



Encore Evaluation

Background:

Since migrating to Sierra a flaw had been identified in Innovative's results list, specifically for title searches and to a lesser extent, keyword searching. The Resource Sharing Committee was tasked to look at the issue and a Discovery Tools Sub-Committee was formed to discuss the possibility of implementing an OPAC enhancement tool (Encore) to improve the patron experience. On September 16, Laurie Shedrick from Innovative was able to demo Encore to the committee but it was felt that we would need to take a closer look at the product before committing. This usability test here is extremely limited and, as standard practice, ongoing testing should be attempted to make sure that Innovative's search algorithm, index rules and cataloging practice align to provide the best possible experience for the user.

Current configuration contributing to title (and keyword) search results (wordy explanation):

Title:

A title search does not use any relevancy rankings (unlike a keyword search) therefore it is fairly unforgiving. A title search for "Infinite Jest" will return only items with that exact search in one of several MARC fields. When Mid-Hudson originally migrated to Millennium, several index rules were established and those have not been reconfigured. Sierra and Millennium title search results have not changed because the index rules have not changed. For title searches the following MARC fields will return results:

240 – Serial title entry (all subfields excluding f and h)

245 – Main title entry (all subfields excluding h and c)

246/7 – Added title (including subfields a, b, g, n, and p)

400 – Series title (discontinued) (only subfields n, p, t and v)

440 – Series title (including subfields a, n, p, v)

490 – Series title (including subfields a and v)

505 – Notes, used for episode titles, song titles and book titles when attached to eReaders (only subfield t)

730 – Added title (all subfields excluding h and x)

740 – see above

8xx – Series information (including subfields k, g, k, l, m, n, o, p, r, s, t, v)



I imagine these indexing rules are the built-in default because it covers nearly all possible references to a title in a MARC record. The issue is not necessarily that these records are included in a title search, many records might be lost if these fields were not included, but the fact that Sierra/OPAC will only display what is found in alphabetical order based on what is returned in the 245. In the OPAC, this is called "system sorted" and is the default (Sierra) display. For example, a patron's title search for "Gone Girl" might likely be for the book by Gillian Flynn but because the book "Gone Girl" has been bought by several libraries that have this title included on their Amazon Kindles, the first 6 entries in the title search will be for variations on the Amazon Kindle because "Amazon Kindle" is what appears in the 245 field and results are returned alphabetically. "Gone Girl" in this instance will have been included in the MARC record in field 505 subfield t.

This title search issue becomes exponentially worse if the title is not "unique" enough. "On the Road" is the example that has been pointed to as especially problematic. "On the Road", as a title phrase, when searched in Sierra, returned results in several MARC fields and returns over 100 results. It seems likely that when a patron or staff member executes this search, they might be looking for the book by Jack Kerouac, and they wouldn't locate this title until first scrolling through over 20 titles. The phrase "On the Road" appears, in addition to the 245, in the 246, 505, 700, and 730 in the test results set. These results are then displayed alphabetically.

Choosing to re-index a title search may cause problems. Is it farfetched to imagine a patron or Sierra user that might be searching the title index for and not finding the Chekhov short story "On the Road" in the book "The Steppe and Other Stories" unless we trace the title search to the 505, which is where "On the Road" appears in this MARC record? In addition, "Amazon Kindle Fire: for Butterfield Library" which has the Kerouac book in the 505, would not be returned either. Taking the brute force approach of eliminating the fields that do not neatly provide title search results (for example, 2xx, 505, 4xx, 7xx and 8xx) might have the unintended consequence of losing a bib record for a searcher just to make the result set for some title searches cleaner.

Keyword:

Keyword indexing is based on these same fields; a relevancy ranking is applied and the results are then displayed in descending chronological order based on the publication date. Multi-word relevance ranking is at work in this type of search using an algorithm that gives weight to results depending on which field the keywords appear. The results are then ordered based on this, and then furthermore, the publication date. Further testing with regard to the index rules for keyword results might help improve some of the results that were found in this usability test and can be considered as a potential solution. The testing here is focused on keyword because a large majority of patron OPAC searches are keyword and that index is our default.



