Competitive Usability and the Catalogue: A Process for Justification and Selection of a Next-Generation Catalogue or Web-Scale Discovery System

TERI OAKS GALLAWAY AND MARY FINNAN HINES

ABSTRACT
This case study demonstrates how competitive usability testing informs the selection and purchase of a next-generation catalogue (NGC) or Web-scale discovery system (WSDS) to enhance a current library catalogue. Using competitive usability techniques, the authors explain how different NGCs and WSDSs solve issues that catalogue users may face when searching for materials in the online catalogue. The goal of this study is to provide a framework that identifies concrete evidence in support of purchase recommendations for an effective system that adequately addresses locally identified issues with catalogue searches. The process of selecting live system implementations from peer institutions is outlined. Steps include surveying library staff about their current library catalogue. Survey results and documented reference questions provided the foundation for user tasks created by testers for use in this study. This multifaceted research design resulted in a case study that captures current issues that users encounter in the discovery and access to library materials and shows how to include competitive usability techniques as part of a purchase rationale while assessing how well a variety of next-generation discovery and access systems address users’ issues.

INTRODUCTION
Libraries aim to connect users to the best possible resources. In the age of Web sites like Google and Amazon, the pressure for libraries to update and improve the find-ability of library resources has increased enormously. Since the mid-2000s, a variety of next-generation catalogues (NGCs) and Web-scale discovery systems (WSDSs), both open source and vendor created, have become available as potential solutions to search problems.
Distinguishing between evolving system types is problematic, and an accepted nomenclature for these systems is still in formation. For the purposes of this article, NGCs will be distinguished from WSDSs that make use of a central index of preharvested publisher and aggregator metadata. Systems that use federated search integration are included in the category NGC.¹

Librarians have documented the process of pioneering a new system to overhaul the traditional catalogue (Antelman, Lynema, & Pace, 2006), of working with a vendor to implement a new system (Allison, 2010), and of how a specific system supports academic research in a specific subject area (Snyder, 2010). Marcin and Morris (2008) describe the comparative features of systems, but there is little in the literature that documents how a library can inform the process of recommending a system for purchase and creating a corresponding rationale for funding.

Few research studies are prescriptive, and best practices for system selection and recommendation are not well documented. Nagy (2011) outlines the stages of the deployment of NGCs as “forming, storming, norming, and performing” (p. 16). The first step in this process is forming, where library staff and users are presented with a new beta system. While this is an important step, it overlooks the need to justify the expense of adding a new system and the process of selecting a system before the forming stage.

Luther and Kelly (2011) set forth some considerations when looking at WSDSs: content, search, fit, and cost. These considerations focus on librarians’ demands to control certain aspects of these systems but do not mention usability testing to inform these decisions. While librarians have expertise in information-seeking behaviors, it is crucial to account for user needs at each stage of the process. Nagy (2011) conducts case studies of academic librarians that were involved with their library’s implementation of an NGC. One such case study involves North Carolina State University conducting usability studies before purchasing and customizing the Endeca system. Yang and Wagner (2010) create a checklist of twelve features indicative of both NGC and WSDS products and evaluate which systems offer what features. The effectiveness of those features is not assessed. Zhang, Ayres, McNair, and Williams (2007) present best practices for the beta-testing process, which may help shape the functionality of a selected system. Olson (2010) addresses the lack of research in this area and outlines several case studies in library management system selection. She finds that libraries often do not approach this decision-making process in a rational manner, and consequently, the best system is not always selected. Rather than selecting a system that solves a defined set of problems, Olson finds that a new system may be chosen to save money, to facilitate a departmental reorganization, or to improve the public perception of the library by implementing something new.

