Unraveling the (Search) String: Assessing Library Discovery Layers Using Patron Queries

Robert Heaton and Liz Woolcott
Utah State University, USA

Introduction

Prompted by a desire to optimize configuration options of their current discovery tool, train and support users better, and systematically assess potential migration options, the researchers formulated two research questions pertaining to that tool: *What are users' general strategies?* and *How effective are those strategies?* These questions are not tidily separable from one another. Users' search behaviors are influenced by the goals they bring to the tool and by the options a tool offers. The effectiveness of a search strategy depends on the tool's indexing and the library's holdings but also on the specificity and accuracy of the search queries used. The researchers do not pretend to an ultimate resolution of these conflicts and the unknown weights of the relative factors but instead hope to decrease to some degree what is uncertain or unknown in these areas.

The setting of this study is Utah State University (USU), a mid-sized land-grant institution located in Logan, Utah. As of fall 2020, USU's enrollment is 22,072 FTE. The Libraries' budget is roughly $12 million, with $6 million dedicated to materials. Its current integrated library system is Innovative Interfaces' Sierra, with the discovery layer Encore Duet. The latter is the result of a partnership between Innovative and EBSCO, interleaving catalog results from the Libraries' Sierra database with article results from EBSCO Discovery Service (EDS).

Methodology

The researchers chose to conduct a log analysis of Encore URLs to approximate answers to the research questions posed. USU Libraries' implementation of Encore incorporates MARC-based records from the catalog (Sierra) alongside non-MARC records (via EDS) from 51 databases. Approximately 2.5 million MARC records and 3.6 million non-MARC records can be found through Encore. The Libraries' homepage features Encore as a single search box. While users have the option to go directly to the library catalog by clicking a link entitled "Books & Media" on the lower half of the webpage, this method is not used very often due to the prominence of Encore's single search box on main section of the webpage. Additionally, the search box does not include an option for advanced search until an initial search is run— and no browse searching is available. This makes it the primary entrance point for users.

When a user conducts a search in Encore, a URL is generated that includes information about the search parameters used by the patron, including elements such as the search terms used, facets selected, advanced searches chosen, or pages visited. Each action a user takes, from selecting a facet to changing a search term or selecting a record to view, will generate a new URL. The researchers ran a report in Google Analytics to download all URLs generated outside of the library's staff IP range on three dates in fall semester 2019. A Monday, Tuesday, or Thursday was selected from each of weeks 12, 13, and 15 of a 16-week semester. These days of the week were chosen because they represented the days with the highest traffic. The latter half of the semester was chosen to minimize the impact on the data of library instruction taking place outside the library.

Within Google Analytics, the researchers created a customized report that time stamped each URL down to the minute. This data was downloaded early in the morning after the date selected and imported into Airtable, a relational database used to code and analyze the data. All URLs were assigned a unique ID

within Airtable, then sorted to segregate URLs that lead to search results from URLs that lead to other features of Encore, including record pages, the homepage, or the advanced search page. Search result URLs were uploaded into Octoparse, a web scraping tool, and each URL was scraped for key data such as the search term used, the number of results, and the title and URL of each result displayed on the page. These results were numbered in the order they appeared and then uploaded into Airtable and assigned a unique ID for search results. Over all three days, 3,604 URLs were pulled from Google Analytics and 68,022 search results were scraped from results lists. These results were coded in three processes within Airtable: URL content, search queries, and known items.

URL content coding included assigning codes to the 3,604 URLs that were pulled from Google Analytics in order to identify the page type that the URL represented as well as any search terms that were used, whether or not advanced search fields or facets were present and which ones, and the page number.

Search query coding included grouping URLs into sessions by identifying similarities in the search parameters and search terms embedded in the URL, and taking into account the time stamps. To do this, the researchers reran the searches in both Encore and a general search engine to identify potential relationships in the search terms that may not be obvious. The URLs within each group were then grouped together and assigned session IDs, then numbered in the order that the researchers anticipated the user took. To provide quality control, the researchers recreated each search session's steps in Encore to verify that the navigation would generate the exact URLs in the same exact order. Once verified, each URL was coded with the action that was taken by the user (selecting a facet, clicking the advanced search, clicking the back button, etc.). Finally, the search terms in each URL were coded for the type of keywords and the order they appeared, as well as what kind of a search it was (known item, topical, etc.), and if it was a known item, the title of the item was recorded.

Known item coding involved taking the list of known items identified by the search query coding and verifying whether the items were available through the library. The researchers coded each item for availability, physical or electronic format, location of material, which entity indexed the item (Sierra or a subscription database), the final content provider, the number of steps it took at access the item if it was electronically available, and the placement of each known item in the search results lists displayed to the user.

Findings

General Observations

Four general categories of search type were identified from the data set: known items, topical, named entity, and other. Known items were coded when a specific book, article, or other resource could be discerned from the user's search. As is the case in other categories as well, user queries varied in their level of specificity. Actual examples of known-item queries include *Macroeconomics: Private and Public Choice by Gwartney, Stroup, Sobel, and Macpherson 15th edition*; *kenneth heineman campus wars*; and *mozart piano sonata no 18 in d major*. Topical searches aimed to retrieve information on a subject, without a specific resource in mind. These used queries such as *traffic analysis logan utah*, *nouns*, and *leaf NDVI transmission*. Named-entity searches reflected a user's interest in some person, place, or thing, whether an author, a geographic focus of research, or topic of study, as long as no recognizable known item shared its name. Examples of this type of search include *Shanghai International Settlement*, *John Steiner*, and *milgram*. A miscellaneous "other" category consisted mostly of resource-specific URLs pasted into the search box.

Known items were the most common search type, appearing in 54.8% of sessions, followed by topical, in 37.9%, then named entity, in 18.4%; other types occurred in just .005% of sessions (Figure 1). Note that a session might include more than one search type, so the total is higher than 100%. Within known items,
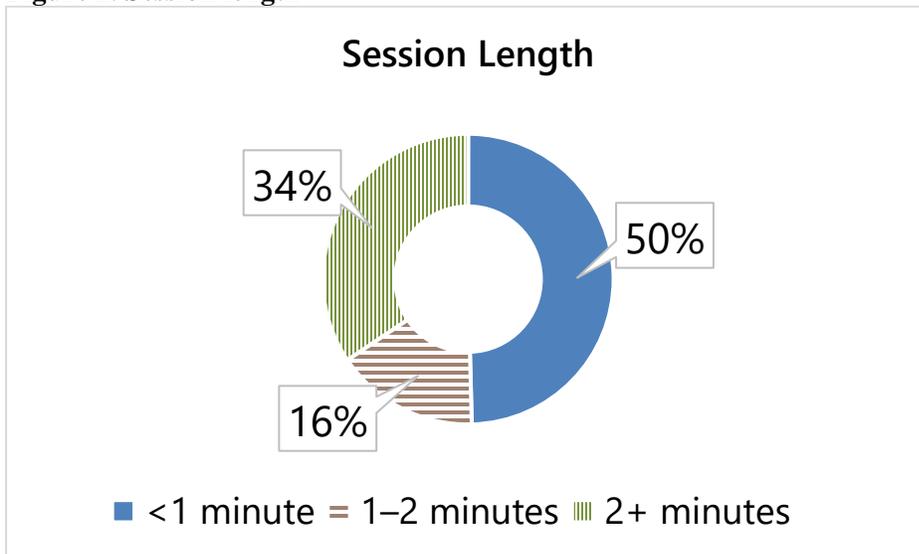
articles (43.3% of known-item sessions) and books (42.2%) were sought in roughly equal proportion, with less interest in other formats (14.5%).

