



ADVISOR REVIEWS—STANDARD REVIEW

WorldCat

Date of Review: February 15, 2008

Composite Score:

★★★★

Reviewed by: Jeffrey Beall

Auraria Library

University of Denver

1100 Lawrence St.

Denver, Colo. 80204

<jeffrey.beall@cudenver.edu>

Product Description

The name is commonly pronounced as “World cat dot org,” but on the site, it refers to itself as just WorldCat. Like other Web 2.0 products, this one is likely to function in an almost-endless beta state as features and improvements are added and made.

WorldCat.org is a free, Web-based version of the master OCLC WorldCat bibliographic database, enhanced with additional data and with most of the common Web 2.0 and Library 2.0 features, such as the ability to create original lists of records and share them. WorldCat.org contains over 100 million records. The additional data includes OCLC’s ArticleFirst, a table-of-contents database that indexes more than 19,000 periodical titles from 1990 to the present and includes about 20 million records. It also includes all the data from the OCLC FirstSearch databases GPO, MEDLINE, and ERIC. These databases cannot be searched individually; you have to search across all of WorldCat.org in every search.

The chief value of WorldCat.org is that it provides free, high quality bibliographic data for tens of millions of published and unpublished items in virtually every format. OCLC has given away the store. However, for their holdings to be exposed on WorldCat.org, libraries have to have a current subscription to OCLC FirstSearch. According to the OCLC Web site:

However, search results do not necessarily return all available matching items, and when a user generates a list of holdings for an item in a particular geographic area, the listings will not necessarily display all libraries in the defined region that actually have the item. This is due to the requirement that WorldCat-cataloging libraries subscribe to FirstSearch in exchange for exposure of their holdings on WorldCat.org (and Open WorldCat partner sites such as Google and Yahoo!).

Search results may not display all matching items in the complete WorldCat database because some items, while being cataloged in WorldCat, are not actually held by any of the WorldCat libraries that presently meet the requirements for Web exposure of holdings. (In addition, some items cataloged in WorldCat may not be presently held by any WorldCat libraries at all, regardless of whether they meet the requirement.)

A user viewing the display of WorldCat libraries in his area that own an item will only see libraries that have it and meet the requirement for holdings exposure. WorldCat libraries in the area that own the item but have not subscribed to FirstSearch will not appear in the listings.

The main page of the site is a simple search box, much like Google’s, except that this is a metadata, and not a full-text, search. The results

are sorted by “relevance,” although what determines relevance is never explained. As an alternative to relevancy ranking, it’s possible to manually sort the search results by record attributes, such as author, title, and date. A column called Refine Your Search appears on the left side of each search results page. This postsearch feature allows users to cluster their search results by author, content, format, language, and year of publication. The “content” limit appears to be based on call number groupings. Once a user selects an individual record to look at, there is a lot of information on the record. For most records, this information is taken from the raw OCLC MARC record. It is rich metadata that provides a standard description of the item, subject analysis, and holdings information. WorldCat here also augments many of the records for books with cover images. The article-level records are not as rich as the records from the main OCLC MARC file. Many are brief records that lack subject metadata.

The main Web 2.0 functionality of the site is the ability to link to and share links to individual records or to user-created lists of records. The site provides a means of creating bibliographies based on any search criteria the user chooses, that is, by subject, author, etc. The lists can be displayed in WorldCat’s default format, or they can be switched to Citations View, which displays the lists in popular citation formats, such as Chicago, Harvard, APA, and Turabian. Users can also add records to one or more of about thirty different social bookmarking sites, such as del.icio.us, Google bookmarks, etc.

It’s also possible to create stable links to individual records and embed these links in Web pages or e-mail messages.

