The current issue and full text archive of this journal is available at
www.emeraldinsight.com/0737-8831.htm

Received 16 April 2010
Revised 28 May 2010
Accepted 9 July 2010

OTHER ARTICLES

Evaluating and comparing discovery tools: how close are we towards next generation catalog?

Sharon Q. Yang
Rider University Libraries, Lawrenceville, New Jersey, USA, and
Kurt Wagner
William Paterson University, David and Lorraine Cheng Library, Wayne, New Jersey, USA

Abstract

Purpose – The purpose of this paper is to evaluate and compare open source and proprietary discovery tools and find out how much discovery tools have achieved towards becoming the next generation catalog.

Design/methodology/approach – The paper summarizes characteristics of the next generation catalog into a check-list of 12 features. This list was checked against each of seven open source and ten proprietary discovery tools to determine if those features were present or absent in those tools.

Findings – Discovery tools have many next generation catalog features, but only a few can be called real next generation catalogs. Federated searching and relevancy based on circulation statistics are the two areas that both open source and proprietary discovery tools are missing. Open source discovery tools seem to be bolder and more innovative than proprietary tools in embracing advanced features of the next generation catalog. Vendors of discovery tools may need to quicken their steps in catching up.

Originality/value – It is the first evaluation and comparison of open source and proprietary discovery tools on a large scale. It will provide information as to exactly where discovery tools stand in light of the much desired next generation catalog.

Keywords Online cataloguing, Libraries, User interfaces, Open systems, Function evaluation

Paper type Research paper

1. Introduction

After all, you can put lipstick on a pig, but it's still very much a pig (Tennant, 2005).

This rhetorical expression is in wide use to describe changes that are superficial but do not change anything fundamental about the subject. Roy Tennant (quoting Andrew Pace) used this as a metaphor for attempts to improve the library catalog user interface in ways that improve the initial look and feel, but that leave the underlying mechanism (and its inherent shortcomings) untouched. The changes in the library OPAC marketplace described by Marshall Breeding in his Library Technology Reports (Breeding, 2007) document the rise from obscurity of a set of open-source, standalone search interfaces that can be installed on top of a vendor-supplied integrated library system (ILS). Without going through the complexity and expense of an ILS migration, a library can implement an open-source, standalone OPAC and gain the advantages of

Library Hi Tech
Vol. 28 No. 4, 2010
pp. 690-709

0737-8831
DOI 10.1108/07378831011096312

a next-generation interface. It is true that the data will still retain any problems arising from the system from which it comes, but the user experience is drastically improved. Indeed, the pig has received an extensive facelift. This article will discuss the extant literature that evaluates next-generation library interfaces, present the features that define such an interface, review 17 user interfaces comparing open-source and proprietary standalones, present a comparison of features, and conclude with some recommendations to those who wish to implement an alternative to their current OPAC.

A discovery tool is often referred to as a stand-alone OPAC, a discovery layer, a discovery layer interface, an OPAC replacement, or the next generation catalog (NGC). Unlike the front end of an integrated library system or ILS OPAC, a discovery tool is defined as a third party component whose purpose is to "provide search and discovery functionality and may include features such as relevance ranking, spell checking, tagging, enhanced content, search facets" (OLE Project, 2009). Discovery tools should not be confused with federated search products. The former "promise to provide a single interface to multiple resources based on using a centralized consolidated index to provide faster and better search results", while the latter search remotely, rely on connectors, and provide "only partial and limited solutions" (Hane, 2009). In addition, a federated search tool usually requires user logon and works in a protected environment, while a discovery layer is open to the public. A federated search tool is dedicated to finding articles across a number of subscribed databases and as such is not within the scope of this paper. Libraries are disappointed with commercial ILS OPACs. Developed as a part of an integrated library system, they have remained relatively static over the years and have not evolved in pace with the discovery and search tools now commonplace at commercial sites such as Amazon. Most of them cannot and will never be able to provide advanced functionalities in order to meet current expectations. It is more practical for vendors and developers to field new OPAC systems that run alongside the older ones than to attempt to alter the proprietary code of ILS OPACs. Most current ILS OPACs do not offer the features of these standalone, next generation catalogs.

Until recently, libraries could do nothing about their outdated OPAC. Proprietary ILS OPACs offered only limited customization. Today, libraries using some of the ILS OPACS can add patches and a limited number of functional improvements by acquiring both free and commercially available plug-ins or add-on modules, but this solution will not completely transform an old OPAC into a next generation catalog. Additionally, libraries may adopt a "Web OPAC wrapper" solution to embed their existing OPAC within another user interface layer (Murray, 2008). The current trend some libraries seem to favor is to simply abandon their current OPAC in favor of one of the new standalone, next-generation discovery tools.

Interfaces may be proprietary or open source. This paper will evaluate both open source and proprietary discovery tools using 12 attributes of next generation catalogs as outlined by Breeding (2007) and Murray (2008). We present a feature-by-feature comparison of the selected interfaces ranked on the number of next generation catalog features found in each system. Today's libraries are faced with a do-or-die proposition: compete successfully with the Amazon/Google interfaces, or be replace by them. By making search interfaces more competitive, feature-rich, social and similar to interfaces found on popular web sites, we are now able to see that we indeed can offer

our users the ability to search, discover, and find in setting comparable to commercial sites.

2. Literature review

A literature review yielded two published studies and one quasi-study that are similar in design to the one described in this paper. The first study was done by two academic librarians in Slovenia, investigating how library catalogs "have tackled the mission of becoming the 'next generation catalogue'" and compared them to Amazon (Murcun and Zumer, 2008). The second study was carried out by two library school faculty members in New Zealand, comparing 22 next generation catalog features on a checklist cross the OPACs in 13 New Zealand academic institutions (Luong and Liew, 2009). The third publication is more descriptive in nature and involves evaluation of folksomonies and tagging in OPACs and discovery layers of four academic institutions in the USA. Additionally, a guest columnist in [journal title] presented a list of "nextgen" catalog attributes and summarized some of the desirable attributes of an evolved library catalog interface.

