

2007

OPAC 2.0: Next generation online library catalogues ride the Web 2.0 wave!

Katie Wilson
Southern Cross University

Publication details

Post-print of: Wilson, K 2007, 'OPAC 2.0: Next generation online library catalogues ride the Web 2.0 wave!', *Online Currents*, vol. 21 no. 10, pp. 406-413.

ePublications@SCU is an electronic repository administered by Southern Cross University Library. Its goal is to capture and preserve the intellectual output of Southern Cross University authors and researchers, and to increase visibility and impact through open access to researchers around the world. For further information please contact epubs@scu.edu.au.

OPAC 2.0: next generation online library catalogues ride the web 2.0 wave!

In today's fast moving world of technology, we can be sure that things will change, but we can't predict the next biggest development. Web 2.0 started as a brainstorming conference on new web applications emerging from the ashes of the Internet dot.com collapse in 2001¹. Web 2.0 includes concepts such as social networking and participation on the Internet, user community involvement and tagging, blogs, wikis and syndicated feeds or alerts. Some of the best known Web 2.0 sites offer a public space to record and categorise data and images using folksonomy tagging for subject keywording. At Flickr (<http://flickr.com>) you can store, tag, share and retrieve your photo images online, in Del.icio.us (<http://del.icio.us>) tag your bookmarked web sites, and on Library Thing (<http://librarything.com>), catalogue books to keep track of what you've read. Amazon.com fine tuned the incorporation of user feedback and engagement into search results and online shopping. Google has turned around user expectations of search and discovery, while challenging the role of libraries in providing access to information with its digitization program.

Library 2.0 is the extension of Web 2.0 to library services, enhancing user services such as community participation and feedback and offering opportunities for online social networking. Some of these ideas are not new to libraries, but the technology is new, and it is enabling the reinvention of services in the library sphere. Some forward looking libraries have started re-examining the function, design and usefulness of the online catalogue², rethinking search and discovery. Integrated library system (ILS) vendors have also started to see the opportunities for extending Web 2.0 applications to the OPAC to transform the user experience and improve the catalogue's usefulness and usability.

OPAC 2.0 is the application of Web 2.0 to online catalogues and is also referred to as next generation, third generation or twenty-first century catalogues.³ The first generation online catalogue in the 1970s provided computerised access to catalogue records using the MARC bibliographic format. The next OPAC development brought keyword searching and Boolean operators to combine keyword search terms. In the 1990s the first Web graphical interfaces and hypertext linking transformed command driven online catalogues to mouse click driven navigation but there was little change to the underlying search and retrieval or discovery functionality. The catalogue remained primarily an inventory of library holdings and a passive exchange of information – users input search terms they hoped would describe what they were

¹ O'Reilly, Tim. What Is Web 2.0? Design Patterns and Business Models for the Next Generation of Software. <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>

² *Rethinking how we provide bibliographic services for the University of California* Bibliographic Services Task Force, The University of California Libraries: Final Report : December 2005 <http://libraries.universityofcalifornia.edu/sopag/BSTF/Final.pdf>

³ Antelman, Kirstin, Emily Lynema and Andrew Pace. Toward a twenty-first century library catalog. *Information Technology and Libraries*, September 2006, p.128-139; Next Generation Summit Search Interface Working Group http://64.233.167.104/search?q=cache:DZLwFGmVG04J:www.orbiscascade.org/staffhome/Next_Generation_Catalog-report.pdf

looking for and OPACs returned lists of catalogue records that tried to match users' requests.

What are the OPAC 2.0 features?

Two strands of OPAC 2.0 developments enabled by Web 2.0 technologies have emerged: (1) extending the usefulness and search features of the catalogue by harnessing more bibliographic MARC and circulation data for searching, and seamlessly incorporating data from other resources; (2) social networking with personalisation and user community tagging and reviewing to provide a richer discovery experience.

OPAC 2.0 offers improved searching experiences through configurable relevance ranking, subject tag maps or clouds, clustering or faceting for filtering and expanding search results, suggestions for additional searching based on authority headings, indexing of data from several sources, such as circulation data (for most popular items) and other catalogues or databases (incorporating results from federated search and OpenURL linking), spelling corrections using the catalogue database contents and community/user reviewing or tagging.

