

Patricia Hswe
LIS591: Independent Study
Director: Tim Cole
1 August 2007

In Search of Relevance: A Comparative Study of Query Runs in Two Interfaces in the IMLS DCC Metadata Repository

Introduction

A central aim of the IMLS Digital Collections and Content (DCC) project, based in Grainger Engineering Library at UIUC, is to make searching across distributed digital collections an interoperable process, particularly by aggregating and standardizing the metadata about those items via the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The guiding research question of the project is, “How can resource developers best represent collections and items to meet the needs of service providers and users?” One outcome of this inquiry was the development of a searchable, item-level metadata repository, made public in summer 2006. Refinements to the search interface are ongoing, and a new version of it was unveiled in summer 2007.

The current IMLS DCC search interface draws upon a database of metadata records that were reprocessed differently compared to the records that were linked to the interface launched in 2006. The changes in reprocessing that have occurred in the last year on the IMLS DCC project are based partly on workflow recommendations submitted by Parmit Chilana, a Master’s student at UIUC’s Graduate School of Library and Information Science. During a practicum she undertook in summer 2006, Chilana reviewed and documented the metadata reprocessing and augmentation steps then being used by the IMLS DCC project. She also looked at the techniques (in the form of scripts and XSLT transformations) that had been implemented by the UIUC Committee on Institutional Cooperation (CIC) metadata portal. Her intent was to understand how the reprocessing and augmentation steps for the CIC metadata portal might be transferred to the work on metadata being carried out by the IMLS DCC project. Chilana’s suggestions for improvements addressed data inconsistencies, mistakes, confusions, and insufficiencies that she discovered during her analysis of metadata records harvested by the IMLS DCC project—which adopted the suggestions, which in part led to an enhanced search interface for the metadata repository.

Post-harvest metadata reprocessing, particularly toward the improvement of metadata quality, has been the topic of a few articles and book chapters in recent years (Cole & Foulonneau, 2007; Foulonneau, Cole, Habing & Shreeves, 2005; Hillmann, Dushay & Phipps, 2004; Dushay & Hillmann, 2003). A main impetus behind these investigations of metadata quality in digital collections is to enhance the search and discovery process for end-users looking for items in them. In this independent study project, in which I essentially picked up where Chilana’s practicum left off, my intent was to assess the effects of these improved reprocessing and augmentation practices, particularly in the context of the results returned after a search—i.e., after a query has been input in the search interface. My overarching research questions, which are inter-related, were:

- Have modifications to metadata reprocessing and augmentation made a difference to retrieval in an end-user search?
- Are there substantive changes in what the user sees?

- Is relevance ranking affected? If so, how?
- What further assessments can be made based on an analysis of these changes?

One method for gauging possible changes to the retrieval experience is to see how the previous search interface performed under a set of conditions and then to test the new search interface against these same conditions. (By the term “interface” here, I also imply the database behind it. Thus, “old interface” signifies the database of metadata records *prior* to changes in reprocessing and augmentation, while “new interface” alludes to the database of differently reprocessed and augmented metadata.) The set of conditions was derived from the IMLS DCC web logs, which track user queries. An informal usability test also was conducted, not only to receive some feedback but also to get a sense of whether this approach to answering the research questions was constructive, in which case the usability test could be expanded upon in future research.¹

Methodology

The research methodology involved a review of the transaction web logs for the IMLS DCC metadata repository dating from November 2006 to May 2007; during this period the old interface, which I will call “Interface A,” was being used. Thirty end-user queries were randomly selected from the web logs to create a test set of queries. (I chose this many queries in case some had to be eliminated, for whatever reason; as will be seen, modifications in the new advanced search necessitated a few eliminations.) Although search limits varied, half of these were multi-word queries, while the other half was made up of single-word queries. (Table I.) In general, according to this period of the web logs, roughly 60% of the queries were single-word, while roughly 40% were multi-word. The topics of the queries randomly chosen for this study ran the gamut, but they referenced mostly people, places, objects, and events.

Table I. Characteristics of queries selected for test runs in both interfaces.*

Kinds of queries	No. of queries	Avg. no. of characters/query	Avg. no. of characters/word	Avg. no. of words/query
Single-word, multi-parameter (e.g., “airport AND Pennsylvania OR Nebraska [author/artist/creator]”)	5	17	7	NA
Multi-word, multi-parameter (e.g., “Black Sunday [title/subject/description] OR Depression [title/subject/description] AND 1935 [date]”)	5	24	6	4
Single-word, single-parameter (e.g., “centennial”)	10	8	8	NA
Multi-word, single-parameter (e.g., “Lower East Side”)	10	15	6	2

*Brackets above enclose what were actual search limits, or parameters, for these sample queries.

