The Next Generation Library Catalog: A Comparative Study of the OPACs of Koha, Evergreen, and Voyager

Sharon Q. Yang and Melissa A. Hofmann

Open source has been the center of attention in the library world for the past several years. Koha and Evergreen are the two major open-source integrated library systems (ILSs), and they continue to grow in maturity and popularity. The question remains as to how much we have achieved in open-source development toward the next-generation catalog compared to commercial systems. Little has been written in the library literature to answer this question. This paper intends to answer this question by comparing the next-generation features of the OPACs of two open-source ILSs (Koha and Evergreen) and one proprietary ILS (Voyager’s WebVoyage).

Much discussion has occurred lately on the next-generation library catalog, sometimes referred to as the Library 2.0 catalog or “the third generation catalog.”1 Different and even conflicting expectations exist as to what the next-generation library catalog comprises:

In two sentences, this catalog is not really a catalog at all but more like a tool designed to make it easier for students to learn, teachers to instruct, and scholars to do research. It provides its intended audience with a more effective means for finding and using data and information.2

Such expectations, despite their vagueness, eventually took concrete form in 2007.3 Among the most prominent features of the next-generation catalog are a simple keyword search box, enhanced browsing possibilities, spelling corrections, relevance ranking, federated navigation, federated search, user contribution, and enriched content, just to mention a few. Over the past three years, libraries, vendors, and open-source communities have intensified their efforts to develop OPACs with advanced features. The next-generation catalog is becoming the current catalog.

The library community welcomes open-source integrated library systems (ILSs) with open arms, as evidenced by the increasing number of libraries and library consortia that have adopted or are considering open-source options, such as Koha, Evergreen, and the Open Library Environment Project (OLE Project). Librarians see a golden opportunity to add features to a system that will take years for a proprietary vendor to develop. Open-source OPACs, especially that of Koha, seem to be more innovative than their long-established proprietary counterparts, as our investigation shows in this paper. Threatened by this phenomenon, ILS vendors have rushed to improve their OPACs, modeling them after the next-generation catalog. For example, Ex Libris pushed out its new OPAC, WebVoyage 7.0, in August of 2008 to give its OPAC a modern touch.

One interesting question remains. In a competition for a modernized OPAC, which OPAC is closest to our visions for the next-generation library catalog: open-source or proprietary? The comparative study described in this article was conducted in the hope of yielding some information on this topic. For libraries facing options between open-source and proprietary systems, “a thorough process of evaluating an integrated library system (ILS) today would not be complete without also weighing the open source ILS products against their proprietary counterparts.”3

Scope and Purpose of the Study

The purpose of the study is to determine which OPAC of the three ILSs—Koha, Evergreen, or WebVoyage—offers more in terms of services and is more comparable to the next-generation library catalog. The three systems include two open-source and one proprietary ILSs. Koha and Evergreen are chosen because they are the two most popular and fully developed open-source ILSs in North America. At the time of the study, Koha had 936 implementations worldwide; Evergreen had 543 library users.4 We chose WebVoyage for comparison because it is the OPAC of the Voyager ILS by Ex Libris, the biggest ILS vendor in terms of personnel and marketplace.5 It also is one of the more popular ILSs in North America, with a customer base of 1,424 libraries, most of which are academic.6 As the sample only includes three ILSs, the study is very limited in scope, and the findings cannot be extrapolated to all open-source and proprietary catalogs. But, hopefully, readers will gain some insight into how much progress libraries, vendors, and open-source communities have achieved toward the next-generation catalog.

Literature Review

A review of the library literature found two relevant studies on the comparison of OPACs in recent years. The first study was conducted by two librarians in Slovenia investigating how much progress libraries had made toward the next-generation catalog.7 Six online catalogs

Sharon Q. Yang (yangs@rider.edu) is Systems Librarian and Melissa A. Hofmann (mhofmann@rider.edu) is Bibliographic Control Librarian, Rider University.
were examined and evaluated, including WorldCat, the Slovene union catalog COBISS, and those of four public libraries in the United States. The study also compared services provided by the library catalogs in the sample with those offered by Amazon. The comparison took place primarily in six areas: search, presentation of results, enriched content, user participation, personalization, and Web 2.0 technologies applied in OPACs. The authors gave a detailed description of the research results supplemented by tables and snapshots of the catalogs in comparison. The findings indicated that “the progress of library catalogues has really been substantial in the last few years.” Specifically, the library catalogues have made “the best progress on the content field and the least in user participation and personalization.” When compared to services offered by Amazon, the authors concluded that “none of the six chosen catalogues offers the complete package of examined options that Amazon does.”