Another opportunity OPAC 2.0 provides is the ability to show more clearly what the online catalogue can do, and how to use it. There is evidence in the literature that people find online catalogues hard to use.⁴ Limitations in the technology have prevented libraries from improving the searchability and success rate of the online catalogue experience. Advanced search screens offering combined Boolean search options and pre- and post-search modifications that provide valuable search filtering options have been relegated to second level access. Helpful instruction pages with explanations and examples are linked from search screens but they are rarely consulted by users. The thinking has been not to clutter the screen with too much information, but this has left users to struggle often despairingly through using the online catalogue.

OPAC 2.0 enables clearer expression of catalogue functionality through the new features by bringing refining options directly to the results screens. Some recent OPAC 2.0 implementations illustrate the possibilities.

NCSU Libraries using Endeca (figure 1) includes several post search choices and clear options to view a record, all on the results screen page:

- Narrow by call number range
- Narrow results by Subject: Topic
- Brief view | Full view
- Sort by relevance

⁴ Antelman, Kirstin, Emily Lynema and Andrew Pace. Toward a twenty-first century library catalog; Calhoun, Karen. *Changing nature of the catalog and its integration with other discovery tools* Final report March 17, 2006, prepared for the Library of Congress <http://www.loc.gov/catdir/calhoun-report-final.pdf>, p.25

The screenshot shows the NCSU Libraries search interface. At the top, there is a navigation bar with links for 'Search the Collection', 'Browse Subjects', 'Services', 'Library Information', 'Community', and 'News & Events'. Below this is a 'MY LIBRARY' section with links for 'Library Account', 'My Course Reserves', 'My Alerts', and 'RefWorks'. The main search area has a search box containing 'digital libraries', a location dropdown set to 'Anywhere', and a 'Go' button. To the right, there is a 'Send search to:' dropdown and another 'Go' button. Below the search bar, the results are displayed for the search term 'digital libraries'. It shows 525 matching items. There are filters for 'Narrow By Call Number Range' and 'Narrow Results By: Subject: Topic' and 'Subject: Genre'. The results list includes three items, each with details like author, published date, format, and online availability. The first item is 'Digital libraries : international collaboration and cross-fertilization : 7th International Conference on Asian Digital Libraries, ICADL 2004, Shanghai, China, December 13-17, 2004 : proceedings' by International Conference on Asian Digital Libraries (7th : 2004 : Shanghai, China), published c2004, format eBook, online view resource online, available at D.H. Hill Library, call number ZA4080 .I578 2004, location Stacks (9th floor).

Queens Library in Queensborough, New York using Aquabrowser (figure 2) extends discovery options offering up front on the results pages:

- related search terms
- spelling variants and translations
- sort options

The screenshot shows the Queens Library search interface. At the top, there is a navigation bar with links for 'My Account', 'My Queens Library', 'Renew & Request Items', 'Web Membership', and 'Print this page'. Below this is a search bar containing 'digital libraries' and a 'Go' button. To the right of the search bar, there is an 'Advanced Search' link. Below the search bar, the results are displayed for the search term 'digital libraries'. It shows 240 results, sorted by relevance. There are two recommended items, each with details like author, published date, format, and online availability. The first item is 'Digital library use' by Ann P. Bishop, Nancy A. van House, Barbara Pfeil Buttenfield, published 2003, format digital (4) ; libraries (3) ; Library (1), series Digital libraries and electronic publishing, subject digital libraries, information technology. The second item is 'Managing the digital library' by Roy Tennant, published 2004, format digital (3) ; libraries (2) ; Library (2) ; Digitization (1), subject digital libraries. On the left side, there is a 'Bibliothek' section with a tree view of related terms like 'Bücherei', 'LIBRARIES', 'school', 'association', 'management', 'media', 'history', 'copyright', 'guide', 'encyclopedia', 'digital libraries', 'catalogue', 'image', 'law', 'congress', 'system', 'photography', 'people', 'science', 'bookshop', 'libraries', 'digital', 'biblioteca'. On the right side, there is a 'Select Location' section with a dropdown menu for 'All locations', a 'Refine' section with links for 'Catalog', 'Format', and 'External Sources', and a 'Login' section with fields for 'Username' and 'Password' and a 'Login' button.