Query terms were recorded, as were the search parameters for advanced search queries. In Interface A, for advanced search, users had the option of selecting a parameter by which to search—such as by “Title/Subject/Description,” “Author/Artist/Creator,” “Type,” “Date,” and “Publisher”—and the option of searching across all the collections in the registry or limiting the search to a specific collection (perhaps more than one, if the user wished). (Figure 1.) Optional Boolean operators took the form of “also must contain,” “or can contain,” and “must not contain.”

Figure 1. Advanced search in Interface A with input of “Civil War and battle.”

In the new interface, which I will call “Interface B,” the advanced search parameters are “Keyword/Phrase Anywhere,” “Author/Artist’s Last Name,” and “Title/Subject Words.” Users still may limit searches to specific collections or search across all of them. (Figure 2.) Boolean operators are not available for selection, however, in Interface B; rather, if more than one search limit is chosen, then the operator “and” is implicit.

Figure 2. Advanced search in Interface B.

The selected queries were input in Interface A, which is attached to an older database of metadata records (<http://imlsdcc.granger.uiuc.edu/oldsearch/search/searchform5.asp>). After each of the thirty queries (or searches), the first twenty results were documented in the following way: the URL for each IMLS DCC metadata record (which describes the object in the result) was copied and pasted into an Excel table, in the order in which they were listed on the results page.² (Figure 3.) In addition, I noted the following for each of the thirty queries: the number of hits returned in total; the number of collections associated with the first twenty hits; and the names of these associated collections.



Figure 3. First three results from query “Civil War and battle” in Interface A. Hot links to IMLS DCC metadata records are circled in red.

These steps were repeated with Interface B (a version of which resides at <http://ratri.grainger.uiuc.edu:8099>), which draws from a database of records reflecting modifications to the reprocessing and augmentation of the metadata. The URLs for the metadata records, linked from the results in Interface B, were entered in the same Excel table for side-by-side comparison with URLs from the old interface. Data for the number of hits, number of collections, and names of collections also were recorded. (Figure 4.)



Figure 4. First three results from the query “Civil War and battle” in Interface B, with hot links to full item records circled in red.

At this stage, I discovered that changes in the now streamlined advanced search made it impossible to reproduce four of the searches in Interface B, because the queries for them engaged the “type” and “date” limits, as well as the Boolean operators “or” and “not”—none of which is available for selection in the new advanced search. Accordingly, these four queries were eliminated, leaving a set of 26 query results from the two interfaces to compare, contrast, and study.

Results

The following tables display data for results returned and the collections associated with them.

Table II. Total number of hits and collections per query in both interfaces*

Query Terms	Interface A: No. of Hits	Interface A: No. of Associated Collections	Interface B: No. of Hits	Interface B: No. of Associated Collections
Trolley AND Pittsburgh	129	1	157	1
Sioux AND dance	17	5	17	4
Oregon AND sports	26	5	26	5
Firland AND sanatorium	19	1	19	1
Three Mormon towns AND Tocquerville	400	1	400	1
Civil War AND battle	110	12	107	10
immigration	194	22	186	16
Computer	436	16	422	10
gardening	225	23	223	15
fountains	727	15	727	13
centennial	292	23	292	15
circus	366	22	366	14
dogs	820	34	827	18
bicycle	771	30	774	18
wrestling	74	10	74	8
textiles	326	16	319	13
Frank Lloyd Wright	28	10	27	7
Puerto Rico	222	12	222	12
Lower East Side	152	12	149	10
Renaissance Society	67	3	66	2
Columbian Exposition	496	11	495	9
Erie Canal	46	4	46	4
Pike Place Market	33	1	33	1
lantern slides	104	9	108	6
performing arts	78	14	79	9
Tehama Indian	115	4	115	2

*In general, mainly because of changes in filtering, there will not be more than twenty collections, shown as associated with results, for any query in Interface B. Nonetheless, this data is shown in order to supplement the data in Table III for purposes of the discussion below.