METHODOLOGY
The explosion of products in the NGC and WSDS market led librarians and library administrators at Loyola University New Orleans to carefully consider how patrons might benefit from a new system. As many peer institutions implemented these systems, a charge was developed to investigate implementations to determine if a rationale for the purchase of either an NGC or WSDS could be identified. For this reason, both types of systems were considered. Vaughan (2011) proposes that WSDSs are not primarily concerned with providing access to traditional library materials from the library catalogue and as such they should not be judged on their ability to do so. However, given competing priorities for limited library funds, a system that effectively searches all library resources is preferable. The primary questions of this study, therefore, were as follows:

• How might an NGC or WSDS enhance findability of materials from the library catalogue?
• Can materials from the library catalogue still be effectively located in a WSDS given the large number of searchable resources?

Selecting Competitive Systems
To answer these questions, competitive usability techniques are used to identify how different NGCs and WSDSs solve issues that users face when searching for materials in the online catalogue. Competitive usability techniques are described by Thomsett-Scott (2005, 2007) in the context of both the library catalogue and the library Web site. The studies involve testing a common set of tasks across a number of systems or Web sites using focus groups and task-based testing. Through this process, a case could be made to administrators and stakeholders on the benefits of a purchase of an NGC or WSDS for the library’s constituents.

Along with the local SirsiDynix online public access catalogue (OPAC), seven NGCs and WSDSs were identified at peer institutions and other Jesuit universities for competitive usability testing. The systems tested were Aquabrowser, EDS, Encore Discovery, Enterprise, Summon, Primo Central, and WorldCat Local. When possible, systems that were paired with the same underlying SirsiDynix integrated library system were selected in an attempt to reflect similar characteristics of a potential local implementation. Open-source systems were not selected due to the staffing limitations of the library and the perceived staff resources required to implement an open-source product.

Identifying Usability Tasks
To ensure that usability testing encompassed the variety of questions that arise about the current library catalogue, the research transaction log was consulted. The log documented substantive questions from students, staff,
and faculty that were directed to a reference librarian via an in-person or telephone interaction. Questions about the catalogue from the most recent six-month period were harvested, coded, and prioritized to reflect the most common user issues. Recognizing that staff have valuable anecdotal information and opinions about the performance of the library’s catalogue, information library staff were surveyed. The process of including staff feedback was beneficial as it allowed for the documentation of positive aspects of the catalogue to be retained in a new system. It also fostered a collaborative decision-making process. A summary of both staff and user issues is reported in Appendix A. For further discovery of library user issues, there are a variety of other tools that could be used, including OPAC search logs, focus groups with various stakeholders, and chat logs for virtual reference questions. Additionally, reviewing the adequacy mean in the Information Control section of the LibQUAL survey over time could show a general trend in satisfaction with systems. Through the reference transaction analysis and staff survey, tasks were developed to reflect areas of user difficulty. Turner (2002) discussed the use of task-based assessment to study both navigational design and labeling, including examples of basic tasks for catalogue studies. Fagan (2010) provides guidance on the development of usability tasks to ensure measurable outcomes and comparability across users. It is also important to reconsider the use of library jargon in usability tasks. According to Letnikova (2008), task designers should take into account some participants’ lack of experience with the library research process. Participants who are unfamiliar with library terminology may have higher failure rates. In many studies, the library Web site and catalogue are tested concurrently as it is often difficult to test them in isolation. Norlin and Winters (2002) present an integrated model of library Web site and systems testing by including tasks that make use of both the Web site and the catalogues and databases. To ensure that usability questions were answerable on the variety of systems, the questions were pretested. One student was recruited to complete a pilot test and to provide feedback. Test questions were altered to reduce the difficulty. Two additional questions related to finding journal articles were added to test the functionality of those systems that included database content.

**Usability Testing**

Students entering the library were invited to participate in usability testing. They were asked if they had approximately twenty minutes to help the library. They were offered candy bars, library mugs, and coffee as an incentive. Four to six students were recruited to test each system under consideration. In Thomsett-Scott’s (2007) catalogue study, each participant tested multiple systems. The study found that testers develop exper-
tise throughout the process so that systems tested last may perform better
than those tested at the outset. In order to avoid the tester learning curve
and to prevent tester fatigue, each participant tested only one system.