In other words, library catalogs in the sample still lacked features compared to Amazon.

The other comparative study was conducted by Linda Riewe, a library school student, in fulfillment for her master’s degree from San Jose University. The research described in her thesis is a questionnaire survey targeted at 361 libraries that compares open-source (specifically, Koha and Evergreen) and proprietary ILSs in North America. More than twenty proprietary systems were covered, including Horizon, Voyager, Millennium, Polaris, Innopac, and Unicorn. Only a small part of her study was related to OPACs. It involved three questions about OPACs and asked librarians to evaluate the ease of use of their ILS OPAC’s search engines, their OPAC search engine’s completeness of features, and their perception of how easy it is for patrons to make self-service requests online for renewals and holds. A scale of 1 to 5 was used (1 = least satisfied; 5 = very satisfied) regarding the three aspects of OPACs. The mean and medium satisfaction ratings for open-source OPACs were higher than those of proprietary ones. Koha’s OPAC was ranked 4.3, 3.9, and 3.9, respectively in mean, the highest on the scale in all three categories, while the whole OPACs were ranked 3.9, 3.6, and 3.6. Evergreen fell in the middle, still ahead of proprietary OPACs. The findings reinforced the perception that open-source catalogues, especially Koha, offer more advanced features than proprietary ones. As Riewe’s study focused more on the cost and user satisfaction with ILSs, it yielded limited information about the connected OPACs.

No comparative research has measured the progress of open-source versus proprietary catalogues toward the next-generation library catalog. Therefore the comparison described in this paper is the first of its kind. As only Koha, Evergreen, and Voyager’s OPACs are examined in this paper, the results cannot be extrapolated. Studies on a larger scale are needed to shed light on the progress librarians have made toward the next-generation catalog.

I Method

The first step of the study was identifying and defining of a set of measurements by which to compare the three OPACs. A review of library literature on the next-generation library catalog revealed different and somewhat conflicting points of views as to what the next-generation catalog should be. As Marshall Breeding put it, “There isn’t one single answer. We will see a number of approaches, each attacking the problem somewhat differently.” This study decided to use the most commonly held visions, which are summarized well by Breeding and by Morgan’s LITA executive summary. The ten parameters identified and used in the comparison were taken primarily from Breeding’s introduction to the July/August 2007 issue of Library Technology Reports, “Next-Generation Library Catalogs.” The ten features reflect some librarians’ visions for a modern catalog. They serve as additions to, rather than replacements of, the feature sets commonly found in legacy catalogs. The following are the definitions of each measurement:

- A single point of entry to all library information: “Information” refers to all library resources. The next-generation catalog contains not only bibliographical information about printed books, video tapes, and journal titles but also leads to the full text of all electronic databases, digital archives, and any other library resources. It is a federated search engine for one-stop searching. It not only allows for one search leading to a federation of results, it also links to full-text electronic books and journal articles and directs users to printed materials.

- State-of-the-art Web interface: Library catalogs should be “intuitive interfaces” and “visually appealing sites” that compare well with other Internet search engines. A library’s OPAC can be intimidating and complex. To attract users, the next-generation catalog looks and feels similar to Google, Amazon, and other popular websites. This criterion is highly subjective, however, because some users may find Google and Amazon anything but intuitive or appealing. The underlying assumption is that some Internet search engines are popular, and a library catalog should be similar to be popular themselves.

- Enriched content: Breeding writes, “Legacy catalogs tend to offer text-only displays, drawing only on the MARC record. A next-generation catalog might bring in content from different sources to strengthen the visual appeal and increase the amount of information presented to the user.” The enriched content
includes images of book covers, CD and movie cases, tables of contents, summaries, reviews, and photos of items that traditionally are not present in legacy catalogs.

- **Faceted navigation:** Faceted navigation allows users to narrow their search results by facets. The types of facets may include subjects, authors, dates, types of materials, locations, series, and more. Many discovery tools and federated search engines, such as Villanova University’s VuFind and Innovative Interface’s Encore, have used this technology in searches.\(^1\) Auto-Graphics also applied this feature in their OPAC, AGent Iluminar.\(^2\)

- **Simple keyword search box:** The next-generation catalog looks and feels like popular Internet search engines. The best example is Google’s simple user interface. That means that a simple keyword search box, instead of a controlled vocabulary or specific-field search box, should be presented to the user on the opening page with a link to an advanced search for user in need of more complex searching options.