Table III. Number of associated collections for first 20 returned hits*

Query Terms	Interface A: No. of Associated Collections for First 20 Hits	Interface B: No. of Associated Collections for First 20 Hits
Trolley AND Pittsburgh	1	1
Sioux AND dance	4	4
Oregon AND sports	3	4
Firland and sanatorium	1	1
Three Mormon towns AND Tocquerville	1	1
Civil War AND battle	5	5
immigration	5	9
computer	2	8
gardening	4	4
fountains	2	2
centennial	1	3
circus	2	1
dogs	2	6
bicycle	1	5
wrestling	5	5
textiles	2	7
Frank Lloyd Wright	6	7
Puerto Rico	3	4
Lower East Side	4	2
Renaissance Society	1	2
Columbian Exposition	5	2
Erie Canal	2	4
Pike Place Market	1	1
lantern slides	1	1
performing arts	1	4
Tehama Indian	1	1

*Six queries in this table are shaded because in both interfaces there is only one collection associated with the first twenty results, and the collection is the same one.

Discussion

As Table II reveals, twelve queries, or 46% of the query set, brought back an equal number of hits in each interface. Interface A showed more results for nine of the queries, or 35% of the query set, while Interface B showed more hits for five queries, or 19%. Thus, even though these queries were being run against the same database, more than half of them resulted in a different number of results.

Table II also shows the number of collections associated with the total number of results returned. Interface A has many more collections related to the results than does Interface B, even though both are drawing on the same database. These differences are due mainly to changes in filtering that have been applied to Interface B; in Interface B, items are no longer being related to sub-collections but, rather, to parent collections. These modifications mean that, although there may be more than twenty collections that relate to a query in Interface B, the end-user will not see more than twenty collections (the top twenty) corresponding with results. As can be discerned in Table II, none of the searches carried out in Interface B returned more than eighteen related collections.

The changes in how items are related to collections in Interface B have resulted in a lower rate of collection recall, which suggests a higher rate of precision in terms of collection relevance. The data appear to bear this out. Table III exhibits the number of collections associated with the

first twenty hits returned for each query. The table has six shaded rows to stand for queries that, in each interface, show only one collection. Because there is no difference in the quantity of related collections for both interfaces, these queries can be eliminated from further comparison and analysis here. This leaves a set of twenty queries. Among them, when the number of collections is noted for the first twenty hits per query in each interface, as Table III above has documented, then Interface B shows more collections associated with the first twenty results for twelve (out of twenty) queries, or 60%; in Interface A it is eight queries, or 40%.

The data are more dramatic when the number of related collections for the first twenty results is set against the total number of related collections for the total number of results. For example, in Interface A the query “computer” has two collections associated with the first twenty results, and for the overall list of results there are sixteen related collections; thus, for the first twenty results Interface A is pulling from only one-eighth, or 12.5%, of the corresponding collections. In Interface B the query “computer” has eight collections associated with the first twenty results and ten collections associated with the total number of hits; for this query Interface B is drawing on 80% of the related collections. For queries showing an increase in associated collections, Table IV below gives the percentage of the total number of collections corresponding with the first twenty results:

Table IV. Queries showing percentage of total number of associated collections

Query Terms	Interface A: % of collections in first twenty results	Interface B: % of collections in first twenty results
Oregon AND sports	60%	80%
immigration	23%	56%
computer	12.5%	80%
centennial	4%	20%
dogs	6%	33%
bicycle	3%	28%
textiles	12.5%	54%
Frank Lloyd Wright	60%	100%
Puerto Rico	25%	33%
Renaissance Society	33%	100%
Erie Canal	50%	100%
performing arts	7%	44%

A look at which collections are relevant to these results can also be valuable. Tables giving this information are provided in the appendix at the end of this paper (Tables A and B). To give a sense of the difference in frequency of associated collections between Interface A and Interface B, consider that the Historic Pittsburgh Collection and the Making of Modern Michigan Collection each appear thirteen times as associated collections in the query results for Interface A. In Interface B, however, Historic Pittsburgh shows up as an associated collection eight times in the query results, while the Making of Modern Michigan Collection appears just four times.

Given the above percentage increases in associated collections, which arguably points to increased collection relevance per query, the question then becomes: how is item relevance affected, if at all? What can be discerned about the ranking of items returned after a query?