Figure 1. Number of sessions by type of search

Number of Sessions by Type of Search

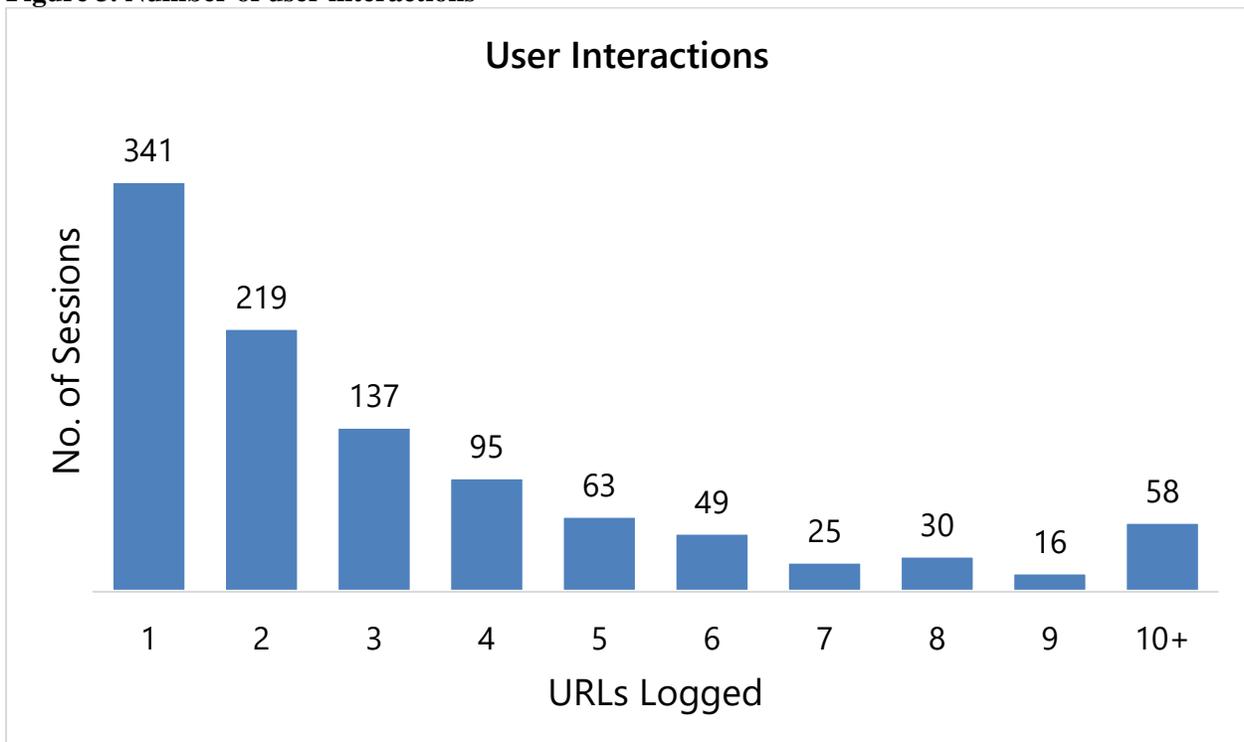| Type of Search	No. of Search Sessions
Known Item	570
Topical	395
Named Entity	192
Other	6
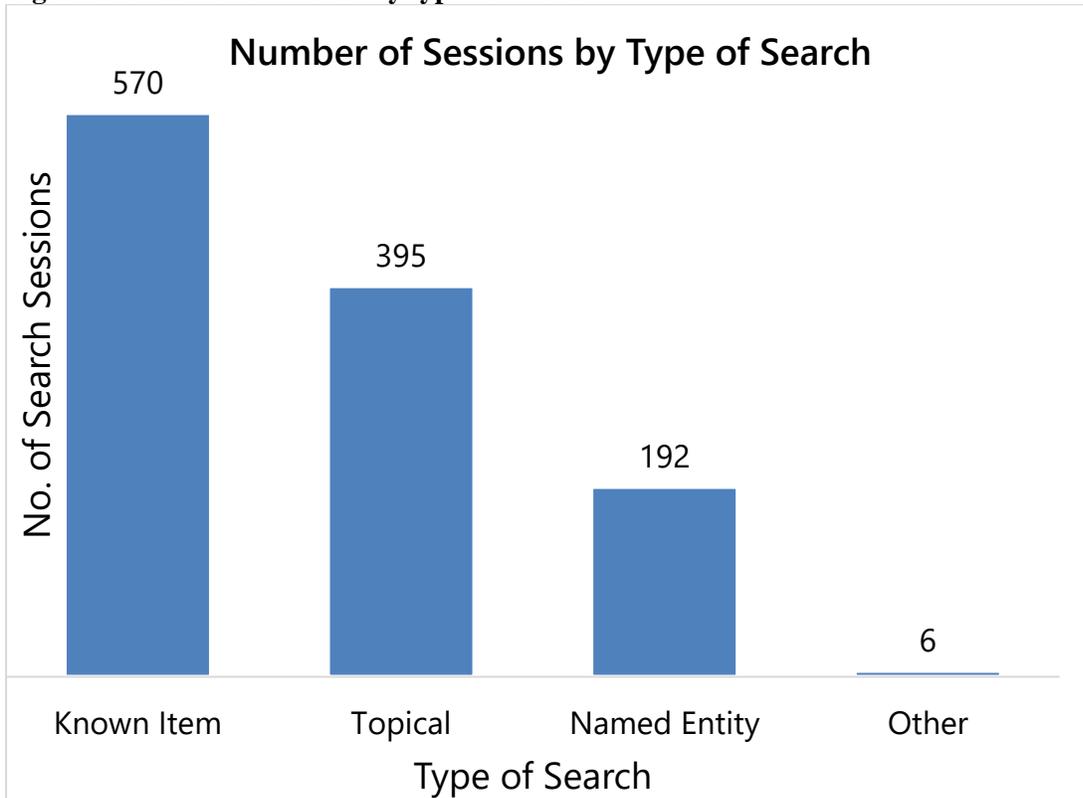
User sessions most often were brief and contained few interactions. Just one-third of sessions were two minutes or longer, and one-half were one minute or shorter (Figure 2). Because of the long tail of longer sessions, the mean session length is around five minutes, but because of the prevalence of extremely short sessions, the median time was one minute or less. A similar trend toward brevity appears when counting pages visited. Each user interaction generates a new URL, and Google Analytics captured the number of URLs visited per session (Figure 3). Most often (33%), users ran a single search; the majority of the time (54%), they visited 1–2 URLs, which might entail running a search and clicking a facet or running a search and viewing a detailed record. Some users conducted in-depth searches, but longer sessions with multiple search adjustments happened much less often than very short sessions.
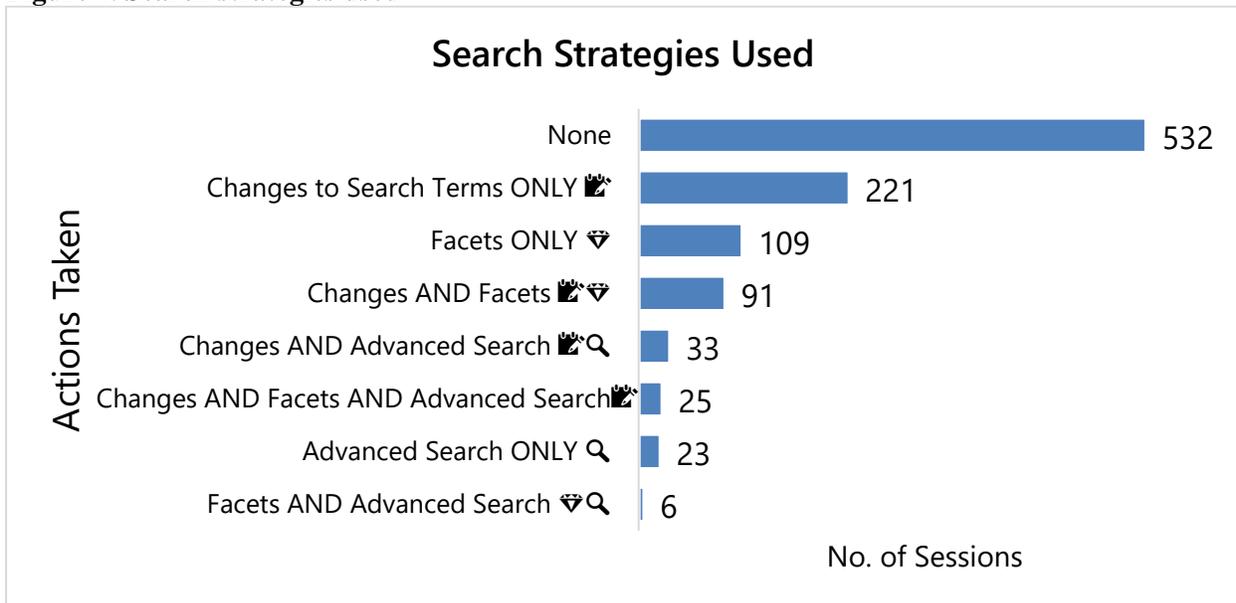
Figure 2. Session length