Multiple laptops (both Windows and Macintosh) equipped with screen-
capture software (Camtasia and Screenflow) were set up around a large
table. A script was read to each student to explain the purpose and proce-
dures of the testing (Appendix B). Students were asked to rank each task’s
difficulty on a number scale to record their perceived difficulty of each
question. It was emphasized that the system’s performance was the focus
of the testing, not the user’s research skills. They were encouraged to de-
scribe their process while performing the different tasks. Combined with
task-based assessment, think-aloud techniques allow a better understand-
ing of the choices made by testers. Guha and Saraf (2005), for example,
asked catalogue users to participate in assessment by explaining how they
conduct their searches in the OPAC and to express whether they are satis-
fied with the catalogues.

RESULTS
Throughout the course of the usability testing, both problems and
strengths of the current integrated library system were documented. Stu-
dents were able to easily find call numbers and determine if an item was
available. Students readily used the “find related results” link. Problem
areas included the advanced search limiters and exporting citations to a
bibliographic utility. A list of system weaknesses was also harvested from
reference transactions. Each system’s functionality was assessed by how
well it solved known user issues with the current OPAC. A list of desired
system options that would potentially address the usability issues was also
developed (Appendix C). The following specifications, grouped by broad
area, indicate the observed and reported solutions that NGCs or WSDSs
may provide.

Search Option Issues
Lack of Format Limiters. When users were asked to identify a particular
format, there were no postsearch filters to help limit results. They often
relied on the call number or title to determine item format. The system
offered “quick searches” and presearch limiters. Students who had prior
library instruction indicated that they found these useful, but those who
did not know about these options had very few postsearch tools at their
disposal. Advanced search labeling confused most users.

Solution: Limiters or filters should be available at the start of search
and as postsearch options. One desirable postsearch approach is to pro-
vide a hierarchical or “tree” approach to limiters in order to provide
groupings of like materials. For example, CDs, cassette tapes, LPs, and
online audio should be grouped under a limiter or filter for sound re-
cordings. By providing the ability to customize the labels of the limiters and what formats are included for each limiter, local preferences for terminology can be provided for users. In tests, users were often confused by library jargon for materials, such as sound recording, audiovisual materials, or digital collections. Users were more successful when presented with more familiar options like CD or DVD. To account for local cataloging practices, these limiters need to be highly customizable in terms of which fields of the MARC and item record indicate the limiter or filter.

Unfamiliarity with Controlled Vocabulary. Students do not use or recognize the utility of controlled vocabulary and preferred to perform natural language searches. This created difficulties in searches for materials about an author rather than by an author.

Solution: Search term completion (autocomplete) with thesaurus integration would assist in this area as well as providing limiters to select controlled vocabulary.

Labeling of Radio Buttons in Basic Search. Basic search-option radio buttons (keyword, browse, exact phrase) were misunderstood, as were options to search by title, author, or keyword.

Solution: Some of these problems may be solved by providing more understandable labels. It may also be advantageous to move these options to an advanced search form.

Unfamiliarity with Boolean Operators. The system performed best with the use of Boolean operators, but almost no students employed Boolean search strategies.

Solution: Implicit use of Boolean, proximity, and truncation options within relevancy ranking would improve system performance in this area.

Misspelled Search Terms. Frequent misspellings were common and yielded no search results.

Solution: Autocomplete and an effective did you mean? function would address this.

Display Feature Issues

Search Terms Not Highlighted. When search terms were not highlighted, students had difficulty assessing the search results, especially when the results are in last-in first-out (LIFO) order. Students were reluctant to click on item records to determine if results were relevant, and if they did, they were presented with little new information that would assist in selection. Users had to find and use tabs to access features—including contents notes and table of contents, which were effective methods to find works within larger works. Some users assumed that a record was a match without verifying the contents notes.

Solution: Remove tabbed display in item records. Provide search term highlighting in results list and item records and search term preview in the results list similar to Google Scholar. These features would support quick assessment of the effectiveness of search terms.
Format Labeling or Icons. When multiple formats were available, users could not quickly identify the preferred format.