If relevance can be understood, in the most basic sense, as a match between a query term and a term in the metadata record, then a close examination of the metadata records for a few results suggests enhanced relevance.³ For instance, the IMLS DCC metadata record for the first

ranked result in Interface A for the query “Erie Canal,” which returned an equal number of hits in both interfaces, shows the words “Erie” and “canal” in the description field, but they appear separately rather than as the phrase “Erie Canal.” In the metadata record for the second ranked result, in the same interface, “Erie Canal” and “Lake Erie” show up in the description field; this is the only instance among the first four results where the query term appears in complete form. In the third and fourth results, the words “canal” and “Erie” are present but only separately. Moreover, in these first four results in Interface A, “Erie Canal” (whether as a single phrase or as separate words) shows up in the metadata records only in the description field. The results are for items belonging to the Making of Modern Michigan Collection.⁴

On the other hand, the first four results for “Erie Canal” in Interface B reveal both an increased frequency of the phrase and differences in where the phrase appears in the metadata record. “Erie Canal” shows up in the title field in all four metadata records; it appears in the subject field for the first two records, while in the subsequent two records just the word “canals” is found in the subject field. (See Figure 5.) Thus, the first four results in Interface B for this query show complete term matching (i.e., “Erie” and “Canal” are a single phrase in the title field). The first three of these four results refer to items belonging to the Maine Music Box Collection; the fourth referenced item belongs to the Digital Past Collection.

The screenshot displays a search results page for the query "Erie Canal". At the top, it shows "1 - 20 of 45 results" and navigation options. The results are listed in a table-like format with the following entries:

- Item 1:** Title: **The Erie Canal : Low Bridge Arranged by Ernst Bacon**. Creator: Bacon, Ernst, Arranger. Subject: Donkeys, Ships, Songs with geographic locations, **Erie Canal**, Songs with piano, Men, Canals. Collection: Maine Music Box.
- Item 2:** Title: **Low Bridge! - Everybody Down : as Fifteen Years On The Erie Canal Words and music by Thos. S. Allen**. Creator: Allen, Thos. S., Composer (1876-1919). Subject: Bridges, **Canals**, Mules, Musical notation, Sheet music covers, Songs with piano, **Erie Canal**, Songs with geographic locations, Songs about work. Collection: Maine Music Box.
- Item 3:** Title: **Down by the Erie Canal lyric and music by Geo. M. Cohan**. Creator: Cohan, George M., Composer (1876-1942). Subject: Cohan and Harris, Artist, Cohan, George M., Lyricist (1876-1942), Songs with geographic locations, Cohan and Harris, Portraits, Songs with piano, **Erie Canal**. Collection: Maine Music Box.
- Item 4:** Title: **Rochester NY - Where the Erie Canal Crosses the Genesee, Apr24/71**. Type: Image. Subject: **Canals**, Rivers, Ships, Cityscapes. Collection: Digital Past.

Figure 5. Screen capture of first four results for query “Erie Canal” in Interface B. Note highlighting of entire query term in both title and subject fields.

Unlike in Interface A, where “Erie Canal” appeared in the description field, in Interface B there is no description field in the metadata records for the first four results. Thus, the title and subject fields seem to hold more weight in Interface B. (In the interest of full disclosure, the description field was one of the search limits omitted in the new interface, but the field does still appear in metadata records in Interface B—just not in these first four.) Further confirmation of this intuition is the fact that in Interface A, the first three items from the Maine Music Box Collection are ranked at the bottom (45th, 46th, and 44th, respectively), and the fourth result ranks 42nd (where, in the IMLS DCC metadata record for this item, “Erie Canal” appears in the title field). The rankings for three of the four above-mentioned items from the Making of Modern Michigan Collection have changed as dramatically in Interface B results for this query: the first, third, and fourth items are ranked as 46th, 44th, and 45th. Interestingly, the second-ranked result for “Erie Canal” in Interface A is ranked eighth in Interface B.

Similar disparities between result rankings in the two interfaces occur with other queries listed in Table IV. What are ranked as the first and second results for “Oregon AND sports” in Interface A are ranked 15th and 25th, respectively, in Interface B; the first two results for this query in Interface B are ranked 23rd and 24th in Interface A. For the query “performing arts,” which shows the sharpest difference in the number of associated collections, the first four results in Interface A rank as 53rd, 65th, 66th, and 67th, respectively, in Interface B. The first, third, and fourth results in Interface B rank 6th, 5th, and 8th in Interface A; the second-ranked result in Interface B is one not included at all in Interface A (which, with 78 hits, has one fewer result than Interface B, which lists 79). As with the query “Erie Canal,” for which the position of the query terms in the result seemed to influence relevance, in Interface A the query terms from “Oregon AND sports” and “performing arts” appeared in the top four results more frequently in the description field of the linked IMLS DCC metadata records, while Interface B showed matches occurring more often in the title and subject fields of a record.