Session Length

34%

50%

16%

■ <1 minute ≡ 1–2 minutes ‖ 2+ minutes

Figure 3. Number of user interactions

User Interactions

No. of Sessions

341

219

137

95

63

49

25

30

16

58

1   2   3   4   5   6   7   8   9   10+
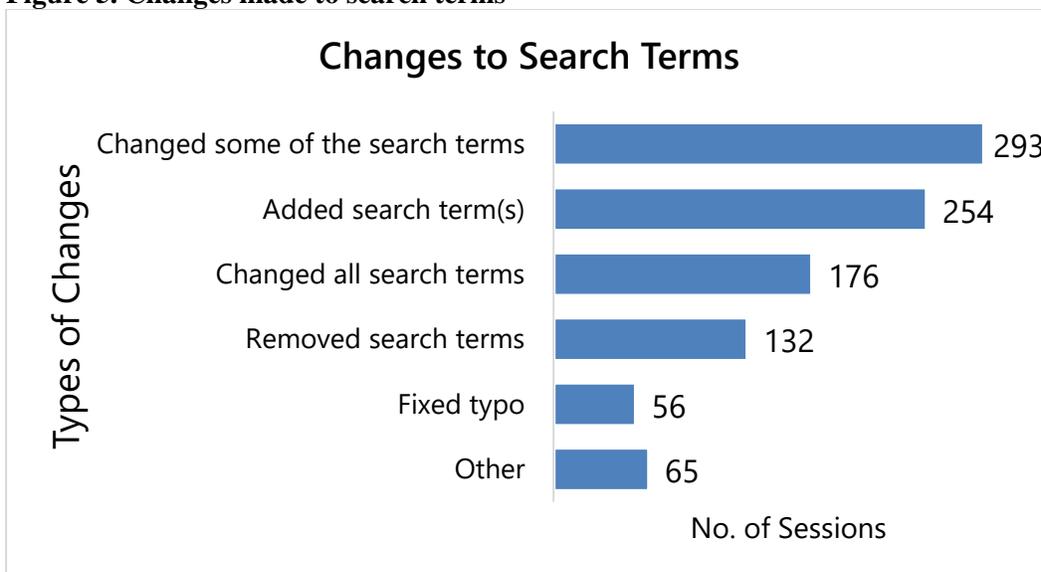
URLs Logged

Search Strategies

Three main search strategies emerged: adjustments to search terms, selection of facets, and use of advanced search. Most often (51% of sessions), users simply searched and viewed results; 21% only changed search terms; 11% selected facets only; and 9% changed their search terms and selected facets (Figure 4). Only 9% of all search sessions included the use of advanced search, generally in combination with other strategies.

Figure 4. Search strategies used

Search Strategies Used

Actions Taken	No. of Sessions
None	532
Changes to Search Terms ONLY	221
Facets ONLY	109
Changes AND Facets	91
Changes AND Advanced Search	33
Changes AND Facets AND Advanced Search	25
Advanced Search ONLY	23
Facets AND Advanced Search	6

Among search strategies, users most often (35.6% of sessions) made adjustments to their search terms (sometimes in combination with other strategies). As users reviewed the number of results returned by a query, as well as the nature of specific results returned, they made a variety of changes to elements of the query while apparently retaining the original search's intent. General categories of the changes made including changing some of the terms (30% of sessions where this strategy was used), adding search terms (26%), changing all of the search terms (18%), removing search terms (13.5%), and fixing typographical errors (5.7%) (Figure 5). Just 6.7% of these sessions applied other changes, such as changing capitalization; adding or removing quotation marks, punctuation or spaces; spelling out acronyms; and manually adding a Boolean operator or asterisk.

Figure 5. Changes made to search terms

Changes to Search Terms

Types of Changes	No. of Sessions
Changed some of the search terms	293
Added search term(s)	254
Changed all search terms	176
Removed search terms	132
Fixed typo	56
Other	65

The likelihood that users would change their search terms differed by the type of search being conducted. Users were more likely to adjust search terms when doing a topical search (users changed terms in 41.6%
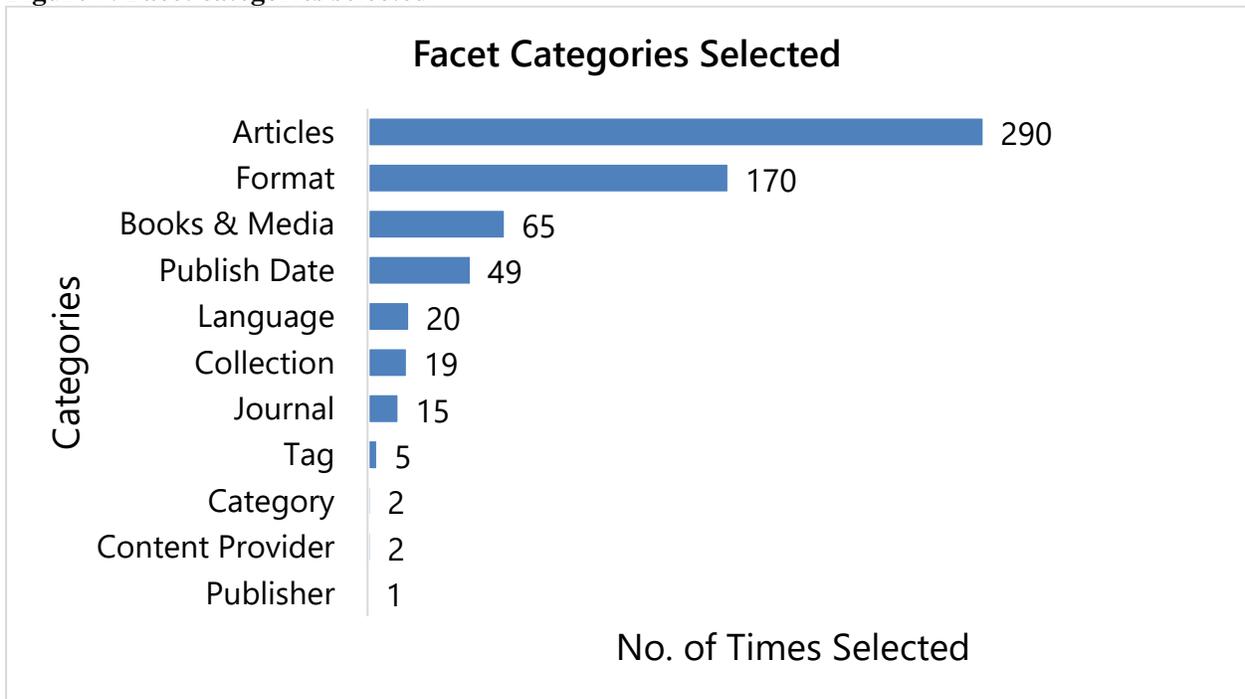
of topical searches) than a known entity (23.1% of those searches) or named entity (22.3%) (Figure 6). This trend is to be expected because topical searches are more likely to be exploratory, flexible, and iterative.