- **Relevancy:** Traditional ranking of search results is based on the frequency and positions of terms in bibliographic records during keyword searches. Relevancy has not worked well in OPACs. In addition, popularity is another factor that has not been taken into consideration in relevancy ranking. For instance, “When ranking results from the library’s book collection, the number of times that an item has been checked out could be considered an indicator of popularity.”\(^3\) By the same token, the size and font of tags in a tag cloud or the number of comments users attach to an item may also be considered relevant in ranking search results. So far, almost no OPACs are capable of incorporating circulation statistics into relevancy ranking.

- **“Did you mean . . . ?”** When a search term is not spelled correctly or nothing is found in the OPAC in a keyword search, the spell checker will kick in and suggest the correct spelling or recommend a term that may match the user’s intended search term. For example, a modern catalog may generate a statement such as “Did you mean . . . ?” or “Maybe you meant . . . .” This may be a very popular and useful service in modern OPACs.

- **Recommendations and related materials:** The next-generation catalog is envisioned as promoting reading and learning by making recommendations of additional related materials to patrons. This feature is an imitation of Amazon and websites that promote selling by stating “Customers who bought this item also bought . . . .” Likewise, after a search in the OPAC, a statement such as “Patrons who borrowed this book also borrowed the following books . . . .” may appear.

- **User contribution—ratings, reviews, comments, and tagging:** Legacy catalogs only allow catalogers to add content. In the next-generation catalog, users can be active contributors to the content of the OPAC. They can rate, write reviews, tag, and comment on items. User contribution is an important indicator for use and can be used in relevancy ranking.

- **RSS feeds:** The next-generation catalog is dynamic because it delivers lists of new acquisitions and search updates to users through RSS feeds. Modern catalogs are service-oriented; they do more than provide a simple display search results.

The second step is to apply these ten visions to the OPACs of Koha, Evergreen, and WebVoyage to determine if they are present or absent. The OPACs used in this study included three examples from each system. They may have been product demos and live catalogs randomly chosen from the user list on the product websites. The latest releases at the time of the study was Koha 3.0, Evergreen 2.0, WebVoyage 7.1. In case of discrepancies between product descriptions and reality, we gave precedence to reality over claims. In other words, even if the product documentation lists and describes a feature, this study does not include it if the feature is not in action either in the demo or live catalogs. Despite the fact that a planned future release of one of those investigated OPACs may add a feature, this study only recorded what existed at the time of the comparison. The following are the OPACs examined in this paper.

**Koha**

- Koho Demo for Academic Libraries: http://academic.demo.kohalibrary.com/
- Wagner College: http://wagner.waldo.kohalibrary.com/
- Clearwater Christian College: http://ccc.kohalibrary.com/

**Evergreen**

- Columbia Bible College at http://columbiabc.evergreencatalog.com/opac/en-CA/skin/default/xml/index.xml

**WebVoyage**

- Rider University Libraries: http://voyager.rider.edu
- Renton College library: http://renton.library.ctc.edu/vwebv/searchBasic
Shoreline College library: http://shoreline.library .ctc.edu/vwebv/searchBasic

The final step includes data collection and compilation. A discussion of findings follows. The study draws conclusions about which OPAC is more advanced and has more features of the next-generation library catalog.

Findings

Each of the OPACs of Koha, Evergreen, and WebVoyage are examined for the presence of the ten features of the next-generation catalog.

Single Point of Entry for All Library Information

None of the OPACs of the three ILSs provides true federated searching. To varying degrees, each is limited in access, showing an absence of contents from electronic databases, digital archives, and other sources that generally are not located in the legacy catalog. Of the three, Koha is more advanced. While WebVoyage and Evergreen only display journal-holdings information in their OPACs, Koha links journal titles from its catalog to ProQuest’s Serials Solutions, thus leading users to full-text journals in the electronic databases. The example in figure 1 (Koha demo) shows the journal title Unix Update with an active link to the full-text journal in the availability field. The link takes patrons to Serials Solutions, where full text at the journal-title level is listed for each database (see figure 2). Each link will take you into the full text in each database.