Solution: Clear format labeling or icons to represent format types.

Guidance to Specific Location. A very frequent question according to the reference transaction form is how to find an item in the stacks. The system performed poorly in guiding the user to the exact location of items.

Solution: Display of library locations with maps improves the usefulness of the call number, prevents confusion in the stacks, and promotes student independence. Real time item availability should be part of the new system to promote independent utilization of library services, especially in environments with evening and nontraditional students.

Relevancy Ranking Issues

Difficulty Locating Collected Works and Anthologies. Students expected the most relevant results to be at the top of the page as is the convention on most e-commerce sites and Google searches.

Solution: A relevancy algorithm could prioritize collected works and anthologies. With WSDSs, it is desirable to have catalogue materials boosted in the rankings.

Ability to Customize Search Results. Librarians would like to provide specific resources for known assignments based on targeted keyword searches.

Solution: Librarians indicated that it would be useful to push course guides and preferred resources to the top of the results list to support local research activities. Access to search logs to see current or popular searches would enhance this feature, as would customizable results for certain search terms.

Additional Tool Issues

Need to Export Citations. Students’ ability to create accurate bibliographies was impeded by the catalogue’s cite/export features.

Solution: Providing users with clearly labeled citations that can be formatted for a variety of citation formats is a tool to avoid unintentional plagiarism.

Visibility of Interlibrary Loan Option. In the case that the library does not own an item, it is useful to indicate that patrons may be able to find it in another library and request it through interlibrary loan (ILL).

Solution: WorldCat and ILL links are available when no or very few items are retrieved. Clear labeling and placement customization options are desirable.

Navigation Issues

Search Box Needed to Revise Searches. Students were confused about where to revise searches because the search box was hidden at the bottom of the page or was not available until they started a new search or advanced search.
Solution: A search box at the top of every page is desired.

Nonfunctional Back Button. Subjects often relied on the back button to revise their search or to start a new search, especially given that there was not a search box on the top of every page. It required a click to bypass the form post-data-error message.

Solution: A functional back button is desired.

Searching for Articles in Catalogue. When asked how to locate an article, some students attempted to find them in the library catalogue.

Solution: In the end it was challenging to test the catalogue in isolation from the library Web site. Placement of links, labeling, and search boxes within the library site need to be purposefully designed to direct the user to the proper system in the case of multiple systems. A single search system that incorporates database content would reduce the need for students to distinguish between search boxes.

Next Steps and Recommendations
During the evaluation period, an attempt to conduct quantitative assessment of the systems’ performances was abandoned due to the small sample group and small number of evaluators. Hertzum and Jacobsen (2001) describe the problems with a small number of evaluators and lack of clear consensus in system evaluation. At this stage, it was of utmost importance to look at the big-picture problems and solutions in each interface. It was more useful to listen as students talked about a search and for the researcher to take notes on specific user issues with each system. Nielsen provides an iterative approach to usability wherein the first round of testing identifies the big-picture problems (Nielsen, 2011). Further rounds of testing with a smaller set of systems that meet basic requirements is an ideal approach.

Analysis of the usability test results indicated that the availability of features was just one consideration, and found that the customization process, labeling, and cataloguing choices of each site also contributed to the success or failure of the system. It was extremely valuable to see the potential of each system in the local context, but it is also important to remember that the selected implementations were customized for other institutions’ users and hence that it was not appropriate to expect them to meet all local user needs. Student search performances on competitive systems helped to identify features that solved users’ issues with the existing system. Results from the usability study revealed that features from both NGCs and WSDSs enhance find-ability of materials from the library catalogue. Rather than an out-of-the-box solution, testing identified the need to focus on the availability of functionalities through customization, enhancement requests, or vendor upgrade processes. Materials from the library catalogue can still be effectively located in a WSDS—even with the large number of resources that are searched—when labeling and navigation are effectively used.
This evidence-based approach is critical in defining a compelling rationale for system purchase. Different institutions will have varying system specifications depending on constituents’ needs. This may dictate a decision on whether both an NGC and WSDS are desired. By presenting a case of clear user problems and potential solutions available, advocates can make a strong case to administrators that limited resources would be used effectively and purposefully. In fact, armed with this new knowledge and given the cost of a new system, some libraries may choose to instead invest resources in improving their current system and Web site.

