

Worldcat Discovery - Making European Libraries More Visible on the Web

by JANET LEES

INTRODUCTION

Firstly may I thank Trix Bakker and the Collection Development Division for inviting me to participate in their session today.

I'm going to address the issue of making library collections more visible on the web and in particular the OCLC Open WorldCat program, the recent European enhancements to this program and the pilot that is currently underway.

I was struck by Eric Lease Morgan's "Top Tech Trends for ALA 2006" posted to the LITA blog on 18th June 2006 (Morgan, 2006). One trend he noted in his post was a "Growing discontent with the library catalogs" in which he stated: "people know libraries contain large quantities of authoritative information, but ease of access and convenience trump such qualities in a heartbeat".

Morgan's comments support the findings of an OCLC study "Perceptions of Libraries and Information Sources" that was undertaken in May - June 2005 by Harris Interactive (Perceptions, 2005). This online survey was completed by 3,348 respondents 14 years and older from Australia, Canada, India, Singapore, United Kingdom and United States. The survey was conducted in English, and was confined to those that had access to the internet. A subset of this survey "College students' perceptions of libraries and information services"³ which may be more pertinent to this group has now been published (College, 2006). This subset comprised 396 college students aged 15-57, undergraduate and graduate from all six countries of the original report.

There are a couple of headline statistics that probably won't surprise anyone in this audience. The first is that the search engine is the start point of choice for almost all information searches - both by college students (89%) and the total respondents (84%).

The second, possibly more reassuring, headline is that the Librarian (person not online service) is the pre-eminent first source of help at the Library (76%).

The third, again not surprising, headline is that since college students are spending more time on the web they are spending less time doing other activities like watching TV (40% less), visiting the library (39% less), reading newspapers (24% less), listening to the radio (19% less) and visiting with friends and family (14% less).

The report findings can be summarised in three areas:

1. on students perceptions and habits:

- 89% college students use search engines to begin information searches whereas only 2% begin their search on the library web site.
- College students like search engines - 93% describe their experience as being either "satisfied" or "very satisfied".
- College students like to self-serve. 54% do not seek assistance when using library resources, although more students sought assistance than respondents overall.
- College students represent a higher level of use of electronic resources than respondents overall.

2. on libraries:

- College students have the highest rate of usage of libraries and broadest use of both p(rint) and e(lectronic) resources of any group.
- College students use the library but they use it less and read less since they began to access the web.

- The most frequent use of the library by college students is as a place to study.
 - Only 10% of college students indicated that their library's collection fulfilled their information needs.
3. on Alternatives to libraries:
- College students use personal knowledge and common sense and cross referencing other sites to judge if electronic information is trustworthy.
 - 50% of college students learn about electronic sources from their teachers, 36% from a library web site and 33% from a librarian.
 - Search engines fit college students' lifestyles better than physical or online libraries. The majority of college students see search engines as a perfect lifestyle fit.

MAKING LIBRARY HOLDINGS MORE VISIBLE TO THE ONLINE SEARCHER

So how can libraries raise their visibility in the search engines results set? We all know that search engines guard their weighting algorithms fiercely and that individual items can easily be lost ... and lost often means that the result does not make the top few pages of the results set. Many libraries individual resources are indexed by the search engines but as individual resources are often lost in long result sets. OCLC by allowing its union catalogue containing holdings of thousands of libraries to be ranked by the search engines could through its members create a 'Libraries Brand', which, it was thought, would make library holdings more visible to the online searcher. To test this hypothesis OCLC began its Open WorldCat pilot in 2004 with two partners Google and Yahoo!.

The partners are allowed to harvest limited field sets in the OCLC WorldCat records and then display the results under the brand 'Find in a Library'. If a user then clicks on a 'Find in a Library' result he is linked through to the OCLC Open WorldCat interface where he can see details of the bibliographic record and a list of libraries holding the item. A geographic filter using postcodes help the user find libraries closest to him and through deep linking set up by OCLC and the library he can be directed to the selected library's OPAC and through authentication potentially the full text / image of the item.

The benefits of this approach are evident for both the searcher and for libraries. It gives library names, collections and services greater visibility in the major search engines and other sites on the open web and allows web searches to find and access library materials and services as part of their normal search engine workflow. It also reveals a portion of the otherwise hidden web at the point of need.

The original partners Google and Yahoo! continue to harvest significant subsets of WorldCat and OCLC is now flagging records with country flags thus allowing the weighting of country holdings to be increased in search engines appropriate country sites. OCLC has also begun to work with Microsoft on their Academic Live Search now in Beta, with Ask.com and with the Google specialist sites Google Scholar, Google Book Search and Google Print.

For those of you who have not looked at Open WorldCat [1] lets briefly walk through the experience from a web searchers perspective. For convenience Google, Yahoo! and Firefox offer toolbar downloads to enable the search box to be positioned on the desktop and OCLC working with OCLC PICA has developed user interfaces in English, German, Spanish, French and Dutch.