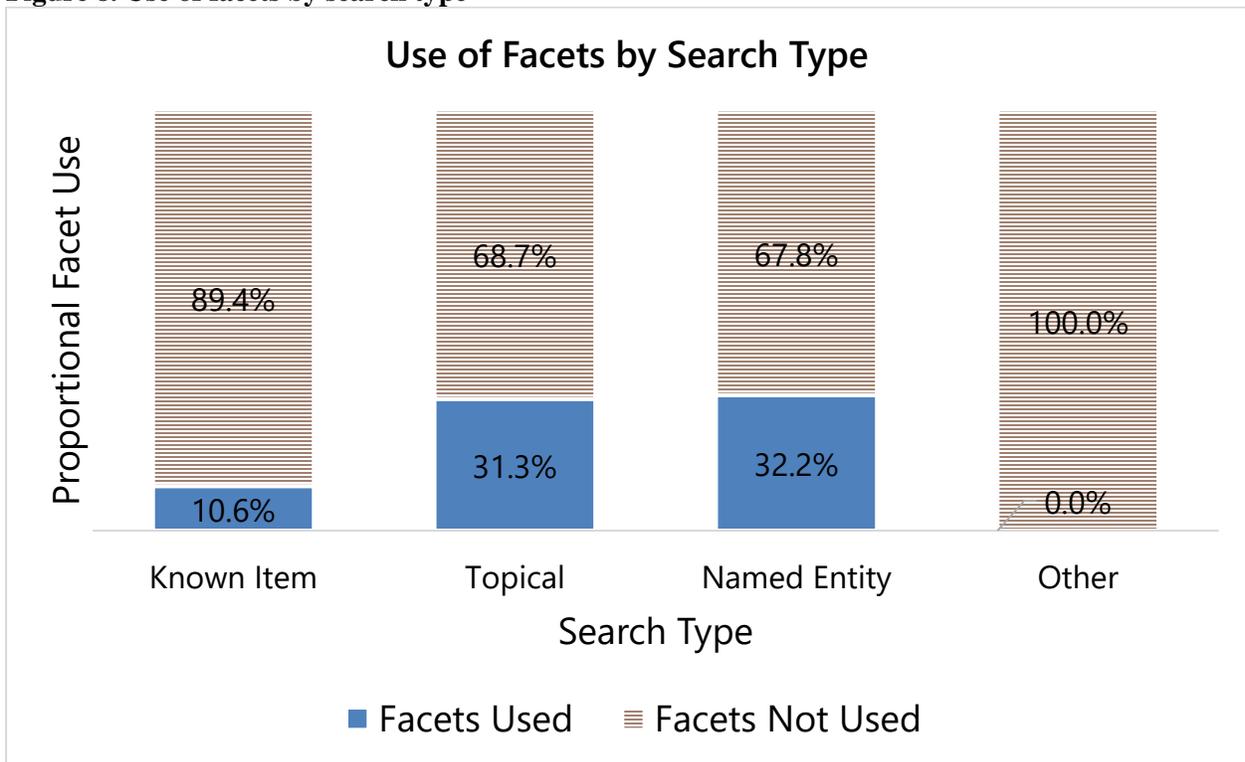**Figure 6. Use of search term changes by search type**
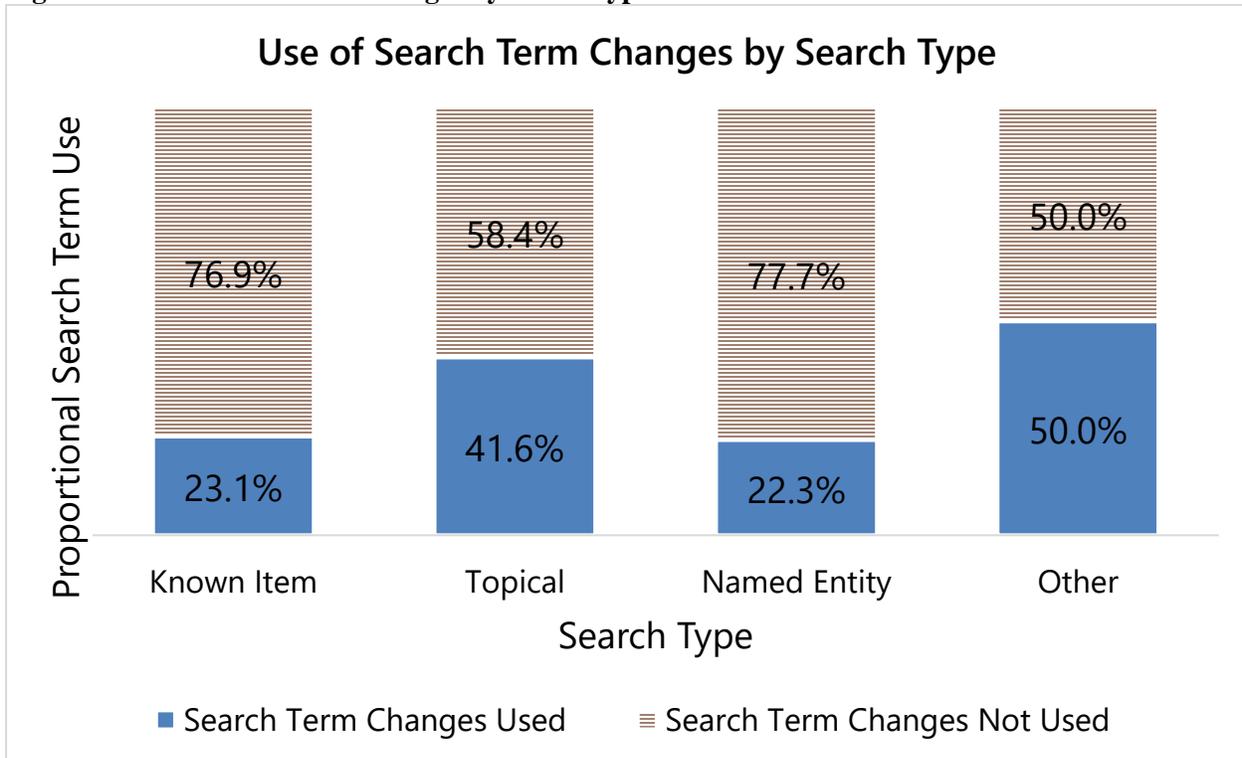
Applying facets was users' second-preferred search strategy, employed in 22.3% of search sessions. In Encore, only those refinements relevant to the current query appear as facets. Therefore, by way of example, if no catalog results were returned (listed as "Books & Media"), it would not be present for users to select. An exhaustive list of individual facets was condensed into eleven facet categories. For example, the "Publisher" category includes such entries as "public library of science," "taylor & francis ltd," "biomed central," and "american library association." The "Publish Date" facet is unique because users can enter a date range rather than select from specific options. Article facets were selected most often (45.5% of sessions where facets were used), followed by format facets (26.6%), followed by "Books & Media" (10.2%) and other categories (Figure 7). Among the 290 sessions where users chose article-category facets, limiting to articles themselves was most selected (153), but limiting to peer-reviewed sources was popular as well (97), followed by limiting full text (73). These sometimes were selected in combination with one another, so the total of individual facet use is higher than the number of sessions that used article-category facets. In the category of format facets, academic journals—which is another way of limiting to articles—was by far the most-selected option (115); other formats were selected much less frequently, such as e-books (18), magazines (12), print books (8), and others. The next-most-popular facets, "Books & Media" and "Publish Date," are standalone facets rather than having subcategories. After those, the language category was used most often to select English (15), with some interest in Spanish (3), French (1), Portuguese (1), and German (1).