LibraryLabs from the National Library of Australia (figure 3) offers:

- related subject headings broken down by subdivision, or subject facets
- suggestions to search by related author or subject
- date clustering

LibraryLabs Searching Bibliographic Records NATIONAL ARCHIVES

Search for: RSS PREFER ITEMS HELD BY YOUR LIBRARIES ONLINE

CLICK TO SHOW RANKING WEIGHTS

Found 102243 in 272ms, showing first 100

Do you want to search for works by association for computing machinery

Possibly related searches:
 For *Digital libraries* See also [Libraries](#)
 For *Digital modulation* See also [Digital communications](#)
 See also [Modulation \(Electronics\)](#)

1. **Digital libraries ... : proceedings of the ... ACM International Conference on Digital Libraries.** ACM International Conference on Digital Libraries. New York, N.Y. : Association for Computing Machinery, 1996- v. : ill. ; 28 cm. score 21157, held by 4 libraries
2. **Digital libraries : research and technology advances : ADL '95 forum, McLean, Virginia, USA, May 1995 : selected papers / Nabil R. Adam ... [et al.], eds.** Forum on Research and Technology Advances in Digital Libraries (1995 : McLean, Va.) Berlin ; New York : Springer, c1996. xiii, 390 p. : ill. ; 24 cm. score 17027, held by 13 libraries
3. **Digital libraries : the ... ACM Conference on Digital Libraries.** ACM Conference on Digital Libraries. New York, N.Y. : Association for Computing Machinery, c1998- v. : ill. ; 28 cm. score 14801, held by 1 library [ONLINE]
4. **Digital libraries : people, knowledge, and technology : 5th International Conference on Asian Digital Libraries, ICADL 2002, Singapore, December 11-14, 2002 : proceedings / Ee-Peng Lim ... [et al.], eds.** International Conference on Asian Digital Libraries (5th : 2002

Topics ordered by occurrence in the search results. Click a topic to add it to the search.

Results by year:

2000-	100%
1990	80%
1980	10%
1970	0%
1960	0%
1950	0%
1940	0%

Who are the players?

North Carolina State University Libraries (NCSU) was the first library to implement an OPAC 2.0 or twenty-first catalogue in response to the “severity of the catalog problem, particularly in the area of keyword searching”.⁵ NCSU purchased the Endeca Information Access Platform software to replace its ILS keyword searching. Endeca (<http://www.endeca.com>) search engine software is used in several major commercial websites and TLC, the Library Corporation, partnered with Endeca to distribute and develop Endeca search for the library market. The Endeca software is separated from the databases it indexes and the user interface, and sits on top of the SirsiDynix integrated library system. This allows customization of the interface but it also requires daily export of bibliographic and holdings data from the catalogue database. The search interface provides faceted browsing, configurable relevance ranking and spelling correction, as well as guided navigation. Exposing Library of Congress subject headings and classification schemes provides the option to narrow or refine a search by call number range, subject – topic or genre, format, library, region, era, language or author (see figure 1). You can also browse the whole catalogue without entering a search term. NCSU went live in January 2006 with its new and improved search experience (<http://www.lib.ncsu.edu/catalog>). Usability studies showed a fifty per cent faster completion of set tasks searching the Endeca catalogue compared to the libraries’ previous catalogue.

For more background on the NCSU Endeca catalogue see <http://www.lib.ncsu.edu/Endeca>.

The State Universities of Florida libraries are also developing a union catalogue for all state university libraries with an Endeca overlay (<http://catalog.fcla.edu>), primarily to provide more user-oriented local OPACs. In the FLCA implementation the bibliographic records are not immediately available the item information (circulation

⁵ Antelman, Kirstin, Emily Lynema and Andrew Pace, *Toward a twenty-first century library catalog* p.128.

data) is retrieved in real time from the underlying Oracle data in the Aleph (ExLibris) system databases.⁶ See the FAQ at <http://www.uflib.ufl.edu/sulopac/endeca/faq.html>

Aquabrowser Library from MediaLab (<http://www.medialab.nl/>) is a third party indexing and discovery product used with many public and some academic library OPACs. The indexing engine is separate to the library system and catalogue data must be exported daily into the AquaBrowser engine. Features include tag maps and a number of facets for filtering search results. AquaBrowser provides a graphical representation of search results, which are displayed contextually by topic or location. Users can select suggested terms for further searching. Aquabrowser also indexes non-MARC data such as federated search results from electronic resources. TLC, the Library Corporation, has partnered with MediaLab to resell AquaBrowser in North America and some Asia-Pacific countries.

Queens Library in Queensborough, New York uses Aquabrowser with its online catalogue <http://aqua.queenslibrary.org/>, see figure 2. Discovery options on the left of the results screen offer related search terms, translations and spelling variants. Facet options to refine searches on the right include Format, Author, Subject, Language, Series, Geographical, Person, and Year. Sort and limit by location are drop down options.