In an expert study in 2008, Mercun and Zummer evaluated six library catalogs: the Slovene union catalogue, Ann Arbor District Library catalogue, Hennepin County Library catalogue, Queens Library catalogues, Phoenix Public Library catalogues, and WorldCat and compared them to Amazon, "which is perceived both as a competitor and a model of an innovative tool" (Murcun and Zumer, 2008). The next generation catalog features used in comparison included search, results page and navigation, enriched content and recommended lists, user participation, user profile and personalization, and other Web 2.0 trends such as RSS feeds, blogs, and instant messaging. They concluded that "none of the catalogues offer as vast a range of features as Amazon does". Their findings offered some insight into current OPACs when compared with next generation catalog.

In a published study in 2009, Luong and Liew (2009) analyzed the OPACs of 13 New Zealand academic libraries against a checklist of 22 advanced features. OPACs of six integrated library systems were chosen in the sample. A comparison was made as to "how libraries using the same integrated library were customizing their interfaces to make them useful to their users" (Luong and Liew, 2009). The features used in comparison are "faceted narrow ability, visual mapping, most-popular ranking, user annotation/comment" as well as more traditional OPAC functionalities such as search types, capability, display, text, layout, and user assistance. The findings indicate that while library OPACs scored high in traditional areas, new features such as tagging, faceted navigation, ranking, and related items are not present.

In a 2009 article and quasi-study, Webb and Nero (2009) evaluated tagging and folksomonies in the OPACs of four academic institutions in the USA: LibraryThing of San Francisco State University Library, Penntags of University of Pennsylvania, Encore of St Lawrence University Libraries, and Aquabrowser of Harvard University Libraries (Webb and Nero, 2009). They observed more value in implementing discovery layers in comparison to ILS OPACs.

In her article "Next generation catalogs: what do they do and why should we care?" Emanuel (2009) characterizes the "nextgen" catalog as having a simpler user interface screen, pulling data from outside sources and including information submitted by users. Overall, Emanuel (2009) says that the next-generation catalog is built to support

the way our users search: entering keywords and then applying limits to the results, rather than a librarian-type search with complex syntax or specific, controlled search language.

While the research by Murcun and Zumer (2008) truly measured the presence of next-generation features in library OPACs, the scope of their study did not include standalone discovery tools. The same can be said about the findings, by Luong and Liew, whose research centered on ILS OPACs. Webb and Nero (2009) included discovery tools such as Encore and Aquabrowser in their observations, but did not focus on the characteristics associated with next generation catalogs. Emanuel (2009) does present the case for the standalone discovery interface implemented alongside an existing ILS and begins to describe desired characteristics, but stops short of an exhaustive comparison of available products. Our literature review did not reveal any research that compared open source and proprietary discovery tools and evaluated progress made by each towards the next generation catalog at the time of this paper's preparation. Therefore the study described in this paper is unique and the first to investigate the development of open source discovery tools versus commercial ones.

3. Investigative procedures

A. Purpose and procedures

The purpose of this study is to evaluate standalone, open source library user interfaces to highlight their developmental progress and adoption of next-generation attributes. This study presents a comparison of open source and proprietary interfaces. Each example being evaluated is ranked based on the number of next-generation features it has. A detailed discussion follows about strengths and limitations of current discovery tools.

The first step in the study involves the compilation of a list of features agreed on by consensus in the library world that the next generation catalog. This list will serve as a checklist for measurement of the presence or absence of next-generation features in the discovery tools. Next, all the major open source and commercial discovery tools were inventoried. For each discovery tool, up to three examples of implementation of the system were selected for examination. When a system is a new release and no implementation sites were identified, a developer's demonstration was used. Some discovery tools were excluded from this study because either they were still under development or no implementations or demonstrations were available for review (e.g. Extensible Catalog and EBSCO Discovery Service). Also excluded from this study were federated search tools such as 360 Search, WebFeat, and Integrated Search. These three products are not library catalogs and only search federated content and are therefore out of our inquiry scope. The final step was to compare each example to the checklist of features and signify the presence or absence of each feature. The findings were tabulated. The conclusion contains a comparison of open source versus proprietary discovery layers.

B. A check-list

We compiled a list of commonly acknowledged features for next-generation catalogs found in the library literature and summarized in Marshall Breeding's Introduction in Library Technology Reports (Breeding, 2007) and Peter Murray's PowerPoint presentation on OPAC discovery layer tools (Murray, 2008).

Discovery tool evaluation check-list:

(1) *Single point of entry for all library information.* The library catalog should be a single search or federated search for all library materials, including pointers to the articles in electronic databases as well as records of books and digital collections. One search should retrieve all relevant materials. Presently, patrons have to search the catalog for books and videos, databases for journal articles, and digital collections and archives for local images and materials.

(2) *State-of-the-art web interface.* Library catalogs should have a modern design similar to commercial, e-business sites. This criterion is highly subjective and as such is difficult to quantify. A next-generation catalog should look and feel like popular sites such as Google, Netflix and Amazon.

(3) *Enriched content.* Library catalogs should include book cover images, user driven input such as comments, descriptions, ratings, and tag clouds. Traditionally, only professionally trained cataloging librarians have the ability to create or add content to bibliographical records.