Figure 7. Facet categories selected

Facet Categories Selected

| Category	No. of Times Selected
Articles	290
Format	170
Books & Media	65
Publish Date	49
Language	20
Collection	19
Journal	15
Tag	5
Category	2
Content Provider	2
Publisher	1

The use of facets depended to a degree on the type of search being performed. Sessions that included topical (31.3%) or named entity (32.2%) searches were much more likely to use facets than were known item (10.6%) or other (0%) searches (Figure 8).

Figure 8. Use of facets by search type

Use of Facets by Search Type

| Search Type	Facets Used	Facets Not Used
Known Item	10.6%	89.4%
Topical	31.3%	68.7%
Named Entity	32.2%	67.8%
Other	0.0%	100.0%

The third and least popular search strategy was advanced searching, applied in under 9% of all searches. This low utilization rate was unsurprising because of how this link is situated in the Utah State University Libraries' web presence. On the library homepage's current design, there is no direct link to the advanced search, just the prominent simple-search box. Users must run a search first, and then from the results page click the "Advanced Search" link immediately below the search box. Advanced-search options include fields to search (Keyword, Title, Subject, Author), Boolean connectors among these fields (AND, OR, NOT), dropdown menus with multiple options each (Format, Collection, Language), and a free-text Date field. Users chose these in various combinations, such as a search of Author AND Title. The total number of sessions in which each option was applied is shown in Figure 9. The default field, Keyword, and the default Boolean operator, AND, were included most often; the next-most-common field, Author, is used one-third less than Keyword, followed by Title and, much less used, Subject. The dropdown menus and date were used quite infrequently.

**Figure 9. Advanced search options applied**
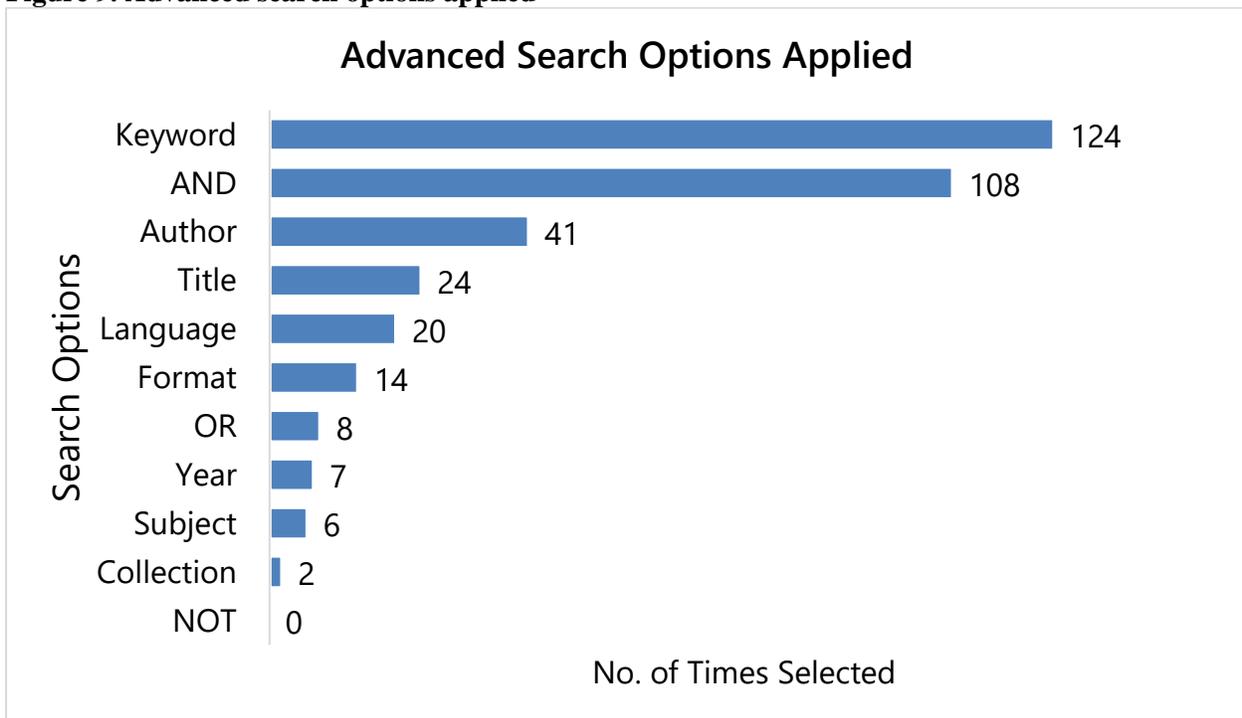
Because advanced searching was so rarely used, it is difficult to pinpoint trends in its use. However, topical (10.8%) and named entity (11.2%) searches were twice as likely to involve an advanced search as were known item (5.3%) searches, and this strategy never appeared in the data set's "other" searches (Figure 10).

Figure 10. Use of advanced search by search type
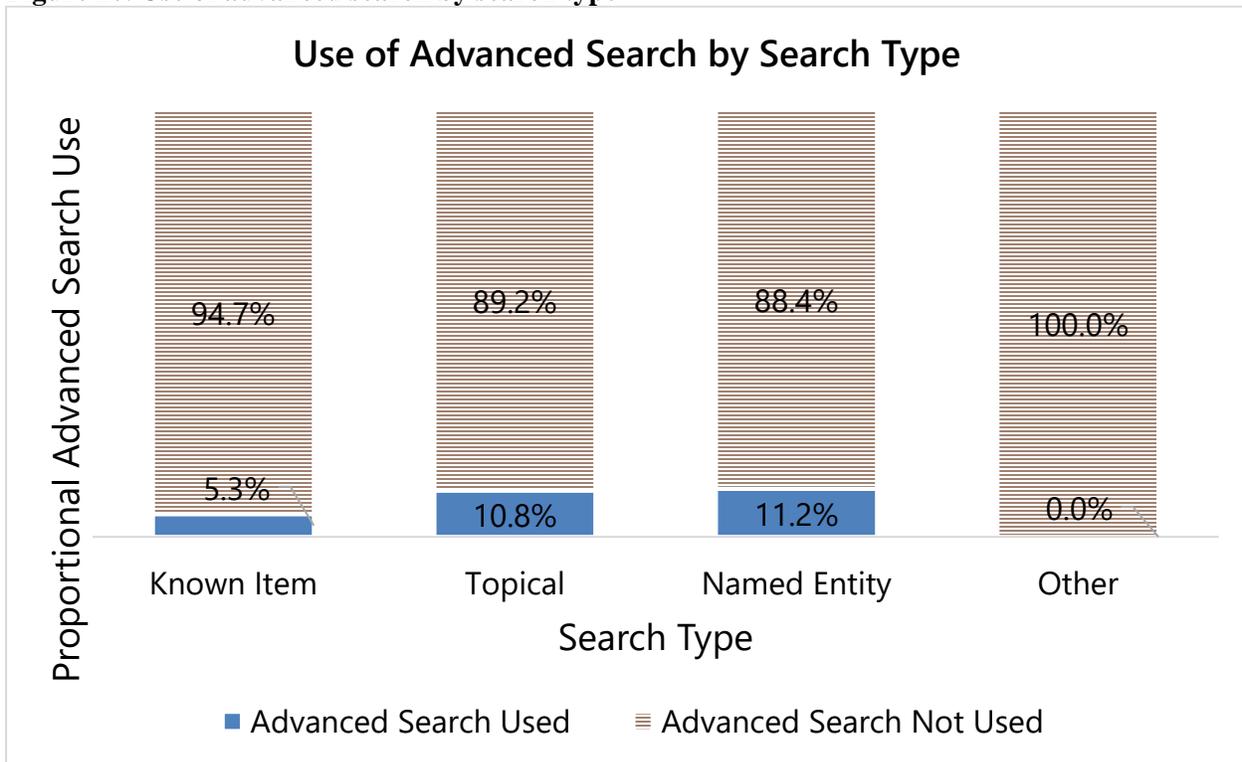
A final observation on search strategies used comes by way of comparison: users' choice of applying facets or conducting an advanced search when an equivalent option existed (Figure 11). As shown above, advanced search was overall much less popular than applying facets, which reflects most of the comparison points shown. Interestingly, the language facet was just as likely to be used in advanced searching as it was in faceting. The Format and Collection facet categories are prominently labeled on the search-results screen and appear just below the "Books & Media" and article-related facets ("Articles," "Full Text," "Peer Reviewed"), which justifies their relative popularity. In contrast, the Language facet category and the Publish Date facet (equivalent to advanced search's Year) fall "below the fold" of the average user's screen. The Language facet is further obscured by being collapsed by default; Publish Date comes after the collapsed Language, Course, Professor, and Tag categories, but its text-entry boxes are exposed by default. Whether because of this relative difficulty of use or because users searching by language were already more inclined toward the advanced-search paradigm, Language was proportionally less popular as a facet option.