State-of-the-Art Web Interface

As beauty is in the eye of the beholder, the interface of a catalog can be appealing to one user but prohibitive to another. With this limitation in mind, the out-of-the-box user interface at the demo sites was considered for each OPAC. All the three catalogs have the Google-like simplicity in presentation. All of the user interfaces are highly customizable. It largely depends on the library to make the user interface appealing and welcoming to users. Figures 3–5 show snapshots from each ILSs demo sites and have not been customized.

However, there are a few differences in the “state of the art.” For one, Koha’s navigation between screens relies solely on the browser’s Forward and Back buttons, while WebVoyage and Evergreen have internal navigation buttons that more efficiently take the user between title lists, headings lists, and record displays, and between records in a result set. While all three OPACs offer an advanced search page with multiple boxes for entering search terms, only WebVoyage makes the relationship between the terms in different boxes clear. By the use of a drop-down box, it makes explicit that the search terms are by default ANDed and also allows for the selection of OR and NOT. In Koha’s and Evergreen’s advanced search, however, the terms are ANDed only, a fact that is not at all obvious to the user. In the demo OPACs examined, there is no option to choose OR or NOT between rows, nor is there any indication that the search is ANDed. The point of providing multiple search boxes is to guide users in constructing a Boolean search without their having to worry about operators and syntax. In Koha, however, users have to type an OR or NOT statement themselves within the text box, thus defeating the purpose of having multiple boxes. While Evergreen allows for a NOT construction within a row (“does not contain”), it does not provide an option for OR (“contains” and “matches exactly” are the other two options available). See figures
Thus Koha’s and Evergreen’s advanced search is less than intuitive for users and certainly less functional than WebVoyage’s.

**Enriched Content**

To varying degrees, enriched content is present in all three catalogs, with Koha providing the most. While all three catalogs have book covers and movie-container art, Koha has much more in its catalog. For instance, it displays tags, descriptions, comments, and Amazon reviews. WebVoyage displays links to Google Books for book reviews and content summaries but does not have tags, descriptions, and comments in the catalog. See figures 9–11.

**Faceted Navigation**

The Koha OPAC is the only catalog of the three to offer faceted navigation. The “Refine your search” feature allows users to narrow search results by availability, places, libraries, authors, topics, and series. Clicking on a term within a facet adds that term to the search query and generates a narrower list of results. The user may then choose another facet to further refine the search. While Evergreen appears to have faceted navigation upon first glance, it actually does not possess this feature. The following facets appear after a search generates hits: “Relevant subjects,” “Relevant authors,” and “Relevant series.” But choosing a term within a facet does not narrow down the previous search. Instead, it generates an entirely new search with the selected term; it does not add the new term to the previous query. Users must manually combine the terms in the simple search box or through the advanced search page. WebVoyage also does not offer faceted navigation—it only provides an option to “Filter your search” by format, language, and date when a set of results is returned. See figures 12–14.

**Keyword Searching**

Koha, Evergreen, and WebVoyage all present a simple keyword search box with a link to the advanced search (see figures 3–5).

**Relevancy**

Neither Koha, Evergreen, nor WebVoyage provide any evidence for meeting the criteria of the next-generation catalog’s more inclusive vision of relevancy ranking, such as accounting for an item’s popularity or allowing user tags. Koha uses Index Data’s Zebra program for its relevance ranking, which “reads structured records in a variety of input formats . . . and allows access to them through exact boolean search
expressions and relevance-ranked free-text queries.19

Evergreen’s DokuWiki states that

the base relevancy score is determined by the cover density of the searched terms. After this base score is determined, items may receive score bumps based on word order, matching on the first word, and exact matches depending on the type of search performed.20

These statements do not indicate that either Koha or Evergreen go beyond the traditional relevancy-ranking methods of legacy systems, such as WebVoyage.

Did You Mean . . . ?