Thus, in the context of item relevance the trend in Interface B reveals frequent and complete term matching in the title and subject fields for the top ranked results, while in Interface A term matching may be partial and tends to occur in the description field.

Usability Testing

Despite the tendency in the results for Interface B toward query term matching as a basic criterion for relevance, an end-user may have a different idea of what relevance is and means. To get a sense of whether this may be so, an informal usability test was arranged, once results from the query runs against both interfaces were tabulated. A simple web interface (Figure 6) was designed to display the list of queries, and each query was hyperlinked to an interface showing a side-by-side comparison of the first twenty results for that query. A library school classmate was recruited for participation in the usability test, chiefly to obtain evaluation and feedback on the query results. In the test the classmate was asked to choose a query; compare the list of the first twenty hits for that query in each interface; and then determine, as best she could, which list seemed to have brought back the most relevant results for the query.



Figure 6. Simple web interface used to select queries during informal usability testing.

The first query selected by the classmate was “gardening,” because—as she said—she knows something about the subject. In Interface A this query brought back 225 results, while in Interface B it brought back 223 results. One of the first observations made by the classmate concerned the difference in result rankings between the two interfaces, namely that the item “Pittsburgh Public School Students Gardening” (a photograph) is ranked ninth in Interface A but second in Interface B. She questioned why the same item had a lower relevance in Interface A, yet a higher relevance in Interface B. The classmate then looked more closely at the first result in Interface A, which is a chapter-long excerpt from a book called *Everyman’s House* by Caroline Bartlett Crane. The classmate launched the search-and-find function to see where the term “gardening” appears in this passage. It appears only once, about two-thirds into the text. At this point the classmate remarked that “this is a problem”—i.e., the fact that this item is ranked first but is not about gardening; she also noted that the item does not appear in Interface B. The classmate then clicked on the eighteenth-ranked result, “A View of the Fields of Gitarama,” which refers to a photograph of terraced gardening. This image struck her as a relevant result. Thus when she discovered that this result was missing entirely

in the first twenty hits in Interface B, despite its relevance to the query term, the classmate expressed puzzlement. (Incidentally, “A View of the Fields of Gitarama” is ranked 54th in Interface B; the query term “gardening” appears in the description field of its metadata record).

The classmate’s observations for the other queries she selected played out similarly to this one. In scanning the results for the third query she chose, which was “fountains,” she noticed that among the results in Interface B, the query term appears in the title and subject fields for the top-ranked results. Upon looking at the metadata records more closely, however, she found this supposed relevance odd, since these results were not examples of “fountains”—a word that, for her, evoked decorative water-bearing structures—but, rather, were results about an apartment complex with “fountains” in its name.

In short, the classmate could not say in the case of each query which interface yielded more relevant results. She seemed more intrigued by the question of what was driving relevance in the results for the two interfaces than by the act of deciding which interface presented results more relevant to the query. This preoccupation may have been caused by the fact that these queries were not ones devised by her, compelling her to “figure” them out. More than once she commented on how a result did not jibe with her perception of what the query should bring back.

Conclusion

In this brief study I attempted to find changes, whether incremental or acute, in the search and retrieval process as carried out in two different search interfaces created for the IMLS DCC metadata repository. The queries comprising my test set yielded results in Interface B that definitely departed from those produced in Interface A. In Interface B, result rankings had shifted—even radically, in some cases—and query term matching, which tended to be more complete or precise than in Interface A, seemed most prevalent in the title and subject fields. The description field appeared to have more weight in Interface A, however. The usability test subject definitely noticed differences in the result rankings between the interfaces. In addition, changes in item-level ranking and relevance may be related to the differences between associated collections (in terms of both their quantity and type—i.e., which collection) for results in Interface A and in Interface B, particularly as noted in Tables II, III, and IV. In other words, in Interface B it does not seem possible to have an increase in the number of associated collections for the first twenty results after a query, without changes in result ranking in this same interface occurring as well. Interface B also displayed a deeper sense of collection context, created by the smaller number of associated collections per set of query results; collection depth, rather than breadth, constitutes a significant difference between the two interfaces.