Figure 11. Choice of use between advanced search and facets
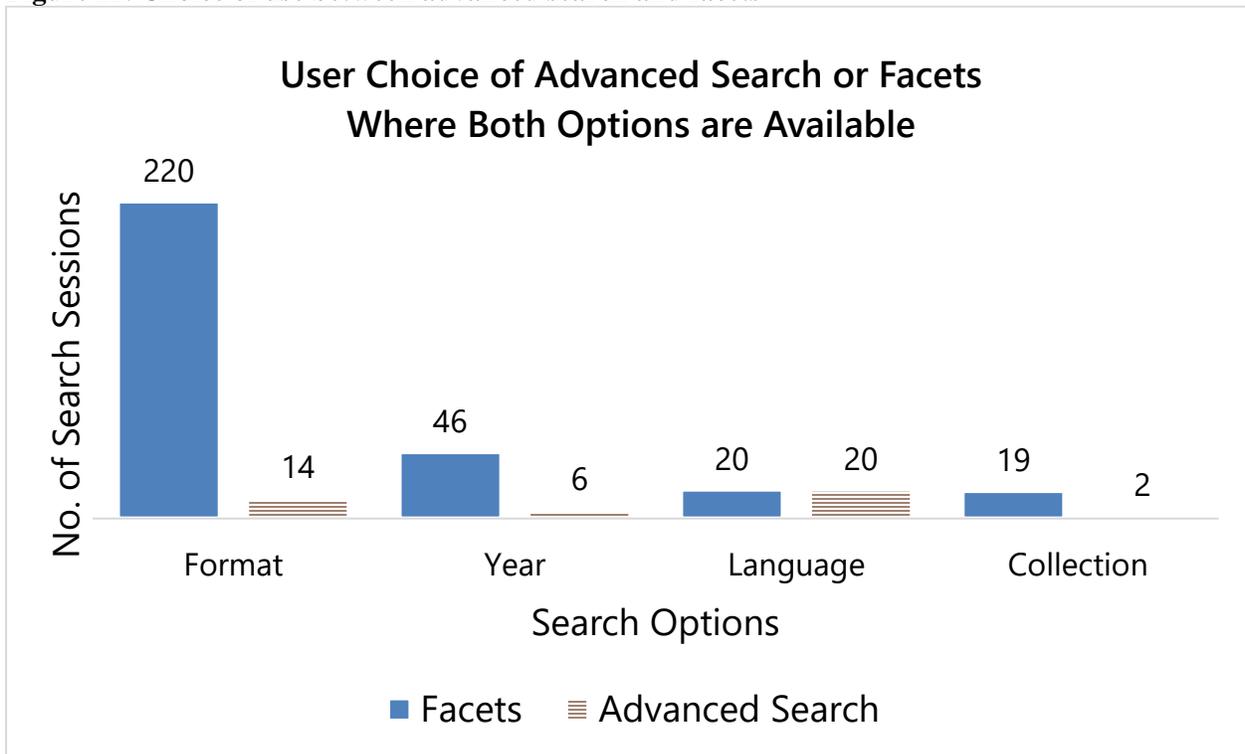
Search Effectiveness

To articulate the effectiveness for search term adjustment, the data set was coded and analyzed to determine if changes to search terms could have been handled by system features (such as advanced search, facets, or "Did you mean?") instead of changes to the search terms. In 75.31% of search term changes, no system feature could have handled the change that was made to the terms (Figure 12). In 12.36% of search terms changes, the alterations to the search terms were a result of using the system features, for example clicking on an authorized name heading to view a list of all publications by that creator. For 12.01% of search term adjustments, the system could have performed the same task as the user's changes to the search term. These could primarily have been handled by advanced search (representing 9.45% of the total search term changes), and less so by the "Did you mean?" (1.89% of changes) and facets (0.67%).

Figure 12. System alternatives to search term adjustments
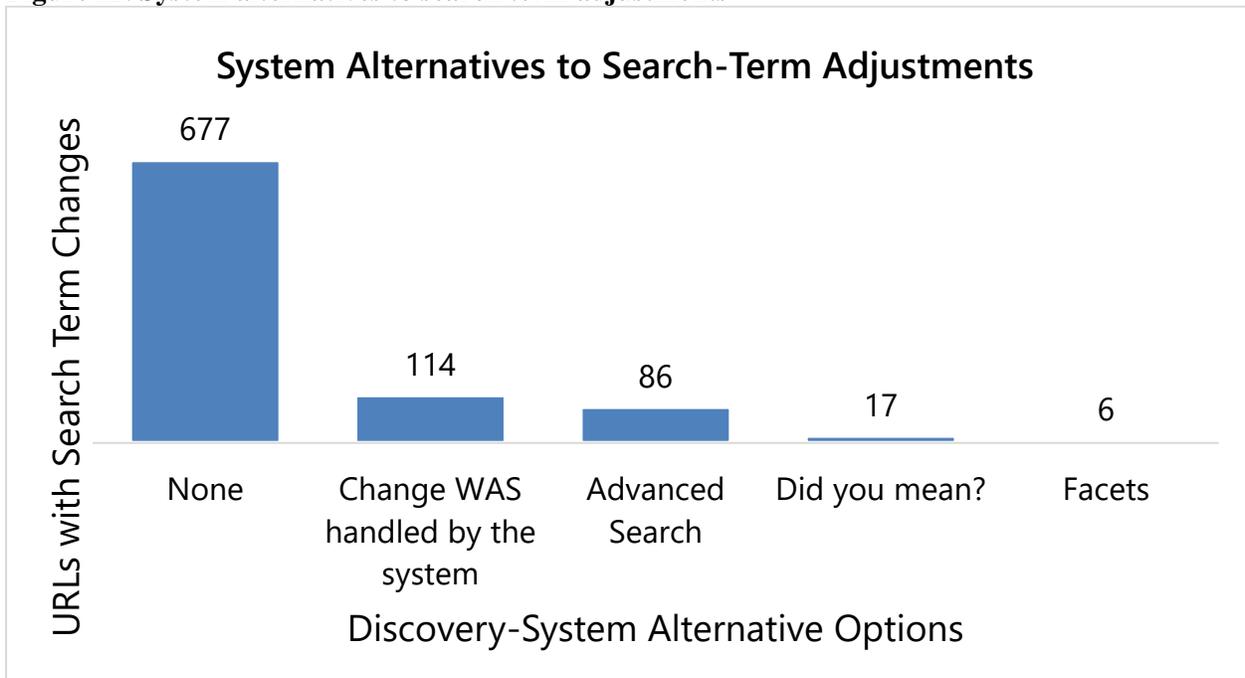
Examining the use of advanced search revealed a surprising trend: 43.7% of advanced searches were done incorrectly. Users appeared to anticipate a phrase search when using the advanced search feature and searched the keyword field solely or two or more keyword fields instead of applying the author, title, or subject indexes or other limiting options available on the page. The results of searching one or more keyword field without the use of quotations to initiate a phrase search brought the exact same results as conducting the same search from the main Encore search box. For instance, one user searched for "Music Therapy" and "sensory disorders" (sans the quotation marks) separated into two keyword fields with no other limiters (Figure 13). This brought up a page with 35,205 results. Running a search from the main Encore search box with all four terms together and no punctuation yielded the exact same set of results. However, putting quotations around each of the terms narrowed results down to 148. The researchers suspect that users are expecting phrase searches when using the advanced search page. This is an area where discovery systems could clarify how the advanced search works or change the parameters to be phrase searches.