The site also allows users to contribute reviews. Reviews include a rating (excellent, very good, good, fair, poor), a title, and the review itself. After the review is written and saved, other users can access it by clicking on the Reviews tab in the main record. Content guidelines for the reviews are very similar to those of Amazon.com; in fact, the whole review feature here is similar to Amazon’s, except that on WorldCat the reviews are hidden behind a tab and are not included on the main screen for the item. There is also a feature called Edit Public Notes, but this feature is not explained in the Help screens. It appears to be a way of adding additional user-generated descriptive or subject data. As is possible in other Web 2.0 services, users of WorldCat.org cannot add their own tags or create new records, but a tagging functionality may be added in the future.

One thing that WorldCat can offer that its competitor sites cannot is its Find in a Library functionality. Each individual record has a Get It section that does two things. It displays nearby libraries that hold the item, sorted geographically. Users’ location data is stored in their profiles, and the system uses this information to create the geographic sort. In most cases, the links to local libraries lead directly to the libraries’ catalog records for the item. OCLC calls this, “interoperabil-

ity with the local delivery environment.” The other thing the Get It section does is link to commercial sites—usually Amazon.com—that sell the item. However, sometimes the site also links to the vendor Baker and Taylor. When this link is specified, a portion of the proceeds from any purchase is shared with a purchaser-selected library, but users have made relatively few such purchases so far.

Additionally, some individual records have two other links on them: Search for Full-Text and Interlibrary Loan. The first of these connects to your institution’s OpenURL resolver, and the second connects to your institution’s interlibrary loan department’s Web page. A library administration mode in WorldCat.org allows libraries to configure what users see when a record for a book the library holds is displayed. Libraries can even add a link here that says Ask a Librarian. This link leads directly to the library’s page.

The interface is available in six languages. Other OCLC products, such as Open WorldCat and WorldCat Local are also based on the WorldCat MARC database. These are separate products, but because they are based on the same data, there is some overlap with WorldCat.org. Sometimes it is hard to specify exactly where one product starts and another ends. It’s easy to confuse them.

Several search plug-ins have been created for WorldCat.org. These include a Facebook application, a Firefox browser search extension, and an ISBN matching functionality in the Google toolbar.