(4) *Faceted navigation.* Library catalogs should be able to display the search results as sets of categories based on some criterion such as dates, languages, availability, formats, locations, etc. Users can conduct a very simple, initial search by their preferred keyword method and then refine their results by clicking on the various results facets.

(5) *Simple keyword search box on every page.* The next generation catalog starts with a simple keyword search box that looks like that of Google or Amazon. A link to advanced search should be provided. The simple search box should appear on every page of the interface as users navigate and conduct searches. Though this feature is considered to be one of the important characteristics in a next-generation catalog, in reality it is not implemented widely. Our survey of sites shows that most libraries do not offer a simple keyword search box as a default start page. Librarians prefer an advanced search and feel that the quick search is more likely to produce results with less precision.

(6) *Relevancy.* Librarians complain that OPAC relevancy results are problematic or that they do not undersand how relevance is determined. The next-generation catalog does better in relevancy ranking with increased precision. In addition circulation statistics should influence the relevancy results. More frequently circulated books indicate popularity and usefulness. They should be ranked higher in the display. Items deemed important enough to have multiple copies should also receive higher relevancy ranking.

(7) *Did you mean ….? A* spell-checking mechanism should be present in a next-generation catalog. When an error appears in the search, there should be a pop-up with the correct spelling or suggestions from a dictionary. Clicking on any of these runs a search.

(8) *Recommendations/related materials.* Commonplace in e-commerce sites, the customer is shown additional items with a suggestion like "Customers who bought this item also bought…" Likewise, a next-generation catalog should recommend books for readers on transaction logs. This should take the form of

"Readers who borrowed this book also borrowed the following . . . " or a link to "Recommended Readings".

(9) *User contribution.* The next-generation catalog allows users to add data to records. The user input includes descriptions, summaries, reviews, criticism, comments, rating and ranking, and tagging or folksonomies. Today's users increasingly look for what other users have to say about items found online, and value what they feel to be their peers' review of items. Tagging clouds can serve as access points and descriptive keywords leading to frequently used items.

(10) *RSS feeds.* Really Simple Syndication allows users to connect themselves to content that is often updated. Next-generation interfaces include RSS feeds so that users can have new book lists, top-circulating book lists, canned searches, and "watch this topic" connections to the catalog on their own blog or feed reader page.

(11) *Integration with social network sites.* When a library's catalog is integrated with social network sites, patrons can share links to library items with their friends on social networks like Twitter, Facebook and Delicious.

(12) *Persistent links.* Next-generation catalog records contain a stable URL capable of being copied and pasted and serving as a permanent link to that record.

C. Open source and proprietary discovery tools

This study included major open source and proprietary discovery tools that authors could identify at the time of writing. Sharon Yang and Kurt Wagner's presentation on open source discovery tools at the Virtual Academic Library Environment (VALE) 2010 Annual Conference was used to identify these products (Yang and Wagner, 2010). Discovery Layer Interfaces in Library Technology Guides by Marshall Breeding (Breeding, 2009) provided confirmation that all relevant products were included. Federated search services such as 360 Search and WebFeat by Serials Solutions, and Integrated Search by EBSCO were not included in this paper as they are not considered to be discovery layers. For each discovery tool, up to three library implementations were used in data collection depending on availability of installations. Generally, the client list could be found from the product's web page. We found that in the case of new products, a live implementation could not always be found. In these cases a demonstration site was used to compile data. Open source discovery tools are considered separately from commercial, proprietary products for the simple reason that the former can be freely implemented, customized and used. They require some local programming and configuration to enable them to search and display data from a traditional ILS. These open source products do not require any sort of contract, or support, as is the case with proprietary systems. The second list is for evolved, next-generation interfaces offered by commercial ILS or interface vendors. The following are two alphabetical lists of sites, one for open source and one for proprietary discovery tools reviewed in this study:

Library sites using open source discovery tools

(1) *Blacklight*

- Stanford University http://searchworks.stanford.edu/

- University of Virginia http://virgobeta.lib.virginia.edu/
- North Carolina University http://historicalstate.lib.ncsu.edu/

(2) *Fac-Back-OPAC (Kochief)*
- Paul Smith's College Book Catalog http://library.paulsmiths.edu/catalog/
- Drexel Libraries collections http://sets.library.drexel.edu/

(3) *LibraryFind*
- Deschutes public library www.dpls.lib.or.us/
- Oregon state University http://osulibrary.oregonstate.edu/

(4) *Rapi*
- Demo by School of Computing, National University of Singapore http://linc.comp.nus.edu.sg/

(5) *Scriblio (WPopac)*
- Plymouth State University http://library.plymouth.edu/
- Cook Memorial Public Library http://tamworthlibrary.org/
- Hong Kong University of Science and Technology http://catalog.ust.hk/catalog/smartcat.php

(6) *SOPAC (Social Opac)*
- Ann Arbor District Library www.aadl.org/catalog
- Allen County Public Library www.acpl.lib.in.us/
- Darien Library www.darienlibrary.org/

(7) *VuFind*
- Colorado State University Libraries http://discovery.library.colostate.edu/
- Yale University http://yufind.library.yale.edu/yufind/
- University of Michigan http://mirlyn.lib.umich.edu/

Library sites for proprietary discovery tools

(1) Aquabrowser by Serials Solutions

(2) Harvard University: http://discovery.lib.harvard.edu/

(3) Queens Library: http://aqua.queenslibrary.org/

(4) Oklahoma State University: www.library.okstate.edu/

(5) BiblioCommons
- Halton Hills Public Library: http://hhpl.bibliocommons.com/dashboard
- Oakville Public Library www.opl.on.ca/
- West Perth Public Library: http://wppl.bibliocommons.com/dashboard

