records in WorldCat was the top priority of all types of librarians, from all types of libraries, inside or outside North America.

Librarians and staff place priority on enriching catalog records with tables of contents data. From a long list of potential data quality enhancements, respondents consistently selected adding tables of contents to bibliographic records as their second choice (behind merging duplicates). This consistency held across nearly all categories of respondents.

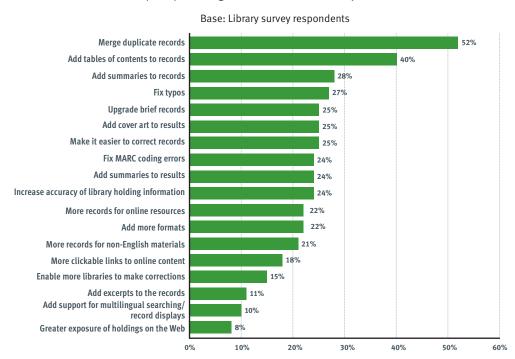
Except for the priorities to merge duplicate records and add tables of contents, significant differences exist in the data quality priorities of librarians by work role, type of library and region. For example, the third most-chosen enhancement for respondents from libraries outside of North America was adding more records for non-English materials; this enhancement did not make it into the top ten list of North American library respondents. Upgrading brief bibliographic records was the fourth most-chosen enhancement by respondents from academic and public libraries, but this enhancement did not make it into the top ten list for special libraries.

Library Survey Findings

To try to compile a comprehensive list of data quality enhancements, reflecting as many points of view as possible, the survey designers looked into the results of many studies of catalog data. The exercise yielded 18 choices to present to survey respondents, who were asked to choose a favorite, plus other enhancements from the survey's list of choices. The chart below gives a ranked order of the choices from all library survey respondents. Merging duplicate records (52%), adding tables of contents to detailed bibliographic records (40%) and adding summaries/abstracts to detailed bibliographic records (28%) are the top three recommended data quality enhancements from all library survey respondents.

Desired Data Quality Enhancements

Which of the following enhancements would you recommend?



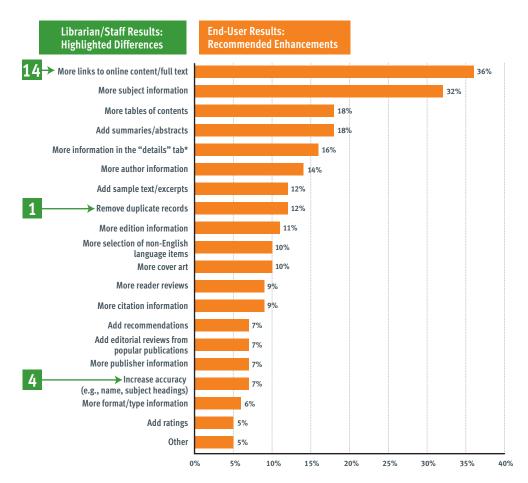
Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (Library survey)

When comparing the data quality enhancements that end users identified with those identified by library staff, both similarities and differences in the data quality enhancements between the two groups emerge. Library survey and end-user survey respondents seem to agree on the importance of adding tables of contents and adding summaries/abstracts; beyond that, various differences emerge. The difference between library and end-user survey respondents' choices pertaining to adding more links to online content seems the most significant; while this enhancement was the first choice of end users, library and staff respondents ranked it in the bottom third of their choices at number fourteen. *Merging duplicate records*, the first choice among library and staff, was ranked eighth by end users, and the library and staff's fourth ranked enhancement, increasing accuracy, was ranked in the end users' bottom third at number seventeen.

Desired Data Quality Enhancements

What changes would be **most** helpful to you in identifying the item that you need?





Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (End-user pop-up survey)

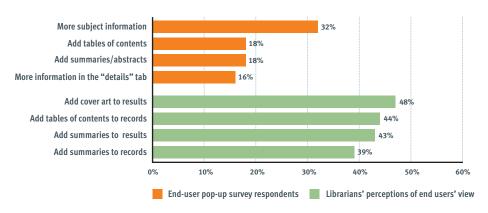
^{*}At the time of the pop-up survey, the WorldCat.org "details" tab contained basic bibliographic information plus enriched data such as table of contents and summary/abstract, if available.