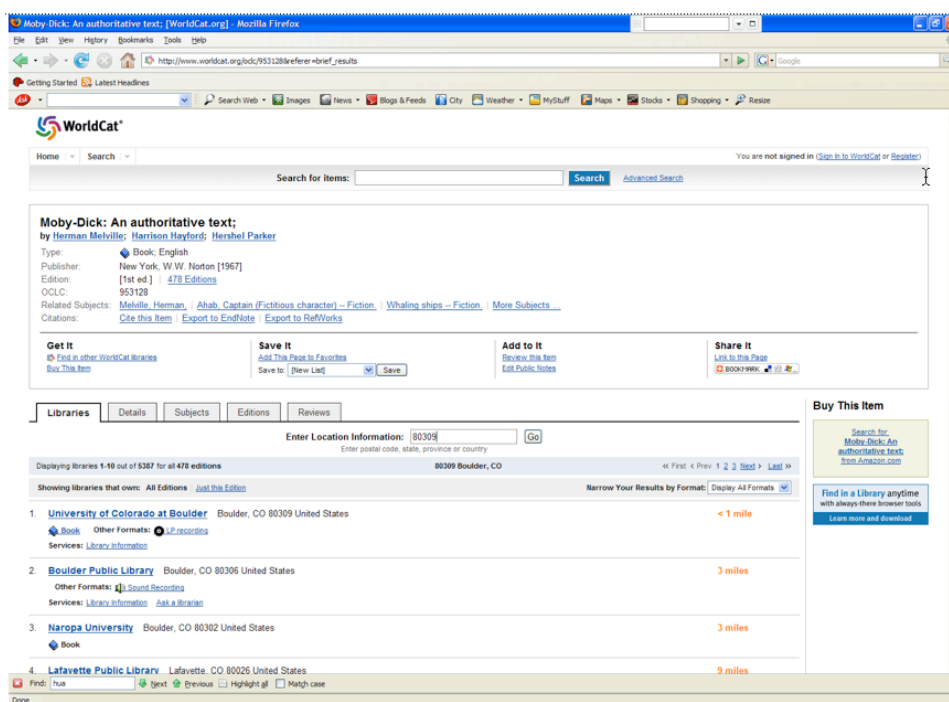
SEARCHING

The site offers a simple and an advanced search. The simple search is intentionally designed to mimic Google’s simple search box. The advanced search capitalizes on the rich metadata in the OCLC MARC records and offers author, title, subject, and number searches. The number searches are a single index that search ISBN, ISSN, and OCLC numbers. The advanced search allows for limiting searches by format, publication date, content (such as fiction or nonfiction), audience (such as juvenile), and language.

The most powerful limiting functionality of WorldCat.org occurs with both the simple and advanced searches and is a column on the left of the results page with links that effectively narrow a search by a number of facets, or characteristics, of the resources that are represented in the metadata. For example, I did a search on “Colorado” and got 387,954 hits. I then clicked on Format: Book (234,565) on the left and limited my search to books. Next I clicked on Year: 1996 (5,259) to further narrow my search. Then I clicked on Language: English (5,184) to narrow my search even further. At this point, the system still offered me several options to further narrow my search, including by author, content (thesis, fiction, or nonfiction), and topic. This last sort, by topic, appears to be chiefly based on the Library of Congress classification number and by the data in the MARC record that indicates whether a publication is a government document.

If you’ve heard of FRBR (functional requirements for bibliographic records) and wanted to see an application of this new library technology, WorldCat.org is a good place to do that, for it provides a FRBR-ized interface. FRBR is a way of clustering search results. It groups together the various versions of a work to reduce the number of separate records in a search results display, thus making the retrieval more compact. For example, when one searches “Moby Dick” in WorldCat.org and clicks on the first record retrieved, there appears a hotlink that says, “469 Editions.” If you click on the link, you get a search result that includes all 469 editions listed individually. Each of these editions contains the same intellectual content, but they differ in things like publisher, format, year of publication, hardback or paperback, etc.

The WorldCat.org search box is freely available as a download and can be placed on any Web site. Also, each record has a link called, Link to This Page. Clicking here generates a short, stable URL that can be linked to from a Web site. According to the OCLC Web site, this feature allows you to “Use a simple URL syntax to create Web-distributable links to a WorldCat item record or integrate into a ‘mash-up’ Web application.”



Critical Evaluation

WorldCat.org is one of the biggest internet giveaways ever. Never has so much data that was at one time proprietary and strictly controlled been made so freely available. This giveaway is in keeping with the Web 2.0 principles of sharing content and user collaboration. Literally hundreds of millions of records representing books, serials, videos, journal articles, and more are now available for free through WorldCat.org. Additionally, library holdings information is included with each record, informing searchers which libraries own a specific work.

FIGURE 1 Sample Screen from WorldCat.org



WorldCat Review Scores Composite: ★★★★★

The maximum number of stars in each category is 5.

Content: ★★★★★ 1/2

The entire content of the WorldCat database, plus tens of millions of records for articles and documents are now freely available to anyone with Web access. Library holding information is also available, provided that certain requirements are met. In some cases, some content is excluded from search results.

Searchability: ★★★ 1/2

The simple search box works just like Google. The FRBR-ized results display makes search results easier to sort through, and the Refine Your Search functionality greatly enhances search precision and recall. Subject searches will miss records lacking subject metadata, and some natural-language searches will miss records containing non-natural-language data.

Price: N/A

Free, but for your own library's holdings to appear in the search results, it has to subscribe to OCLC's FirstSearch product.