(6) Encore-Innovative Interfaces Inc.
- St Lawrence University: www.stlawu.edu/library/
- Syracuse University: http://library.syr.edu/find/
- University of Houston: http://info.lib.uh.edu/

(7) *Endeca-Endeca*
- North Carolina State University: www.lib.ncsu.edu/endeca/
- McMaster University: http://library.mcmaster.ca/
- University of Central Florida: http://ucf.catalog.fcla.edu/cf.jsp

(8) *One Search: Follett* (hosted and require login)
- Follett: http://onesearch.fsc.follett.com/onesearch/
- Pima Public Library: http://onesearch.fsc.follett.com/FIACollection/?custnum = 0200947000&searchterm = &remoteapp = OneSearch.dll&screenclass = com.follett.fiacollection.screens.FirstScreen&Command = Search

(9) *Primo-Ex Libris*
- Vanderbilt University: www.library.vanderbilt.edu/
- University of Iowa: www.lib.uiowa.edu/
- Emory University: http://web.library.emory.edu/

(10) *SirsiDynix Enterprise-SirsiDynix*
- Warren County Library: www.warrenlib.com/ (call to confirm)
- Fort MacLeod RCMP Centennial Library: www.chinookarch.ab.ca/client/hq
- Caroline County Public Library: www.caro.lib.md.us/library/

(11) *Summon by Serials Solutions* (now Proquest)
- Dartmouth College Libraries: http://library.dartmouth.edu/
- University of Calgary: http://library.ucalgary.ca/
- University of Sydney: www.library.usyd.edu.au/

(12) *Visualizer-VTLS*
- Demo-Networked Digital Library of Thesis and Dissertations: http://thumper.vtls.com:6080/visualizer/
- Demo: http://thumper.vtls.com:7080/visualizer/
- Upper Arlington Public Library: www.ualibrary.org/index.php

(13) *WorldCat Local-OCLC*
- University of Connecticut: http://uconn.worldcat.org/
- Indiana University: www.indiana.edu/~kolibry/worldcatlocalfaq.shtml
- SUNY: http://sunysccc.worldcat.org/ca/ http://library.ucalgary.ca/ooooniversity of Calgary Lib

D. Data collection

Each of the 12 next-generation catalog attributes discussed in Section B, was checked against the sites in Section C. Features were marked "present" (✓) when they were seen at least once in a production or demonstration installation, otherwise, the feature was marked "absent" (x)We were careful not to rely solely on the product web sites for confirmation of the presence of a feature. Given the nature of open-source applications, where functionality may be feasible yet not actually implemented, we went to the

production sites wherever possible to confirm our findings, which are recorded in Tables I and II.

4. Evaluation and comparison

A. Evaluation

A single point of entry for all library resources: Federated search is the holy grail of discovery layers. "The pursuit of a Discovery Layer seem to be driven by the need to present one, strong and stable user interface over many disparate sources of information" (Williams, 2008).Without this capability, a discovery tool can be hardly considered complete. While many discovery tools indicate on their web sites that federated search is an integral part of the package, a reality check shows that most discovery tools covered by this study are not performing federated search except Summon and LibraryFind. Some discovery tools give the false impression of a unified interface by adding a tab on the top menu bar for databases and other resources, but in reality a user has to search the catalog, databases, and digital resources separately. Encore performs a pseudo federated search by a button called "Results from Article Databases". Clicking on this button presumably will lead users to a login and execution of the same search across the databases.

The reason why most discovery tools in live examples do not include all library resources is not clear, nor is it within the scope of this paper. Conventional federated search engines such as 360 Search, WebFeat, and EBSCO Integrated Search use connectors (software programs) to individual databases, while discovery tools use a different approach by extracting data and building indexes to resources. As no uniform standards exist for these disparate resources, it is hard to develop a search mechanism dealing with resources that are vastly different in design. Like federated search engines, discovery tools may have to negotiate with database vendors to build pointers or keyword indexes to databases. Is it possible that different discovery tools cover a limited number of different databases as federated search interfaces do today? Federated search tools can hardly serve as OPACs. They require authentication and only operate in a protected environment. Most lack the advanced features of the next-generation catalog. The following is the ranking of discovery tools based on federated searching capability:

(1) LibraryFind and Summon.

(2) Encore.

(3) Rest of the discovery tools.

State-of-the-art interface: Most discovery tools in this study have attractive user interfaces. Most have faceted navigation on one side and colorful book cover images and tags on display. Therefore most of the discovery tools received endorsement in this category except Rapi and Scriblio. Figure 1 is a screen shot from Encore, which is, admittedly, proprietary, but leads the group of this category of next-generation interfaces.

Rapi has a very basic, simple user interface with text only display (see Figure 2). It does not possess the color and design of a modern OPAC. Scriblio is built on the WordPress blog platform and has a highly customizable user interface. Scriblio often serves as the base structure of a web site with searching capability and blends into the rest of the environment rather than as a distinctive discovery layer. When compared

Table I. Open source discovery tools

	BlackLight	Fac-Back OPAC	LibraryFind	Rapi	Scriblio (WPOPAC)	Sopac (Social OPAC)	VuFind
Single point of entry	x	x	✓	x	x	x	x
State-of-the-art web interface	✓	✓	✓	x	x	✓	✓
Enriched content	✓	✓	✓	x	✓	✓	✓
Faceted navigation	✓	✓	✓	x	✓	✓	✓
A simple box of keyword search with a link to advanced search	✓	✓	✓	✓	✓	✓	✓
Relevancy	x	x	x	x	x	x	x
Did you mean . . .?	✓	x	✓	x	✓	x	✓
Recommendations	x	x	✓	x	✓	✓	✓
User contributions	x	x	✓	x	✓	✓	✓
RSS feed	✓	x	✓	x	✓	x	✓
Integration with social networking sites	x	x	x	x	✓	x	✓
Persistent link	x	✓	✓	x	✓	x	x
Total NGC features	6	5	10	1	9	6	9