Working with vendors to see if they could accommodate preferred system specifications was the next phase of this approach to system selection. Approaching vendors with a clear sense of feature availability or with an awareness of the timeline of development for system specifications helps avoid feature overload. Some functionality may be easily implemented through simple vendor-side customization, while others require customization with a higher level of in-house expertise using the vendor’s API, for instance. If desired features are scheduled for a future release, it is important to find out if the status of the features is confirmed or guaranteed. Otherwise, features may be slow to come to fruition. In each instance, it was useful to have vendors respond to the status of requests in writing.

Working with vendors on system requests was the task of a small working group. This allowed the group to stay on task during this information-seeking stage. Once a small set of systems that met the basic requirements was selected, a larger group of stakeholders was consulted. Live vendor demonstrations and site visits were then scheduled. When possible, these visits were scheduled with product managers and trainers along with sales representatives. This phase involved reexamination of a small set of systems. This provided an opportunity to do usability tests with additional constituent groups.

It became clear at this point in the process that determining effective labeling for limiters or filters and navigation remains a critical need. Card sorting techniques could be utilized to complete this process. The results of the initial round of usability tests could be used to customize a group of demo systems to incorporate labeling and customization preferences so that systems are representative of local labeling choices and preferences.

Even with this multifaceted approach to making a selection, the testers were not able to identify one single system that provided an out-of-the-box implementation strategy. To reiterate, it is imperative to consult with vendors to determine the level and type of customizations available. Local user needs may be as diverse as the populations served, and it may be desirable to harvest a list of desired elements from a variety of systems. Financial and staffing realities may result in the necessity for sacrifice; some system specifications may need to be sacrificed in order to meet budgetary constraints. Prioritization of desirable features will facilitate this process. Libraries may find that some systems may not be appropriate due to
the expertise and staffing hours needed for administrative support and customization. Consulting with vendors may improve the likelihood that more specifications are addressed by the initial purchase. While the main goal of these usability tests was to improve find-ability of library resources through the justification for the purchase of an NGC or WSDS, continued and periodic usability tests are necessary to keep abreast of possible system customizations and enhancements.

APPENDIX A

STAFF SURVEY RESULTS

- Locating desired materials can be difficult.
- Library instruction is necessary for users to learn how to use the system.
- Catalogue relevancy rankings are needed.
- Sorting options are lacking.
- Facets would be useful.
- Record exporting or a citation button associated with item records is desired.
- Retain the feature to mark items to put in a folder without a log-in.
- Retain enhanced content (book jackets and tables of contents).
- Customized quick search functions work well.
- Finding similar items with one click works well.

REFERENCE TRANSACTION LOG FINDINGS

- The majority of user issues related to keyword or subject searches.
- There is a curricular emphasis on the arts and religion and students frequently need assistance with these subject areas.
- Users often inquired about how to find a known item in the catalogue.
- Users requested assistance at the desk on how to find specific formats like e-books and DVDs.
- Help is needed to find items beyond the library’s collection.
- Locating a work within a larger work, for example, a song on a CD or a poem in an anthology, is difficult.
- Users wanted to know how to find the library’s journal holdings.
- Assistance was needed to find a specific subject area in the library stacks.
- Users asked for help finding known items for which they did not have complete information.

APPENDIX B

USABILITY SCRIPT AND TASKS

(Read to student) Below are some common questions that students have. Please start at the library homepage to search for the answer to each ques-
tion after I read it aloud. Describe your process out loud as you go. For example:

“First I’m going to type ‘insomnia’ in the search box because I think that is the right keyword to use to find what I need. I’m looking through the results, and I see a book with the title Understanding Sleeplessness. This looks like it could be good.”