Figure 13. Example of a dual keyword advanced search

Search effectiveness is a difficult metric to assess for topical searches using log analysis, particularly because the desired outcome is hard to determine and measure. Known items represent a concrete, measurable need, where the success or failure of the search can be more closely estimated. Therefore, the researchers decided to focus on known item searches in this log analysis and look towards more qualitative means for assessing topical searches in the future.

Users did not show a solid pattern of employing any of the three search strategies for known items. Facets were used in only 14% of search sessions for known items and changes to search terms were only used in 18% of search sessions for known items. The primary reason for using search term changes was to fix a typing error. Still, the rate at which either of these strategies was applied for known item searches indicates that the initial search failed to bring a recognizable copy of the item to the top of the search results.

The collection provided access to 77% of the known items searched for by users. Known items generally appeared within the first few results, with the number 1 position being the most common (Figure 14). Of all known item searches, under 5% of the time, the item was not included anywhere in the results (represented by the "N/A" in Figure 14). In 30 of 609 known items, the item appeared in the search results list more than once. This is due to multiple records for different editions, formats, or content providers.

**Figure 14. Position of known items in search results**
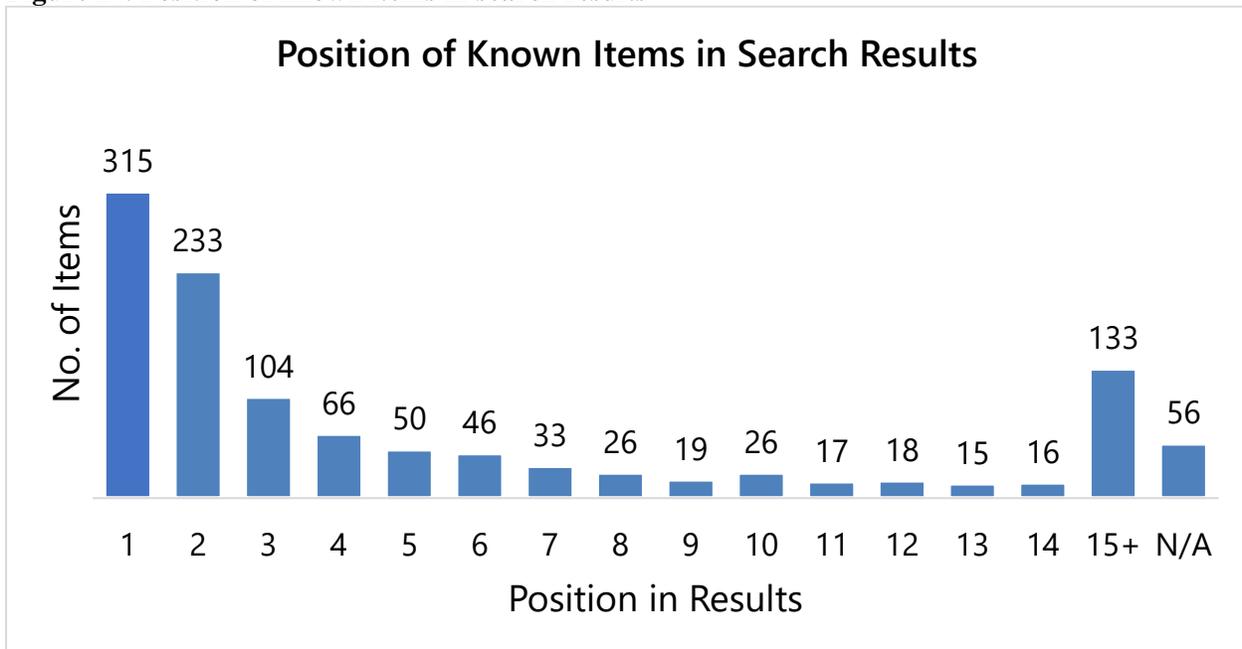
The researchers examined the availability of known items to determine if they should appear in search results, then analyzed how often they were displayed in users' search results, and how many steps it took to access the item. As reflected in the data in Table 1, Encore correctly listed access to an item for 80.59% of all known items. Of just the items that appeared in search results, 87.7% of items were available. The remaining 12.3% represented items which were listed but not accessible through Encore, such as indexed citations for publications which the library did not own a subscription. Of the known items that were available through the library, 87% were found in Encore and 13% were not. This discrepancy was often due to searches for unindexed sections of a known item (such as a book chapter), complex titles (such as a piece of sheet music), known items which are commonly cited in titles of other works (such as the Diagnostic and Statistical Manual of Mental Disorders or DSM–5), or user error.

Table 1. Discoverability vs. deliverability of known items

		Available	
		Yes	No
Appeared in Results	Yes	406	57
	No	61	84

For electronic resources, the number of steps it took to access an item was measured and recorded by the researchers. This metric does not reflect the number of steps it took the user to access the item, because this data was not available from the log analysis. It is an area that the researchers will examine in qualitative studies in the future. For the sake of understanding the users' environment, though, the researchers measured the shortest path from the search results page in Encore to the full text of the electronic resources. In 60.34% of electronically accessible known items, a single click could access the full text, and 98.56% of available electronic resources could be accessed in three or fewer clicks (Figure 15). No resource took over five steps to access.