Only Evergreen has a true “Did you mean . . . ?” feature. When no hits are returned, Evergreen provides a suggested alternate spelling (“Maybe you meant . . . ?”) as well as a suggested additional search (“You may also like to try these related searches . . . ”). Koha has a spell-check feature, but it automatically normalizes the search term and does not give the option of choosing different one. This is not the same as a “Did you mean . . . ?” feature as defined above. While the normalizing process may be seamless, it takes the power of choice away from the user and may be problematic if a particular alternative spelling or misspelling is searched purposefully, such as “womyn.” (When “womyn” is searched as a keyword in the Koha demo OPAC, 16,230 hits are returned. This catalog does not appear to contain the term as spelled, which is why it is normalized to women. The fact that the term does not appear as is may not be transparent to the searcher.) With normalization, the user may also be unaware that any mistake in spelling has occurred, and the number of hits may differ between the correct spelling and the normalized spelling, potentially affecting discovery. The normalization feature also only works with particular combinations of misspellings, where letter order affects whether a match is found. Otherwise the system returns a “No result found!” message with no suggestions offered. (Try “homosexuality” vs. “homo-sexuality.” In Koha’s demo OPAC, the former, with a missing “s,” yields 553 hits, while the latter, with a misplaced “s,” yields none.) However, Koha is a step ahead of WebVoyage, which has no built-in spell checker at all. If a search fails, the system returns the message “Search Resulted in No Hits.” See figures 15–17.

Recommendations/Related Materials

None of the three online catalogs can recommend materials for users.

User Contributions

Koha is the only system of the three that allows users to add tags, comments, descriptions, and reviews. In Koha’s OPAC, user-added tags form tag clouds, and the font and size of each keyword or tag indicate that keyword or
Figure 9. Koha enriched content

Figure 10. Evergreen enriched content

Figure 11. Voyager enriched content

Figure 12. Koha faceted navigation

Figure 13. Evergreen faceted navigation

Figure 14. Voyager faceted navigation
Nevertheless, the user contribution in the Koha OPAC is not easy to use. It may take many clicks before a user can figure out how to add or edit text. It requires user login, and the system cannot keep track of the search hits after a login takes place. Therefore the user contribution features of Koha need improvement. See figure 18.

**RSS feeds**

Koha provides RSS feeds, while Evergreen and WebVoyage do not.

### Conclusion

Table 1 is a summary of the comparisons in this paper. These comparisons show that the Koha OPAC has six out of the ten compared features for the next-generation catalog, plus two halves. Its full-fledged features include state-of-the-art Web interface, enriched content, faceted navigation, a simple keyword search box, user contribution, and RSS feeds. The two halves indicate the existence of a feature that is not fully developed. For instance, “Did you mean . . . ?” in Koha does not work the way the next-generation catalog is envisioned. In addition, Koha has the capability of linking journal titles to full text via Serials Solutions, while the other two OPACs only display holdings information. Evergreen falls into second place, providing four out of the ten compared features: state-of-the-art interface, enriched content, a keyword search box, and “Did you mean . . . ?” WebVoyage, the Voyager OPAC from Ex Libris, comes in third, providing only three out of the ten features for
the next-generation catalog.

Based on the evidence, Koha’s OPAC is more advanced and innovative than Evergreen’s or Voyager’s. Among the three catalogs, the open-source OPACs compare more favorably to the ideal next-generation catalog than the proprietary OPAC. However, none of them is capable of federated searching. Only Koha offers faceted navigation. WebVoyage does not even provide a spell checker. The ILS OPAC still has a long way to go toward the next-generation catalog. Though this study samples only three catalogs, hopefully the findings will provide a glimpse of the current state of open-source versus proprietary catalogs.

ILS OPACs are not comparable in features and functions to stand-alone OPACs, also referred to as “discovery tools” or “layers.” Some discovery tools, such as Ex Libris’ Primo, also are federated search engines and are modeled after the next-generation catalog. Recently they have become increasingly popular because they are bolder and more innovative than ILS OPACs. Two of the best stand-alone open-source OPACs are Villanova University’s VuFind and Oregon State University’s LibraryFind.21 Both boast eight out of ten features of the next-generation catalog.22 Technically it is easier to develop a new stand-alone OPAC with all the next-generation catalog features than mending old ILS OPACs. As more and more libraries are disappointed with their ILS OPACs, more discovery tools will be implemented. Vendors will stop improving ILS OPACs and concentrate on developing better discovery tools. The fact that ILS OPACs are falling behind current trends may eventually bear no significance for libraries—at least for the ones that can afford the purchase or implementation of a more sophisticated discovery tool or stand-alone OPAC. Certainly small and public libraries who cannot afford a discovery tool or a programmer for an open-source OPAC overlay will suffer, unless market conditions change.

References

4. Ibid.
9. Ibid.
11. Ibid., 26–27.
15. Ibid.
16. Ibid.