1. The user enters a search in the search engine of choice.
2. A result including the phrase 'Find in a Library' is selected.
3. The Open WorldCat entry provides a short bibliographic display.
4. A post code is entered for the users geographic location.
5. Open WorldCat shows libraries holding the item ranked by distance from the users postcode.
6. Once the user selects a holding he can link through to the library's OPAC and also other library services such as virtual reference and ILL.

Other features include a 'permalink' to help people cite specific records in the Find in a Library interface. The permalink eliminates session specific information from a linking url that can get in the way of using the links on an ongoing basis. People who wish to link to WorldCat records from course syllabi, bibliographies etc can find the permalink useful. Open WorldCat also provides "buy it" links, which link to online booksellers including Amazon.com as an alternative to borrowing.

The Open WorldCat pilot served to provide the proof of concept and searches clicking through to the Open WorldCat interface have grown steadily since the pilot began and now ranges from 5-9 million each month with a seasonal pattern. There are currently some 17,000 libraries whose holdings are made visible through Open WorldCat and there were nearly 150,000 referrals to these libraries in May 2006. There have been a total of 120 million referrals since the project began with 83% of users who arrive at the "Find in a Library" interface proceeding to a library OPAC.

Libraries who wish to have their holdings made visible in Open WorldCat must have their holdings included in WorldCat. Traditionally this has required libraries to either catalogue directly in WorldCat or batchload their holdings. Batchloading is undertaken by some UK CURL libraries, South African libraries who contribute to SACAT, some national bibliographies like that of the Czech Republic and national union catalogues like the Polish NUKAT libraries.

THE WORLDCAT DISCOVERY MODEL

However this only creates a 'partial World' Cat and excludes the many millions of library holdings that are represented in European national and regional union catalogues who up to now have had no affiliation with OCLC. OCLC PICA therefore worked with OCLC to create a European model which is gaining in popularity both within and beyond Europe called the WorldCat Discovery model. This model creates a relationship between union catalogue owners and OCLC for the exchange of bibliographic records and holdings that can be harvested through the Open WorldCat programme. Thus individual libraries can have their holdings made visible in Open WorldCat whilst they continue to catalogue in their regional or national union catalogue. The addition of holdings in this way also qualifies the libraries in the union catalogue to become OCLC members and participate in the global cooperative.

A joint OCLC / OCLC PICA working group was formed in January 2005 to develop and test the WorldCat Discovery model. This group has worked to

- Load Link UK and the Dutch NCC databases to create sufficient library holdings to conduct a pilot in the UK and Netherlands. A further pilot is planned for Germany as several German union catalogues and some individual libraries have agreed to load their data for this purpose.
- Implement the ZING Update (more formally SRU record update [2]) standard to synchronise record identifiers between the union catalogues for automatic updating.
- Collaborate on the use of FRBR (Functional Requirements for Bibliographic Records) in display sets.
- Specify, develop and implement European enhancements to the Open WorldCat interface including the language interfaces, European geospatial data and 'buy it' options

The pilot is currently underway in the UK and Netherlands and will continue until the end of August. Feedback from the pilot libraries will be used to further refine the service to meet the needs of European libraries and, if successful, to pave the way for the model to be used in other parts of the world.

Finally, based on the success of the Open WorldCat project, OCLC has now decided to launch (in August) its own destination portal called WorldCat.org which will include the entire 70 million records and billion holdings that make up the WorldCat database. Librarians and other web site owners will be able to add the WorldCat.org search box to their sites. OCLC plans to further develop the services available through WorldCat.org to continue to help libraries increase their visibility in the web world and deliver services that match end users needs.

Thank you

NOTES

1. Open WorldCat information available at <http://www.oclc.org/worldcat/open/default.htm>
2. Documentation at <http://www.loc.gov/standards/sru/record-update/>

REFERENCES

College Students' Perceptions of Libraries and Information Resources. A Report to the OCLC Membership. Dublin, Ohio : OCLC, 2006. <http://www.oclc.org/reports/perceptionscollege.htm>

Morgan, Eric Lease: "Top Tech Trends for ALA 2006 : 'Sum' pontifications", June 18th, 2006.
<http://litablog.org/2006/06/18/eric-lease-morgans-top-tech-trends-for-ala-2006-sum-pontifications/>

Perceptions of Libraries and Information Resources. A Report to the OCLC Membership Dublin, Ohio : OCLC, 2005.
<http://www.oclc.org/reports/2005perceptions.htm>

WEB SITES REFERRED TO IN THE TEXT

LinkUK. <http://www.linkuk.org/>

NCC - Dutch Central Catalogue. <http://picarta.pica.nl/DB=2.4/LNG=EN/>

WorldCat.org – Beta version. <http://www.worldcat.org/>