Potential Solution: Encore and forcing an improved Keyword search

This solution does not change any of the problems with the title search. The difference is that Encore forces the OPAC user to either select advanced search to perform a title search or perform a much improved keyword search. The Encore interface does not offer the dropdown selection (keyword, title, author, etc.), the default is one search for databases and the catalog. The indexing we have in place for keyword search may or may not need to change however the relevancy rankings should improve. We asked Jessica Kerr from Woodstock and Nina Acosta from MHLS to take a closer look at Encore: by using a test version of the database supplied by the vendor, the MARINet OPAC from California which uses Encore, and our current OPAC (midhudsonlibraries.org). The flaw in this test is that we cannot completely recreate the variables (bib records being searched and customized indexes being set) to understand how our database will return results if Encore were to be implemented.



Jessica's Results:

Keyword search: "On + The+ Road" in midhudsonlibraries.org

Order of result	Item Type
15	eAudiobook
23	Book on CD
24	Print Book

Keyword search: "On + The+ Road" on MARINet

Order of result	Item Type
1	Print Book
4-6	Further Print Editions
8	Book on CD

Keyword search: "On + The+ Road" on Innovative's demo environment

Order of result	Item Type
1	eBook
2	Book on CD
3-4	Print Material
5	Book on Tape

Keyword search: "On + The+ Road" on Amazon

Order of result	Title
1	Book (movie tie-in)
2	Instant Video
3	The Road by Cormac McCarthy
4	DVD



With our current OPAC search, a patron would need to return results on the second page to find the print books. If our patrons also use the keyword “Kerouac”, the results improve for midhudsonlibraries.org but the patron would need to know how to spell that.

Keyword search for “Help” with the intention of finding Kathryn Stockett’s book or the record by The Beatles.

midhudsonlibraries.org:

Order of result	Item/Material type
33	Music from the motion picture
37	Blu-Ray
39	Deluxe edition (with two items attached)
42	DVD
43	Book club in a bag
57	eAudiobook
58	eBook
65	Beatles CD (1 st Beatles entry)
69	Book on CD
73	Book with the majority of items attached
74	Large Print

MARINet

Order of result	Item/Material type
1	DVD
2	Book
3	Book on CD
4	Large Print
5-7	Books with same title different author
8	Blu-ray
9	DVD different movie
10	Beatles music CD



Innovative Demo Product

Order of result	Item/Material type
1	Book
2	Book on CD
3	Large Print
4	Book with different author same title
5	Beatles VHS
6	Book with different author same title
7	Article PDF (help-seeking)

Amazon

Order of result	Item/Material type
1	Beatles Blu-Ray
2	Book
3	Book with different author same title
4	DVD
5	Beatles CD

In this test, midhudsonlibraries.org fails spectacularly. The result list is dominated on the top with the keyword “Help” in the title somewhere and is displayed in descending order based on publication date, not popularity as MARINet seems to be. The first target bib record is not returned until result number thirty nine and because this version of The Help by Kathryn Stockett is the deluxe edition, this isn’t even the “normal” bib record this search is being tested against. The standard bib record for the book is not returned until result seventy three. Encore had much improved results.

Addition Encore OPAC enhancements that deserve closer inspection:

Overdrive API: If MHLS chose Encore, the system would be able to take advantage of an Overdrive API. This API will allow the patrons to manage their Overdrive accounts without leaving the OPAC.

Federated Search: Encore is also a federated search interface so Encore is returning indexed results from our databases in keyword searching. An improvement from traditional federated search by

Better Interface: The interface is not unlike Sierra by allowing a patron to narrow down their search using facets.



Issues:

Encore would need to be implemented prior to migrating to Boopsie.

In May, we had a quote for Encore at 42,875/year (3 year commitment) plus a 2,500 set-up fee.

Would not allow for look and feel design adjustments, we would be able to add our logo and not much else.