Librarians' Perceptions of What End Users Want

Library survey participants who work directly with users were asked to predict what enhancements their end users would recommend. From a list of 18 enhancements, library survey respondents who work directly with users were asked to select one. then other enhancements to WorldCat that would be the most helpful for their libraries' users. In terms of discovery-related enhancements, nearly half of the library and staff respondents (48%) felt that adding cover art to search result lists would be helpful to their users, followed by 44% who felt adding tables of contents to records and 43% who felt adding summaries/abstracts to search result lists would be helpful. Adding tables of contents and adding summaries/abstracts were also among the enhancements chosen by the end users themselves.

Top Discovery-related Data Quality Enhancements

What changes would be **most** helpful to you in identifying the item that you need? (End-user pop-up survey) Which of the following enhancements would be most helpful for your patrons? (Library survey)

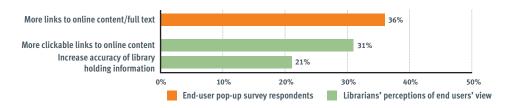


Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (End-user pop-up survey and library survey)

On the delivery side, more links to online content topped both librarian (perception of end users' views) and end-user survey respondents' enhancement choices. Librarians and staff also predicted that end users would be helped by increasing the accuracy of library holdings information—an enhancement that undoubtedly would be helpful in improving the end user's delivery experience, but of which end users would likely be unaware.

Top Delivery-related Data Quality Enhancements

What changes would be **most** helpful to you in identifying the item that you need? (End-user pop-up survey) Which of the following enhancements would be most helpful for your patrons? (Library survey)



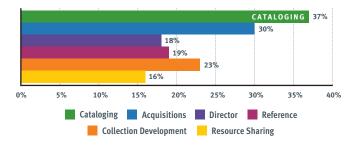
Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (End-user pop-up survey and library survey)

Librarians' Desired Data Quality Enhancements by Job Responsibility

The findings suggest some significant differences among librarian and staff enhancement choices based on job responsibility. For example, cataloging respondents were significantly more likely to recommend upgrading level 3 or other brief records compared to directors, reference, collection development and resource sharing staff.

Upgrade Brief Records

Which of the following enhancements would you recommend?

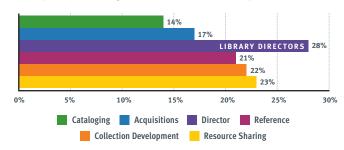


Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (Library survey)

Director respondents were significantly more likely than acquisitions and cataloging respondents to give priority to more clickable links to online content.

More Clickable Links to Online Content

Which of the following enhancements would you recommend?



Source: Online Catalogs: What Users and Librarians Want, OCLC, 2009 (Library survey)

Librarian and Library Staff Recommendations

Based on the research, librarians and library staff have the following recommendation for library catalogs:

- Merge duplicate records
- Make it easier to make corrections to records, possibly using a social cataloging approach
- Place more emphasis on the accuracy and currency of library holdings
- Add enrichment content to the catalog with tables of contents, summaries and cover art
- Seek out more communication about what users say they want.

Conclusions

Catalogs have many constituencies, both inside and outside the library. This study's results suggest that end users place a high priority on enrichment data (tables of contents, summaries, etc.) and on links to online content, both text and media. Librarians and library staff are also important catalog users, and their data quality priorities tend to be different than end users' priorities. If their data quality needs are met, they can do their jobs more efficiently and effectively. And, some of librarians' data quality requirements (to be able to fix mistakes and control heading forms, for example), while not shared by end users, play a role in fulfilling end users' needs.

As noted in the introduction, this report is written for those readers who are seeking to define requirements for improved catalog data. It is for those readers that the following recommendations are made:

- Examine and compare the library's investments in bibliographic work, catalog management, linking functionality and enrichment content (tables of contents, summaries, etc.) and rebalance them as appropriate to better meet end users' catalog data quality requirements.
- Within the library community and with relevant organizations, explore how to
 obtain or produce enrichment content (tables of contents, summaries, etc.)
 through data mining, the use of APIs, partnerships with publishers and vendors,
 and collaborative library projects.
- Encourage the appropriate organizations to complete research and development to improve relevance ranking in online catalogs. Explore the feasibility of redeploying classification data (and the terms associated with classification numbers) and other existing data to improve relevance ranking.
- Pay more attention to the library's delivery services and the data elements that support a positive experience for the end user.
 - Explore the feasibility of home or office delivery of materials, or other faster, more convenient options for the delivery of physical items. Be willing to change policies and workflows, and to change how library human and financial resources are deployed to make this happen.
 - Explore consortial borrowing and lending options. As appropriate, be willing
 to link from the catalog to purchase options and to digitized materials on
 demand. Consider setting a library standard (e.g., within 24 hours) for getting
 requested materials to users.
 - Invest effort in improving the library's linking metadata management and interoperability with licensed and open-access data.
 - Consider adding more digital content that end users can easily link to from the catalog, and associate print descriptions with their digitized counterparts.
 - Where possible, link from the catalog to excerpts or snippets, both text and media.
- Examine the local editorial changes being made to bibliographic records and analyze which ones directly assist end-user discovery and delivery, and which do not matter as much. Redesign procedures and workflows to focus human expertise on what matters most to end users and which must be done manually.
- Libraries will not be able to accomplish what is needed by going it alone.
 Collaboration and a coordinated approach involving many organizations (and even end users) will be required. As noted in section 1 of the Library of Congress Working Group final report, traditional library workflows, featuring the same corrections being done multiple times at multiple libraries, are costly and redundant.¹¹ The right mechanisms for collaboratively sharing the effort of data quality improvements could assure better experiences for end users of catalog data at less cost to libraries.