What these changes assert about relevance is harder to determine, perhaps as especially evidenced by the informal usability test I conducted. Query term matching makes for relevance at a rudimentary level. The matches (between terms in the query and terms in the metadata record) detected by my usability test subject did not convince her of relevance; she also was not satisfied by the rankings in the query results. Essentially, she had her own expectations of relevance that were not met by the query results or the rankings she reviewed. What is needed is a more sophisticated approach to relevance measurement that enables test subjects to acquaint themselves with the IMLS DCC repository first, perhaps through a period of extended browsing, in preparation for launching their own searches—i.e., searches *relevant to them*, to begin with, and which they care about. (For instance, the subject in my usability test based her first query selection, “gardening,” on her knowledge of the topic, which suggests further the subjective nature of relevance judgment.) In

addition, in such a test situation, it may be valuable to permit subjects to refine one or two of their searches or queries, because the process of narrowing a search term can lead to a more granular sense of relevance and thus a better idea of what a result should look like.⁵ (At the same time, it would be necessary to balance this qualitative component with quantitative instruments of measurement.) Moreover, the practice of refinement allows test subjects a chance to get to know the workings of a user interface more completely.

The question of how one understands and begins to employ a user interface—of what constitutes user interface literacy—is topmost in the minds of usability experts.⁶ An enhanced command of this question would be constructive for what the IMLS DCC project is aiming to achieve in terms of making its repository more usable. Models for designing usability tests to study the search and retrieval process in a digital context, which this study could look to, are definitely available, but my sense is that they may be more prevalent in the world of web search engine development, which is different from the world of a digital library or a digital collection project. “Use” itself means one thing in the context of search via a web search engine and another thing in the context of search in a digital collection, even one as heterogeneous in content as the IMLS DCC project (heterogeneity being one of the commonalities between what is searched by web search engines and what is searched in aggregations of digital collections). Thus, in dedicating attention to relevance through a more rigorously designed usability test, I also would want to consider as thoroughly as possible the areas where usability testing of web search engines helps in the context of a digital collection and areas where it falls short.

While the foregoing would support a qualitative measurement of differences between the two interfaces, there could be changes in the way that quantitative assessments are pursued as well. For instance, if I were to undertake this study again, or carry it further, I would not only look more closely at ranking differences (by recording where the first ten rankings in Interface A are in Interface B, for instance, and vice versa) but also try to discern patterns beyond what I found this time around (e.g., with respect to the title, subject, and description fields). I might focus more on single-word queries, since they are the most prevalent, or at least replicate the proportion of single-word queries to multi-word queries in my test set. In addition, I would research library and information science literature for more guidance about constructing a set of sample queries that is free from as much bias as possible.

Another future step or direction to take is to investigate the search algorithm employed by the project, which could introduce promising leads about recall, precision, and relevance. The question of weighting by title and subject fields (which query results tended to show) especially makes the exploration of algorithmic approaches compelling. However, attention to algorithms may make more sense after study results are in.

All in all, I found this independent study a valuable learning experience. It cast into sharp focus the need to make metadata “work” for users of digital collections and the possibilities through which such a need may be researched. Part of the challenge for an aggregated digital collection project such as the IMLS DCC is the diversity of the content, which points up the heterogeneity of the users of the content—which makes understanding those users, so as to serve them better, an almost Sisyphean task for librarians. Nonetheless, there is solid evidence that aggregated metadata can be applied toward the creation of digital library services (Cole & Foulonneau, 2007), and the IMLS DCC project affirms this. Increased attention to, and study of, user needs in a digital collection context will help make it even more so.

References

Arms, Caroline R., and William Y. Arms. "Mixed Content and Mixed Metadata: Information Discovery in a Messy World." *Metadata in Practice*. Chicago: ALA, 2004. 223-237.

Cole, Timothy W. and Muriel Foulonneau. "Using Aggregated Metadata to Build Digital Library Services." *Using the Open Archives Initiative Protocol for Metadata Harvesting*. Westport, Conn.: Libraries Unlimited, 2007. 161-183.

Dushay, Naomi, and Diane Hillmann. "Analyzing Metadata for Effective Use and Re-use." September 2003. *DCMI Conference Papers 2002-2005*. 30 July 2007. <http://www.nl.go.kr/dcpapers/501_Paper24.pdf>.

Foulonneau, Muriel, Timothy W. Cole, Thomas G. Habing, and Sarah L. Shreeves. "Using Collection Descriptions to Enhance an Aggregation of Harvested Item-Level Metadata." *International Conference on Digital Libraries. Proceedings of the 5th ACM/IEEE-CS Joint Conference on Digital Libraries*. New York: ACM Press, 2005. 32-41.

Hillmann, Diane, Naomi Dushay, and Jon Phipps. "Improving Metadata Quality: Augmentation and Recombination." October 2004. *DCMI Conference Papers 2002-2005*. 30 July 2007. <http://www.nl.go.kr/dcpapers/Paper21_pdf>.