Encore is the OPAC 2.0 product from Innovative Interfaces Inc., vendor of the Millennium integrated library system (http://www.iii.com/encore/main_index2.html). Encore operates on top of the online catalogue but is integrated with the ILS and does not require separate data export or re-indexing. It provides relevance ranking, faceted searching by format, language and collections, user tagging, and "best bet" suggested resources as well as dynamically generated search suggestions. Encore is due for release later in 2007.

Primo, the OPAC 2.0 product from the ILS vendor Ex Libris (<http://www.exlibrisgroup.com/primo.htm>), integrates with existing library environments and can be used with back-end software product from Ex Libris and other library system vendors. The user experience or presentation layer is separated from the back-end system infrastructure. Primo optimises discovery and delivery across all institutional resources.

The National Library of Australia has developed a prototype OPAC 2.0, Library Labs (<http://li01.nla.gov.au>), for potential use with Libraries Australia, see figure 3. Library Labs uses the open source search engine Lucene to index 16 million bibliographic records with holdings information from a copy of the Australian National Bibliographic Database (ANBD) as at January 2006. It extracts topics and relationships from records retrieved from a full text search and presents search results that are:

- relevance ranked
- clustered by date, subject (hierarchy and facet), form, genre, author, Dewey & LC classification
- augmented with suggestions based on subject authority matching

⁶ Bennett, Rich. Post to NGC4LIB listserv, May 5, 2007.

- with spelling suggestions from the database

Record details are shown with:

- related records from OCLC's XISBN service
- data including ratings and reviews from Amazon
- for selected holding libraries, deep links into their OPAC

The clustering currently implemented in the Library Labs prototype is based on metadata culled primarily from the subject entries, the leader and 008 fields in a MARC record. It offers users the opportunity to refine search results by subject, by Conspectus discipline, Dewey Decimal Classification, and by the format, literary form or nature of a work. LibraryLabs is Opensearch compliant so that it can be searched through other web interfaces and NLA is planning to build Z39.50 and SRW/SRU gateways to Lucene.⁷

OCLC offers WorldCat.org local, a discovery and delivery interface for use by a library or library consortium. It allows libraries to combine the metadata of OCLC member libraries worldwide and customise WorldCat.org for local discovery and delivery. It features faceted browsing, cover art, and a multilingual interface, as well as customised relevancy ranking. Libraries can apply their own local branding and information. The University of Washington Libraries have implemented the first public beta version of WorldCat Local (<http://uwashington.worldcat.org> or from the link on <http://www.lib.washington.edu>). The search returns results from the University of Washington Libraries, the Summit consortium libraries (most academic libraries in Washington and Oregon), the entire WorldCat database and a selection of article citations from ERIC, Medline, GPO and ArticleFirst (see figure 4).

The screenshot shows the University of Washington Libraries website interface. At the top, there is a navigation bar with links for 'Off-Campus Access', 'Ask Us', 'Your Library Account', and 'Survey - Your Voice Counts!'. Below this is a search bar containing the text 'digital libraries' and a 'Search' button. The search results are displayed in a list format, sorted by 'Relevance'. The first result is 'Digital libraries' by William Y Arms; NetLibrary, Inc. The second result is 'DIGITAL LIBRARIES - Mass Digitization Projects' by Roy Tennant. The third result is 'INFOTECH - DIGITAL LIBRARIES - Evolving the Resolving' by Roy Tennant. On the left side of the search results, there is a 'Refine Your Search' section with various filters such as 'Author', 'Content', and 'Library Science, Gen...'. The page also includes a 'Home' link and a 'Learn More' link.

⁷ Boston, Tony and Alison Delitt. Relevance ranking of results from MARC-based catalogues: from guidelines to implementation exploiting structured metadata. Information. Online, Sydney January 2007 <http://il01.nla.gov.au/InfoOnline2007Paper.doc>

WPopac is an online catalogue search interface developed inside the framework of WordPress blog management software. Casey Bisson of the Lamson Library, Plymouth State University developed the WPopac as an OPAC 2.0 testbed to experiment with formatting and display of bibliographic data as well as user interactivity. WordPress plugins are written to use and display catalogue data anywhere on the page.⁸ The software takes advantage of some of the unique MARC data and extracts real-time holdings information from the online catalogue. This requires export of data from the library management system for indexing within the WPopac software (<http://www.plymouth.edu/library/opac/>, see figure 5). Casey Bisson recently received the Mellon Award for Technology Collaboration for his OPAC 2.0 project, now called Scriblio. Cook Memorial Library in Tamworth, New Hampshire (<http://tamworthlibrary.org/>) is also testing Scriblio. Its OPAC offers Browse and Search options as well as sharing Flickr photo, tags and blogs.