In a world of unlimited human and financial resources, a data quality program for a library's online catalog could meet all end users' needs and all librarian and staff needs. In a world of limited resources, library leaders must make choices, creatively

deploy the resources they have, and balance competing data quality requirements. A data quality program that strikes a balance between what end users and librarians/ staff want and need, but gives an edge to the desires of end users, seems more likely to assure the library will continue to thrive in the end-user communities it serves.

Notes

- 1. Hill, Janet Swan, "Is it worth it? Management decisions related to database quality," Cataloging and Classification Quarterly 46 (1) (2008): 5-26.
- 2. Miksa, Francis L., Charles Ammi Cutter: Nineteenth-Century Systematizer of Libraries, (dissertation, University of Chicago, 1974) p. 373, www.ischool.utexas.edu/~miksa/ dissertation. html. Cutter's 1876 objectives of the catalog are quoted in Miksa's comprehensive study of Cutter's contributions to librarianship:
- To enable a person to find a book of which the author, the title or the subject is known.
- To show what the library has by a given author, on a given subject or in a given kind of literature.
- To assist in the choice of a book as to its edition (bibliographically) or as to its character (literary or topical).
- 3. A comprehensive review of past and present cataloging theory may be found in The Future of Cataloging: Insights from the Lubetzky Symposium, April 18, 1998, University of California, Los Angeles, edited by Tschera Harkness Connell and Robert L. Maxwell, Chicago: American Library Association, 2000.
- 4. Davis, Carol C., "Results of a survey on record quality in the OCLC database," Technical Services Quarterly 7 (2) (1989): 43-53.
- 5. IFLA Study Group on the Functional Requirements for Bibliographic Records, Functional requirements for bibliographic records: final report, UBCIM publications (München: K.G. Saur, 1998) new ser., v. 19.
- 6. Joint Steering Committee for the Development of RDA, Full Draft of RDA, 2008, www.collectionscanada.gc.ca/jsc/rdafulldraft.html.
- 7. De Rosa, Cathy, Lorcan Dempsey and Alane Wilson, The 2003 OCLC Environmental Scan: Pattern Recognition: A Report to the OCLC Membership (Dublin, OH: OCLC Online, 2003).
- 8. De Rosa, Cathy, Perceptions of Libraries and Information Resources: A Report to the OCLC Membership (Dublin, OH: OCLC, 2005).
- 9. Case, Donald O., Looking for information: a survey of research on information seeking, needs, and behavior (San Diego CA: Academic Press, 2002) p. 140. The Principle of Least Effort asserts that information seekers will use information sources that are the easiest and most convenient, even when better or more authoritative sources are available, but not as easy or convenient to use or find.
- 10. Byrum, John D., "Recommendations for urgently needed improvement of OPAC and the role of the National Bibliographic Agency in achieving it" (paper prepared for the 71st IFLA General Conference and Council, Oslo, Norway, August 14–18, 2005), www.ifla.org/IV/ifla71/papers/ 124e-Byrum.pdf.
- 11. Library of Congress, On the Record: Report of the Library of Congress Working Group on the Future of Bibliographic Control. Washington, D.C.: Library of Congress, recommendation 1.1, "Eliminate redundancies," 2008, www.loc.gov/bibliographic-future/news/lcwg-ontherecordjan08-final.pdf.

For updates and more information about Online Catalogs: What Users and Librarians Want please visit:

www.oclc.org/reports/



6565 Kilgour Place Dublin, Ohio 43017-3395 1-800-848-5878 +1-614-764-6000 Fax: +1-614-764-6096 www.oclc.org

ISO 9001 Certified