Table II. Proprietary discovery tools

	Proprietary tools				
	Aquabrowser	BiblioCommons	Encore	Endeca	One Search
Single point of entry	x	x	x ✓	x	✓
State-of-the-art web interface	✓	✓	✓	✓	✓
Enriched content	✓	✓	✓	✓	x
Faceted navigation	✓	✓	✓	✓	x
A simple keyword search box	✓	x	✓	x	x
Relevancy	x	x	x	x	x
Did you mean …?	✓	✓	✓	✓	x
Recommendations	x	✓	✓	✓	x
User contributions	x	✓	✓	x	x
RSS feed	x	x	x	x	x
Integration with social networking sites	x	✓	x	x	x
Persistent link	✓	x	x	✓	x
Total NGC features	6	7	7.5	6	2
	Primo	SirsiDynix Enterprise	Summon	Visualizer	WorldCat Local
Single point of entry	x	x	✓	x	x
State-of-the-art web interface	✓	✓	✓	✓	✓
Enriched content	✓	✓	✓	✓	✓
Faceted navigation	✓	✓	✓	✓	✓
A simple box of keyword search with a link to advanced search	x	x	✓	✓	✓
Relevancy	x	x	x	x	x
Did you mean …?	✓	x	✓	x	✓
Recommendations	✓	✓	x	x	✓
User contributions	✓	x	x	x	✓
RSS feed	✓	x	x	✓	x
Integration with social networking sites	✓	x	x	x	x
Persistent link	x	x	✓	x	x
Total number of NGC features	8	4	7	5	7

Figure 1.
State of the art user interface-Encore

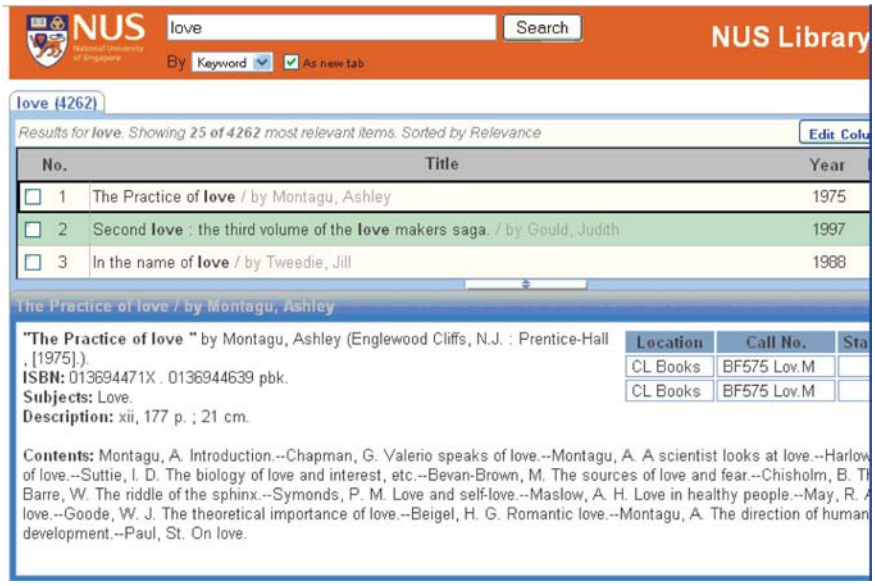

Figure 2.
Rapi

with other discovery tools, Scriblio is visibly different in its way of displaying contents even though it has more of the next-generation features than many other discovery tools (see Figure 3).

Enriched contents and user contributions: Almost all the discovery tools provide cover images, but not every discovery tool allows users to contribute and share data.

Figure 3.
Scriblio

Traditionally only cataloguers are authorized to edit and maintain integrity of bibliographical database in an OPAC. It is a revolutionary step and new concept to allow user contribution to library records. Eight out of 18 discovery tools reviewed in this study have this feature. The following is a list of discovery tools that allow user contribution to enrich the catalog contents in varying degrees:

(1) BiblioCommons (8): Tags, comments, summaries, quotes, notices, age, videos, rating

(2) LibraryFind (3): Tags, reviews, rating

(3) Primo (2): Tags, reviews

(4) Scriblio (2): Tag, comments

(5) Sopac (2): Tags, rating, reviews

(6) VuFind (2): Tags, comments

(7) WorldCat Local (2): Tags, reviews

(8) Encore (1): Tags.

Faceted navigation: Faceted navigation is a standard feature in all the discovery tools covered in this paper except Rapi and One Search by Follett. Faceted navigation is very creatively described and integrated across the products we evaluated. The number of facets and the way they function vary widely, but the most commonly seen configurations are:

(1) Access/library/location/collection

(2) Author/creator

(3) Availability/available
(4) Call number/classification LCC/Dewey Range
(5) Content/content type
(6) Format/type/material type/resource type/form/genre
(7) Journal title
(8) Keyword
(9) Language
(10) Organization as (author)/provider/corporate author
(11) Publication year/publication date/published date/publish date/creation date/decades
(12) Publisher
(13) Region/geographic/continent/place
(14) Series title
(15) Source
(16) Tag
    - By tag-genre
    - By tag-tone
    - By tag-theme
(17) Target audience
(18) Topic/subject/subject term.