Leighton, H. Vernon, and Jaideep Srivastava. "First 20 Precision among World Wide Web Search Services (Search Engines)." *Journal of the American Society for Information Science* 50.10 (1997): 870-881.

Russell, Daniel M. "What Are They Thinking: Searching for the Mind of the Searcher." 20 June 2007. Presentation at the Joint Conference on Digital Libraries. 30 July 2007. <<http://dmrussell.googlepages.com/JDCL-talk-June-2007-short.pdf>>.

References Consulted

Rose, Daniel E. "Reconciling Information-Seeking Behavior with Search User Interfaces for the Web." *Journal of the American Society for Information Science*. 57.6 (2006): 797-799.

Rose, Daniel E., Daniel Orr, and Raj Gopal Prasad Kantamneni. "Summary Attributes and Perceived Search Quality." *Proceedings of the 16th International Conference on World Wide Web*. New York: ACM Press, 2007. 1201-1202.

Rose, and Danny Levinson. "Understanding User Goals in Web Search." *Proceedings of the 13th International Conference on World Wide Web*. New York: ACM Press, 2004. 13-19.

Appendix

Table A: Collections Associated with First Twenty Results in Interface A

Query Term (s)	No. of Collections	Names of Collections
Trolley AND Pittsburgh	1	Historic Pittsburgh
Sioux AND dance	4	Heritage Colorado, Amer. Missionary Assoc., Western Waters DL, Amer. Journeys
Oregon AND sports	3	King Co. Snapshots, Columbia River Basin Ethnic Archive, Western Waters DL
Firland and sanatorium	1	King Co. Snapshots
Three Mormon towns AND Tocquerville	1	Museums and Online Archive of California (MOAC)
Civil War AND battle	6	Digital Past, Illinois Alive!, IMLS DCC, NC Experience, Southern Homefront
immigration	5	Making of Modern Michigan, Historic Pittsburgh, Heritage CO, Digital Past, Florida Environments Online
computer	2	Making of Modern Michigan, Historic Pittsburgh
gardening		Making of Modern Michigan, Historic Pittsburgh, Africa Focus, Heritage CO
fountains	2	Making of Modern Michigan, Historic Pittsburgh
centennial	1	Making of Modern Michigan
circus	1	Making of Modern Michigan
dogs	1	Making of Modern Michigan, Historic Pittsburgh
bicycle	1	Making of Modern Michigan
wrestling	5	Making of Modern Michigan, Historic Pittsburgh, Africa Focus, Heritage Colorado, Digital Past
textiles	2	Historic Pittsburgh, Africa Focus
Frank Lloyd Wright	6	Making of Modern Michigan, Historic Pittsburgh, Education by Design, Heritage Colorado, Charles W. Cushman, Digital Past
Puerto Rico	3	Making of Modern Michigan, Digital Past, Flora and Fauna of the Great Lakes
Lower East Side	4	Making of Modern Michigan, Historic Pittsburgh, Heritage Colorado, Charles W. Cushman
Renaissance Society	1	Historic Pittsburgh
Columbian Exposition	5	Feeding America, Historic Pittsburgh, Heritage Colorado, Charles W. Cushman, Digital Past
Erie Canal	2	Making of Modern Michigan, Historic Pittsburgh
Pike Place Market	1	King Co. Snapshots
Lantern slides	1	Education by Design
Performing arts	1	Historic Pittsburgh
Tehama Indian	1	Charles W. Cushman, MOAC