Contract: N/A

The site provides links to a document entitled "OCLC WorldCat.org Services Terms and Conditions" [<http://www.oclc.org/worldcat/policies/terms/>](http://www.oclc.org/worldcat/policies/terms/) that provides a standard list of terms and conditions. The contract stipulates that data from this product may not be used for commercial purposes, for cataloging, or for systematic harvesting of data. The disclaimers make clear the data is provided on an "as is" basis.

By freely exposing this data, OCLC hopes to bring Web users to libraries and to library services. Each of the records has a stable URL and is indexed in the main internet search engines. So internet search engine users just might end up visiting a library's Web site or the library itself as a result of a search. Library data is functioning as advertising for libraries. This is a true example of repurposing data.

This resource is still relatively new, and as in many Web 2.0 services, development will be ongoing. A few minor problems with the site will certainly be repaired soon. For example, in Citations View, colons are deleted from titles. There is also a problem with searching for lists; if you search "HIV," you only get results that contain those three letters, like "Archivos." Unlike other sites, you cannot stay logged in to WorldCat.org on the same computer. You have to log in again on each visit. There's a way to mark "favorites" but it's not what you might expect. Favorites are saved in your browser and not on the site itself.

WorldCat.org is not a full-text database; it is a bibliographic database. But the searching is full-text. That is to say, users do a full-text search on metadata. This can be problematic at times because the data in the records is not all in natural language. For example, content standards used in MARC records prescribe the use of abbreviations in headings, such as "Dept." for "Department." These differences can cause missed retrieval. Moreover, the system does not effectively provide alphabetically-sorted browse displays as online catalogs do. Search results can be sorted by author or title, but the displays are limited to ten titles per screen, and the title always appears at the head of each entry, leaving the searcher to wonder why the sort element isn't always visible.

There is no way to see MARC data in WorldCat.org. But the entire database is not based on MARC metadata. Citation databases, like

ArticleFirst, are in some other format. Moreover, the ArticleFirst data is only table of contents data—it does not include subject metadata. Therefore users performing a subject search might think that their retrievals are complete, when in fact most of the records from the article databases will be excluded from the search results, because they don't contain any subject metadata.

Creating lists can be messy: if you mark records from a search result screen, then go to another screen, those selections are lost. So to create a list, you have to select records from a search result screen, add them to the list, rerun the search, and then add additional items.

Authentication

The site is free, but it might recognize your IP address, and you may have to enter your U.S. zip code; this makes the Find in a Library feature work. Also, for the article databases, according to the OCLC Web site,

Users recognized as being on your library network through IP authentication also see fulfillment options you may offer for a particular article. These can include a direct link to full text

Contact Information

URL: [<http://www.worldcat.org/>](http://www.worldcat.org/)

Bob Schulz [<schulzb@oclc.org>](mailto:schulzb@oclc.org) is OCLC's product manager for WorldCat.org. There is also a Feedback link on the site itself that leads to a form users can fill out to report problems or ask questions.

from your subscription-based online databases; a link to your library's OpenURL resolver; or a link to your article delivery service.

Libraries can enter their IP addresses in the FirstSearch administrative module.

References

Free, David. (2007). "New features from WorldCat.org." *College and Research Libraries News*, v. 68, no. 10, p. 622.

Hansen, R. (2007). "WorldCat.org." *Choice: Current Reviews for Academic Libraries* vol. 45, Issue 3, p. 436.

Quint, Barbara (2007). "OCLC's WorldCat Local: A Promising Development for Library Patrons." *Information Today*. NewsBreaks

April 23, 2007. Online <<http://newsbreaks.infotoday.com/nbReader.asp?ArticleId=35939>>.

About the Author

Jeffrey Beall is Catalog Librarian at Auraria Library, University of Colorado at Denver. He previously worked as a cataloger in Widener Library at Harvard University. He earned his M.S.L.S degree from the University of North Carolina at Chapel Hill. He also holds an M.A. degree in English from Oklahoma State University. His recent research interests include access problems caused by typographical errors in bibliographic databases and bibliographic database maintenance. ■