Simple keyword search box with a link to advanced search: A simple keyword search box with a link to advanced search on every page of the OPAC is an attempt made by libraries to imitate Google and some other popular internet search engines. The purpose is to make sure the user always has a search box at hand wherever they go within the interface. The simple keyword box should appear at every step along the way as a user navigates through each screen. Such a keyword search box is generally referred to as "quick search" in the interfaces we evaluated. Aware of the current focus on next-generation functionality, many vendors have supplied this feature out of the box, but some libraries, skeptical of its value, refused to implement it. The quick search does not encourage precision and is thought by library instruction staff to mislead users. Many libraries replaced the quick search and choose instead to default to a basic or advanced search. Therefore, we believe libraries are at odds with their users with regards to the value of this next-generation feature.

Vendors of discovery tools in this study make this feature highly configurable. Libraries operating the same discovery tools display different searches. Some of the discovery tools maintain this simple keyword search box consistently at every step during a search. For instance, libraries running on Blacklight, Encore, and Aquabrowser start with and keep a simple keyword search box on every OPAC page (see Figure 4).

Some discovery tools start with a quick search box, but display a more complex search with a pull down menu once inside the search interface. Most libraries start with

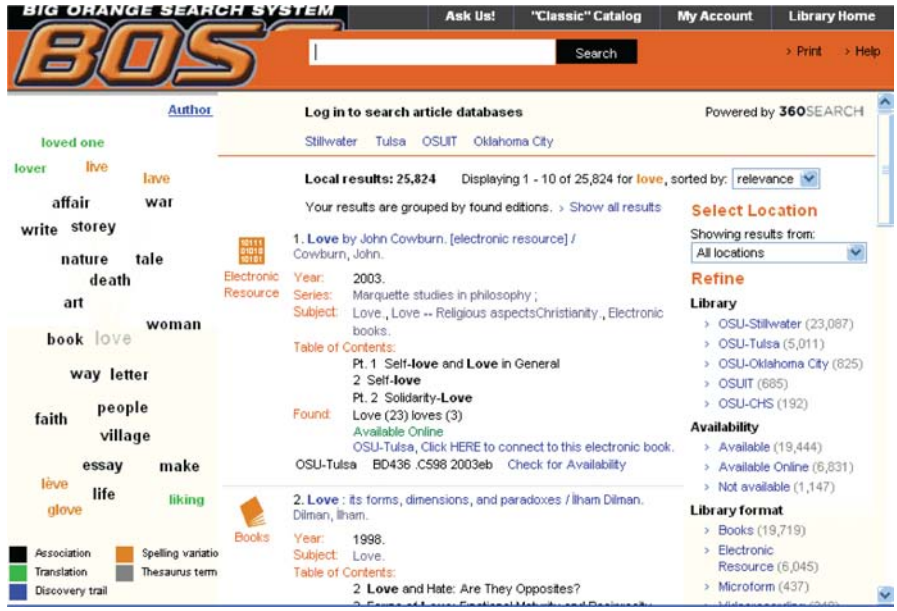

Figure 4.
An example of a simple keyword search box with a link to advanced search, Aquabrowser

and maintain on every page a complex search box with a pull down menu, offering standard search keys such as keyword, author, title, and call number (see Figure 5).

The discovery tool will receive a check (✓) indicating the presence of this next-generation feature only if at least one implementation of this discovery tool displays a simple keyword search box with a link to advanced search on every page.

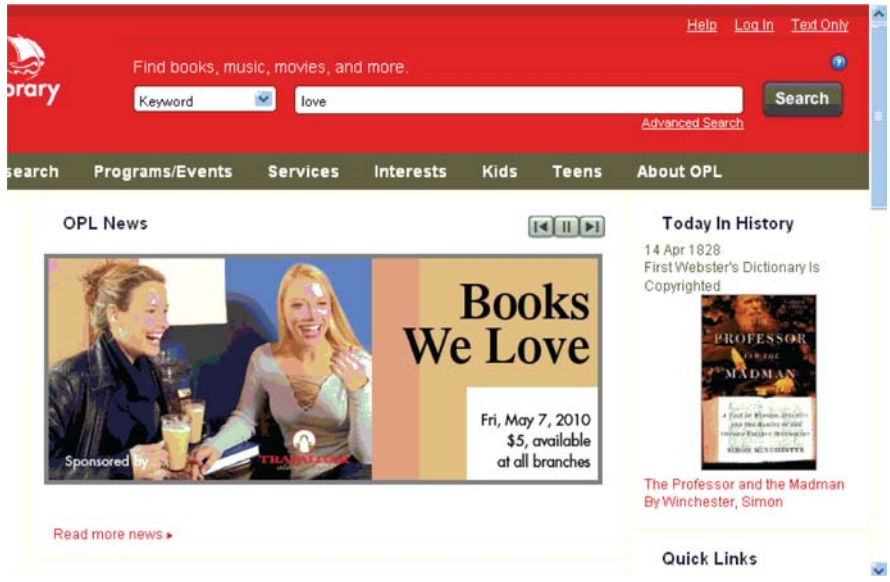

Figure 5.
A discovery tool that does not display a simple keyword search box

Relevancy: A feature that is missing from all the discovery tools is relevancy. This study did not measure relevancy in the traditional sense of the ranking algorism, but rather observed if circulation statistics played a part in search result display. So far no discovery layers have linked circulation statistics to the relevancy ranking. None of them can rank search results based on the number of times items have been circulated or the number of copies an item has. Therefore all the discovery tools received the x (x) sign indicating the absence of this feature.

Did you mean ….?: It is easy to find out if this feature is missing or present in the discovery tool by entering a misspelled word into the search box. When a system does not have this feature, a response will be displayed indicating there were no items under this term. When a system has this feature, it will display a message such as "Did you mean …?" or "Your search has found no hits. Please choose from the following terms … ". By clicking on the recommended term(s), the user will resume the search. In both cases the feature is counted as present if the discovery recommends one term or a list of terms for users to choose from.