After attempting each question, please rate the difficulty of each task on a scale of 1 to 5.

1 = most difficult
2 = difficult
3 = moderate
4 = easy
5 = very easy

Table 1. Usability Tasks

<table>
<thead>
<tr>
<th>Question</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a book about the Industrial Revolution.</td>
<td></td>
</tr>
<tr>
<td>Find a book that contains the poem “Birches” by Robert Frost.</td>
<td></td>
</tr>
<tr>
<td>Export a citation of this book.</td>
<td></td>
</tr>
<tr>
<td>What is the call number for Bibles?</td>
<td></td>
</tr>
<tr>
<td>Where are they in the library?</td>
<td></td>
</tr>
<tr>
<td>Find an e-book about Gabriel Garcia Marquez.</td>
<td></td>
</tr>
<tr>
<td>Find a CD that has the song “How High the Moon.”</td>
<td></td>
</tr>
<tr>
<td>Find a periodical/journal/magazine that is about psychology.</td>
<td></td>
</tr>
<tr>
<td>Find a DVD about Africa.</td>
<td></td>
</tr>
<tr>
<td>Find a novelist who writes fiction similar to Agatha Christie’s works.</td>
<td></td>
</tr>
<tr>
<td>Find the reference book The Merck Index.</td>
<td></td>
</tr>
<tr>
<td>What other local libraries might have the book titled Aquarius Conspiracy if we don’t have it in our library?</td>
<td></td>
</tr>
<tr>
<td>Renew books that you have already checked out.</td>
<td></td>
</tr>
<tr>
<td>Find a book and an article about Sojourner Truth.</td>
<td></td>
</tr>
<tr>
<td>Find the following article:</td>
<td></td>
</tr>
</tbody>
</table>

APPENDIX C

DISCOVERY SYSTEM SPECIFICATIONS

Search Options

• Limiters or filters available at the start of search and as postsearch options
• At the postsearch, a hierarchical or “tree” approach to limiters to provide like materials in groupings
• Highly customizable facets for material type
• Search term completion (autocomplete) with thesaurus integration
• Postsearch limiters for controlled vocabulary
• Customizable advanced search form
• Implicit use of Boolean, proximity, and truncation options within relevancy ranking
• Prioritizing primary sources in relevance ranking
• “Did you mean...”
• No tabbed display in item records
• Search term highlighting in results list and item records
• Search term preview in the results list similar to Google Scholar
• Clear format labeling or icons to represent format types
• Display of library locations with maps
• Real time circulation availability

Relevancy Ranking
• Collected works and anthologies prioritized in relevancy ranking
• Catalogue materials boosted in the rankings (WSDSs)
• Ability to push course guides and preferred resources to the top of the results list to support local research activities
• Access to search logs to customize relevance

Navigation
• Search box at the top of every page
• Functioning back button

Additional Tools
• Clearly labeled citations that can be formatted for a variety of citation formats
• WorldCat and ILL links available when no or very few items are retrieved

Note
1. Vaughan (2011) provides a definition of Web-scale discovery and the differentiation between Web-scale discovery products and next-generation discovery layers.

References

Teri Oaks Gallaway is the library systems and Web coordinator at the J. Edgar and Louise S. Monroe Library at Loyola University New Orleans. She performs librarian liaison duties including instruction and collection development for the sociology, criminal justice, and honors departments. She has given presentations at Electronic Resources & Libraries, the Handheld Librarian, and Educause on promoting and managing electronic resources and on incorporating technology skills into library instruction.

Mary Finnan Hines is the learning commons and outreach librarian in the J. Edgar and Louise S. Monroe Library at Loyola University New Orleans. Her position focuses on user services, and she leads the Student Library Advisory Committee. She provides collection development and instruction to the languages and cultures department. She has presented at the Louisiana Library Association’s 2011 Annual Conference as well as IFLA’s 2009 satellite conference in Bologna, Italy.