Table B: Collections Associated with First Twenty Results in Interface B

Query Term (s)	No. of Collections	Names of Collections
Trolley AND Pittsburgh	1	Historic Pittsburgh
Sioux AND dance	4	Amer. Missionary Assoc., Heritage West, Western Waters DL, Amer. Journeys
Oregon AND sports	4	Western Waters DL, King Co. Snapshots, Columbia River Basin Ethnic Archive, Museum and Online Archive of California (MOAC)
Firland and sanatorium	1	King Co. Snapshots
Three Mormon towns AND Tocquerville	1	MOAC
Civil War AND battle	5	Teaching with Digital Content, Western Waters DL, Digital Past, Southern Homefront, Illinois Alive!
immigration	9	King Co. Snapshots, Teaching with Digital Content, Western Waters DL, Illinois Alive!, MOAC, Amer. Missionary Assoc., Columbia River Basin Ethnic History Archive, Central Florida Memory, North Carolina in Black and White
computer	8	Teaching with Digital Content, INFOMINE, Digital Past, Western Waters DL, Historic Pittsburgh, King Co. Snapshots, Heritage West, Making of Modern Michigan,
gardening	4	Western Waters DL, Historic Pittsburgh, MOAC, Western Trails
fountains	2	Digital Past, Charles W. Cushman
centennial	3	Heritage West, Digital Past, Western Waters DL
circus	1	Charles W. Cushman
dogs	6	Charles W. Cushman, Historic Pittsburgh, Western Trails, Digital Past, Teaching with Digital Content, Honore Daumier Lithographs
bicycle	5	Digital Past, Charles W. Cushman, King Co. Snapshots, Western Waters DL, Central Florida Memory
wrestling	5	King Co. Snapshots, Western Waters DL, Africa Focus, Digital Past, Heritage West
textiles	7	Heritage West, Africa Focus, MOAC, Historic Pittsburgh, Tennessee Documentary History, Digital Past, Charles W. Cushman
Frank Lloyd Wright	7	Charles W. Cushman, Heritage West, Historic Pittsburgh, Making of Modern Michigan, Digital Past, MOAC, Education by Design,
Puerto Rico	4	INFOMINE, Western Waters DL, Digital Past, MOAC
Lower East Side	2	Charles W. Cushman, Amer. Missionary Assoc.
Renaissance Society	2	Historic Pittsburgh, Heritage West
Columbian Exposition		Teaching with Digital Content, Digital Past
Erie Canal	4	Maine Music Box, Digital Past, Making of Modern Michigan, Historic Pittsburgh
Pike Place Market	1	King Co. Snapshots
lantern slides	1	Education by Design
performing arts	4	Historic Pittsburgh, Heritage West, Teaching with Digital Content, Africa Focus
Tehama Indian	1	Charles W. Cushman, MOAC

¹ I should add here that the new interface also reflects changes in the search algorithm, which I did not explore in this study—not for lack of interest but for lack of time.

² I analyzed ranking (and, by extension, relevance) in the first twenty results returned after a query, but if I were to continue this study, I might look more closely at the concept of precision as explored by Leighton and Srivastava (1997). In their investigation of the performance of several web search engines popular at the time, Leighton and Srivastava constructed a test suite of fifteen questions (collected from actual queries from the undergraduate reference desk at their institution) and devised a measurement they called “first 20 precision,” which “rates the services based on the percentage of results within the first 20 that returned that were relevant” (Leighton & Srivastava, 1997). Since the concept of precision is crucial to that of relevance, it may well be worthwhile to explore ways to measure precision in the context of the results retrieved for searches executed in the IMLS DCC metadata repository.

³ Obviously, relevance—as an intensely researched and even hotly debated topic—betrays the conceptual simplification of it that I offer in passing above. However, I carried out a short-term research project. If I were to expand on what I did in a future study, then I would attend to the definition and issue of relevance much more explicitly and thoroughly. I elaborate on this point in the “Conclusion” section of this paper.

⁴ In truth, results for each query run in Interface B showed rankings different from results for queries run in Interface A. I focus on the queries in Table IV mainly in an attempt to imply the influence of collection association on result relevance and ranking.

⁵ The act of scanning a results page—something intuitively done during a search process—not only expands a user’s ability to judge relevance but also to refine one’s search terms. In the usability test for this study, the test subject was observed scanning the results and comparing them between the two interfaces, but the limitation of not being able to probe further (i.e., by refining the search term) may have hindered her from determining which list of results was more relevant (and thus which interface more useful).

⁶ The idea of user interface literacy is borrowed from a keynote address presented at the 2007 Joint Conference on Digital Libraries, given by Daniel M. Russell, an Uber Tech Lead for Search Quality & User Happiness at Google, Inc. Discussing at length the question of what people are really doing when they are searching, Russell proposes an intensive investigation of what might be called the “gestalt” of the user search experience—the meso, micro, and macro levels of information-seeking behavior: “meso” stands for the “field study” context of the search process (how users in the real world are searching for information); “micro” represents the observation of the user at a microscopic level (such as tracking eye movement); and “macro” is the attempt to reconcile what the search logs disclose with actual behavior (bringing “outside behavior back to where we can see the signals”). See D.M. Russell, “What Are They Thinking: Searching for the Mind of the Searcher,” <http://dmrussell.googlepages.com/JDCL-talk-June-2007-short.pdf>. Retrieved 30 July 2007.