The screenshot shows the Lamson Library OPAC interface. At the top, it says 'Lamson Library catalog of books and more (under development)' and 'Plymouth State UNIVERSITY'. Below the header, there are navigation links: 'Library Home', 'What is this?', and 'Report a bug'. A search bar contains the text 'digital libraries' and a 'Search' button. The main content area shows 'Searched for:' with the keyword 'digital libraries'. Under 'Search Results:', there is a book entry: 'Exploring The Digital Library : A Guide For Online Teaching And Learning' by Kay Johnson and Elaine Magusin, published in 2005. The entry includes a URL, subject keys, and a small image of the book cover. On the right side, there is a 'Refine Your Search' section with a 'subjkey' list containing various categories and their counts, such as 'United States (51)', 'Academic Libraries (46)', and 'Libraries (41)'. There is also an 'author' section with 'Association Of Research Libraries, Systems And'.

John Blyberg at the Ann Arbor District Library has developed SOPAC, a “social OPAC” a set of social networking tools integrated into the AADL catalogue (<http://www.aadl.org>) which uses Innovative Interface’s integrated library system. It gives users the ability to rate, review, comment-on, and tag items.⁹ SOPAC is run through Drupal, an open source content management system. SOPAC does not index data outside of the Innovative system. Drupal acts as a proxy agent, querying the OPAC and reformatting the returned data.

Sebastian Chan and team at the Powerhouse Museum in Sydney have developed an OPAC 2.0 catalogue inhouse

⁸ Bisson, Casey, maisonbisson, WPopac: an OPAC 2.0 testbed , February 9, 2006 <http://maisonbisson.com/blog/post/11133/>

⁹ Blyberg, John <http://www.blyberg.net/2007/01/21/aadlorg-goes-social/> January 21, 2007

(<http://www.powerhousemuseum.com/collection/database/>), using collection records from the museum's Emu collection database (see figure 6). The OPAC 2.0 software recommends similar search terms to increase discovery and encourage browsing of the collection. It tracks searches and ranks search results based on user interactions, and users can tag objects for later access¹⁰.

More resources

To find out about the latest OPAC 2.0 developments join NGC4lib (<http://dewey.library.nd.edu/mailling-lists/ngc4lib/>), the Next Generation Catalogs for Libraries mailing list started by Eric Lease Morgan. The purpose of the listserv is to discuss topics including:

- Who are the primary intended audiences for a library's "card catalog"?
- Considering the changing nature of information access in an Internet environment, how is an electronic "card catalog" of today different from the one designed ten or fifteen years ago?
- What kind of content should these "card catalogs" contain?
- To what degree are these things "catalogs" (as in inventory lists), and to what degree are they finding aids?
- To what degree should traditional cataloging practices be used in such a thing, or to what degree should new and upcoming practices such as FRBR be exploited?
- How would such a thing get created and by whom?

¹⁰ Chan, Sebastian. PowerHouse Museum launches Web 2.0 styled collection search, June 8, 2006. <http://www.powerhousemuseum.com/dmsblog/index.php/2006/06/08/powerhouse-museum-launches-web-20-styled-collection-search/>

- What are some of the functionalities of "next generation" catalog?

Buy or build?

Two paths for OPAC 2.0 development exist currently - commercial and open source third party products added to online catalogues, and integrated library system vendors building their own or partnering with third party products. Do you build your own using third party products or should you buy through your ILS or another vendor? Factors to consider will be available budget and inhouse technological expertise.

Purchasing an existing OPAC 2.0 product through an ILS vendor or a third party vendor carries a price tag as well as some local staff time for configuration and customisation. Do-it-yourself developments using third party products involve the software cost (unless it's open source) and investment in local staff expertise and development time.

OPAC 2.0 provides opportunities to vitalise the online catalogue search experience and is something that all libraries will need to investigate. Join the NGC4LIB listserv, read about current developments, and talk to other libraries and colleagues about ideas and plans. Partnering with similar or neighbouring libraries may be an economical way to share costs and expertise.