Recommendations: A discovery tool has several ways to recommend materials to a user. Typically, a statement like: "The library patron who borrowed this item has also borrowed the following items" appears following the search results. This functionality emulates that of sites such as Amazon and libraries value this opportunity to encourage users to borrow related items. Other approaches involve links to "Similar items", or "Similar Subjects". Any form of recommendations of additional items received a check indicating the presence of this next-generation feature.

RSS feed: A characteristic orange colored icon is present in the discovery tool if provision is made to provide an RSS feed. The presence or absence of RSS functionality was noted for each discovery tool.

Integration with social networking sites: It is easy to determine the presence or absence of this feature in the interfaces we sampled. Like the RSS feed, a library can add this feature by installing the third-party-supplied coding. It is not the concern of this study to distinguish if a function is native or an add-on. Rather it is counted as present if it existed in the discovery tool at the time of this study. Otherwise it is counted as missing.

Persistent link: Sometimes a persistent link is called permanent link in a discovery tool. Generally this feature is native and comes with the discovery tool. It is counted as present or absence depending on its availability in the discovery tool.

B. Open source discovery tools

The following table (see Table I) lists all the 12 features of a next-generation catalogs in the left column, with a check (✓) or an x (x) to indicate features a discovery layer tool possessed or missed. The names of the open source discovery tools are in the top row. The findings are summarized in Table I.

The most important feature of the next generation catalog is federated searching and a single of point of entry for all library resources. Many of the open source discovery tools claimed this capability as "a single-search interface to aggregate digital content that would otherwise be siloed" (Blacklight Project Team, 2009) and as a goal to enable users "to search and browse through all of your library's resources by replacing the traditional OPAC" (Villanova University, 2010). However, federated

search is missing from the most of the sites being reviewed. LibraryFind is the only discovery tool that demonstrated federated search.

As mentioned earlier, another next-generation feature missing is "relevancy" where in no example that we examined could we determine that circulation statistics were incorporated into ranking search results.

Based on Table I, we can conclude that among open source discovery tools, LibraryFind is the discovery tool that demonstrated most of the next-generation characteristics. VuFind and Scriblio are ranked second, with nine features present and three features missing for each system. In spite of their equal ranking, VuFind is a better tool in many ways, especially its user interface being far more pleasant than Scriblio. Sopac and Blacklight each possess six out of 12 features. Rapi possesses the fewest next-generation features. Developed as a class project by students in Computing Science Department of National University of Singpore, Rapi is not widely implemented by libraries even though new releases are still coming out. The following is a ranked list of open source discovery tools based on the number of NGC features each demonstrated:

(1) LibraryFind (10).
(2) VuFind and Scriblio (9).
(3) Sopac (6) and Blacklight (6).
(4) Fac-Back-OPAC (5).
(5) Rapi (1).

C. Proprietary discovery tools

In Table II, which follows is a summary of presence and absence of next-generation features in proprietary discovery tools.

Though proprietary discovery tools claim that they are next-generation catalogs with federated search capability, this capability was only seen in Summon. While most commercial discovery tools got zero point in federated searching, Encore got 0.5 because it demonstrated a feature that was one step away from federated searching. In Encore there is a button called "Search for Journal Articles". Clicking this executes the same search in the databases. This does not live up to its billing as a single-search of all resources. As with the open source discovery tools, the commercial examples do not take circulation statistics into consideration when ranking and displaying search results. Based on the score each commercial discovery tool received, Primo ranked first. Ranked the lowest is One Search by Follett. It is not clear if One Search should be considered a federated search engine or discovery tool. It is the only system in this study that searches across library resources, but does not have its own distinctive display features:

(1) Primo (8).
(2) Encore (7.5).
(3) BiblioCommons, Summon, Worldcat Local (7).
(4) Aqubrowser and Endeca (6).
(5) Visualizer (5).
(6) SirsiDynix Enterprise (4).
(7) One Search (2).

D. Comparison

The following is the final ranking of all the discovery tools covered in this study based on the number of presence of next-generation catalog features each displays. Listed on the left are the names of the discovery tools and on the right is a numeral indicating how many of next-generation features the discovery tools have displayed out of a total of the 12 on the checklist.

The findings clearly indicate that none of the discovery tools are truly next-generation catalog if all 12 criteria must be present. Federated searching and relevancy based on transaction data are missing from all discovery tools.. It is also apparent that open source discovery tools have taken the lead in progress towards a next-generation goal. On the top of the ranking are three open source discovery tools: LibraryFind, VuFind, and Scriblio. The authors were particularly impressed with LibraryFind and VuFind, which they consider superior in many ways. Vendors of proprietary discovery tools are, naturally, more conservative in what they offer under their corporate flag:

(1) LibraryFind (10).
(2) VuFind and Scriblio (9).
(3) Primo (8).
(4) Encore (7.5).
(5) BiblioCommons, Summon, and Worldcat Local (7).
(6) Aqubrowser, Endeca, Sopac, and Blacklight (6).
(7) Fac-Back-OPAC and Visualizer (5).
(8) SirsiDynix Enterprise (4).
(9) One Search (2).
(10) Rapi (1).

5. Limitations

This study is based on real life examples or demonstrations of discovery tools. It is not based on what a discovery tool claims it can do or is capable of doing, but how libraries use them and how they perform in real life. The discovery tools may be capable of a feature by design, but due to political or technical reasons the implementation site may not have activated this feature for various reasons. Therefore this study may declare a feature missing even though the vendor may claim it is there. This study goes by what a discovery tool does and how libraries use them, not what it has the potential to do.

A discovery tool may not come with a next-generation feature, but a site may be able to add this feature through coding they create or adapt from another source. The added feature is not native for that discovery tool, but counted as present. Therefore a feature may be missing in a discovery tool, but labeled as present in this study because it was present during the review. Typical scenarios involve features such as RSS feed and integration with social network sites.

There was a subjective side to this study. Human judgment was called for on some occasions. For instance, "state of the art web interface" is purely subjective. There are no clearly expressed criteria for the purpose of this study except those in the minds of the authors and their experience in the evaluation of these tools.

The number of sites being reviewed for each discovery tool was limited and varied due to the difficulty in identifying available sites for the purpose of this study. Some discovery tools had only demonstrations from vendors. Therefore some data was collected from demonstration and some was from real life implementations of a discovery tool. There might be a difference between a demonstration site and a production site of a discovery tool. Consequently the findings may be affected.

Finally, this study attempted to cover all the major discovery tools, but there are some new ones that could not be included such as EBSCO Discovery service and Extensible Catalog. EBSCO Discovery Service is too new a release that there is no demo or live examples available for review at the time of this study. XC is not being fully developed yet and there was no demonstration or installation for review. Additionally, the distinction between federated search tools and discovery tools are increasingly blurred. One Search by Follett may be considered a federated search tool. Therefore this study is not comprehensive in scope. The authors made a careful effort to include the alternate catalog products most likely to be considered by libraries, but admit that they did not include all products that exist for this purpose.

6. Conclusion

"One might think of the term next-generation as describing something new whose development is forthcoming. Libraries seek next-generation catalogs here and now as these interfaces exist in e-commerce and we have heard our users ask why the library's interface is poor by comparison. Libraries do not necessarily have to wait" (Breeding, 2007). This study shows that the next generation catalog is becoming current generation catalog as predicted by Breeding (Breeding, 2007), but federated searching and relevancy ranking remain problematic areas that need attention by open source community and proprietary vendors. True federated searching is, and will always be, the promised land of next-generation catalog and discovery tools. A discovery tool is not complete without this federated search capability. Libraries, vendors and the open source community must continue to cooperate and work in a spirit of optimism and collegiality to make the true next-generation catalog a reality.

References

Blacklight Project Team (2009), *Blacklight*, available at: http://projectblacklight.org/ (accessed 31 March 2010).

Breeding, M. (2007), "Introduction", *Library Technology Reports*, Vol. 43 No. 4, pp. 5-14.

Breeding, M. (2009), *Library Technology Guides – Key Resources in the Field of Library Automation: Discovery Layer Interfaces*, available at: www.librarytechnology.org/discovery.pl?SID=20100321898651123 (accessed 15 March 2010).

Emanuel, J. (2009), "Next generation catalogs: what do they do and why should we care?", *Reference User Services Quarterly*, Vol. 49 No. 2, pp. 117-20.

Hane, P.J. (2009), "New discovery tools for online resources from OCLC and EBSCO", *NewsBreaks*, available at: http://newsbreaks.infotoday.com/NewsBreaks/New-Discovery-Tools-for-Online-Resources-From-OCLC-and-EBSCO-53468.asp (accessed 27 March 2010).

Luong, T.D. and Liew, C.L. (2009), "The evaluation of New Zealand academic library OPACs: a checklist approach", *The Electronic Library*, Vol. 27 No. 3, pp. 376-93.

Murcun, T. and Zumer, T. (2008), "Next generation of catalogues for the next generation of users: a comparison of six library catalogues", *Program: Electronic library and information systems*, Vol. 42 No. 3, pp. 243-61.

Murray, P. (2008), "Discovery tools and OPAC", PowerPoint presentation at NISO Forum on Next Generation Discovery Tools: New Tools, Aging Standards, available at: http://dltj.org/article/discovery-layer-video-tour/ (accessed 27 January 2010).

OLE Project (2009), *Data Dictionary*, available at: http://oleproject.org/overview/ole-reference-model/data-dictionary/ (accessed 27 March 2010).

Tennant, R. (2005), "Digital libraries: 'lipstick on a pig'", *Library Journal*, Vol. 130 No. 7, p. 34.

Villanova University (2010), *About VuFind*, available at: http://vufind.org/ (accessed 31 March 2010).

Webb, P.L. and Nero, M.D. (2009), "OPACs in the clouds", *Computers in Libraries*, Vol. 29 No. 9, pp. 18-22.

Williams, M. (2008), *New Jack Librarian*, available at: www.newjackalmanac.ca/2008/03/ontario-scholars-portal-yours-to.html (accessed 7 April 2010).

Yang, S. and Wagner, K. (2010), "Open Source Stand-Alone OPACs", Rutgers University, Piscataway, NJ, PowerPoint presented at VALE /NJ ACRL/NJLA CUS 11th Annual Users' Conference, The Future Is Now: Meeting the Needs of Our Users, Busch Campus Center, January 8, available at: www.valenj.org/vale/breakoutsection/c-open-source-standalone-opacs

About the authors

Sharon Q. Yang has been working in the library automation field for 18 years. She received her Master's degree in Science in 1988, Certificate for Advanced Librarianship in 1989, and Doctor of Library Science in 1997, all from the School of Library Service, Columbia University in New York City. While working on her doctoral degree, she started working at Wagner College as Head of Technical Services/Systems in 1990. She acquired her computer knowledge through experience at work and by taking college classes and attending workshops in computer science. Now she works as Systems Librarian and Associate Professor with Tenure in Rider University Library, New Jersey. Sharon Q. Yang is the corresponding author and can be contacted at: yangs@rider.edu

Kurt Wagner is Assistant Director and Head of Library Information Systems for the David and Lorraine Cheng Library, William Paterson University, Wayne, New Jersey, USA.