Figure 15. Electronic access steps for known items
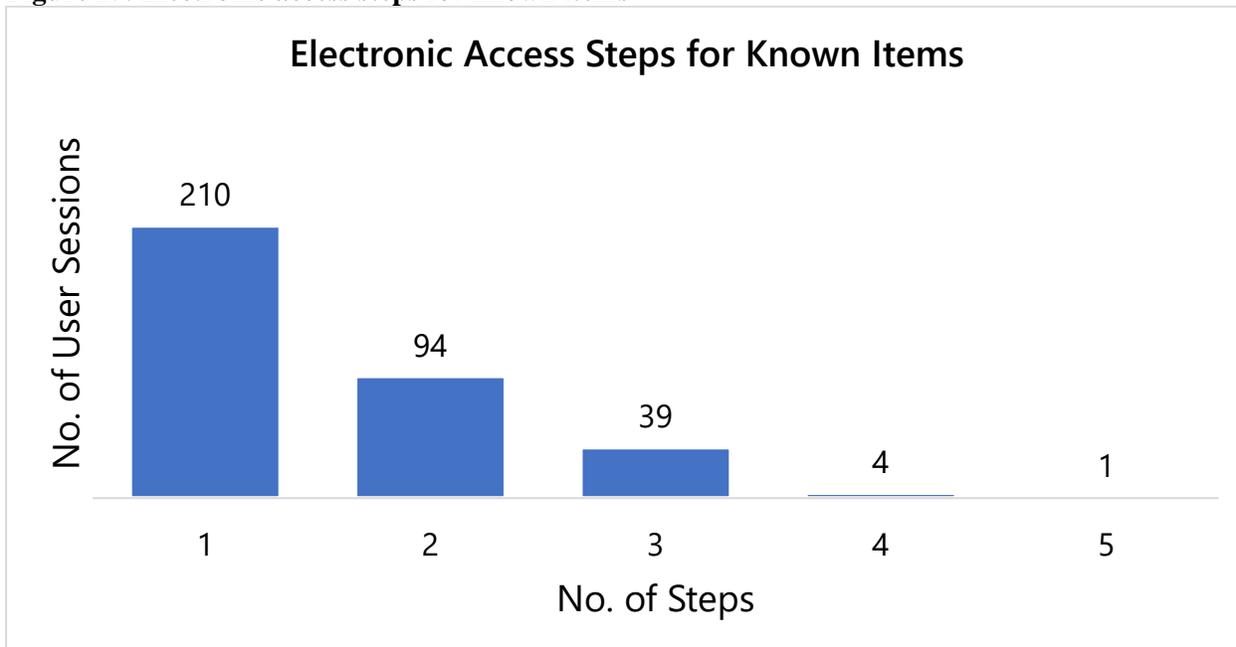
Determining the circulation of physical items was a difficult and imperfect metric to capture. Due the system storage of checkout dates, two sets of data had to be extracted to approximate checkouts for specific items: 1) items that were currently checked out and 2) items with a last checkout date within two weeks of the data collection dates. These two data sets were combined and deduped to triangulate an approximate list of bibliographic records along with their checkout dates. These were compared against the bibliographic record numbers of items that appeared in search results lists for known items. Of the available physical known items, only 14% of items provided a match, indicating that even though users searched for known items and Encore displayed them in a search results list, a large majority were not ultimately checked out by users. However, as the available data couldn't provide a simple list of checkout dates per bibliographic record number, it is not a complete data set and the researchers emphasize that impact from this finding should not be considered conclusive.

Discussion

Revisiting the Research Questions

The first question driving this research was *How did users conduct their searches?* The findings show that users' strategies change depending on the type of search (Table 2). The most popular type of search, known item, is the fastest and least complex, averaging just two pages visited and three minutes per session. In comparison, topical searches led users to visit twice as many Encore pages and spend twice as long, with similar increases in the use of search strategies. When a session includes both known-item and topical searching, many of these metrics double again. Although these findings are expected, they are encouraging: at least at a high level, the search logs show thoughtful and sophisticated engagement with the tool that the library promotes most heavily to its users.

Table 2. User engagement and strategies by search type

Measure of engagement	Known-item-only searches	Topical-only searches	Both
Average # pages visited	2 pages	4 pages	7 pages

Measure of engagement	Known-item-only searches	Topical-only searches	Both
Average search session time	3 minutes	7 minutes	14 minutes
Uses facets	10.6%	31.3%	25%
Uses advanced search	5.31%	10.8%	15%
Uses search term changes	22.08%	40.66%	82.5%

The second research question examined *How effective are users' strategies?* In other words, *Did Encore generally provide what users were looking for?* Most often, this was in fact the case. Users' most common strategy, modifying their search terms, did not contravene features that the tool made available. Encore's relevancy ranking interpreted users' searches correctly, bringing at least known items to the top. Users make short, targeted visits to Encore, apparently more interested in full-text availability than the full breadth and depth of the tool's indexing. Encore generally delivered the desired full text and did so with few intervening steps. To a lesser but still meaningful extent, the discovery of items in the libraries' physical collections remains an important consideration in choosing and implementing a discovery solution.

Unexpected Answers and Unanswered Questions

A serendipitous glitch during our data collection raises an additional point in favor of maintaining a discovery tool rather than only a traditional library catalog. Library staff updated an API password with EBSCO but did not pass the new credentials to Innovative. Because of this, data collected on one day pulled up only catalog results, excluding all article data normally provided by EBSCO. The error was caught and rectified by reprocessing those URLs after updating the password. By comparing the data sets, the researchers found that with only catalog results available, 53% of user searches would have returned zero results, and the average number of search results returned decreased by 35%. If the library were to contemplate moving to a catalog-only search, it would require a significant change in user behavior. Users' actual queries seem much better calibrated to a web-scale discovery tool than a traditional catalog.

As is often the case, this study raised a number of questions beyond those that it set out to answer. Answers to some of these might be found in the data already collected while others require different types of exploration. In some cases, users went along with the tool's defaults, but in others they moved past the first option given. To what extent might a tool's defaults influence users' search strategies? Or in contrast, what user behaviors would likely be consistent no matter the tool? The answers to these questions could save libraries considerable time shopping for tools or customizing a local implementation on the basis of options that are not likely to become important to users, despite librarians' best intentions.

The library would also benefit from a more nuanced understanding of successful and failed searches. Among those failures already identified, how can the library intervene to reduce specific types of failure? A qualitative approach could reveal whether users are satisfied with what they succeed in finding and how frustrated they are when they fail to find a known item or sufficient information on a topic. As a different approach to the question of perceived success, how would instructors' or librarians' judgment of the quality and relevance of retrieved sources compare to that of students?

Finally, just as user searching and discovery-tool performance are intricately intertwined, they are further complicated by questions of collection quality. What to offer more of, what to offer less of, and even how to gauge the value of an online download or physical checkout are ever-present concerns that cannot be separated from questions of discovery tools' support of both discovery and delivery. The researchers

Conclusion

This study's log-analysis method led to valuable insights into both how users are conducting their searches and how well Encore responds to those searches. Users in the data set most often conducted known item searches (54.8% of sessions), followed by topical (37.9%), named entity (18.4%), and other (.005%). Of the known items sought by users, 77% were contained in the library's local or licensed collections; books and articles were comparably prevalent among known-item searches (42.2% and 43.3%, respectively), with a breadth of other materials desired less often. Users kept search sessions brief (66% were under two minutes) and basic (54% included one or two page interactions). Most search sessions (51%) included no special strategies, with 35.6% including changes to search terms, 22.2% applying facets, and 8.4% using advanced search, most often with multiple strategies combined. A tool feature could only have accomplished modified search terms 12% of the time, suggesting that users applied the strategy effectively. Users neither used advanced search often nor used it well; 9% of sessions included an advanced search, and 43.7% of these reflected users' misunderstanding of its features. When conducting topical searches, users spend more time, visit more URLs, and employ each of the search strategies at roughly twice the rate of known-item searches.

Applicability to three domains of library practice motivated this research: configuring facet labels, link placement, and other details of the current Encore implementation; providing useful support to discovery users via instruction, reference, and online tutorials; and preparing for formal exploration of other discovery systems for a potential migration. All of these are part of the library's near-future plans, and the high-level takeaways from this study's findings will inform the library's approach in each domain. First, the choices made should account for users' most common needs, i.e., known items first, followed by topical searching. The library's tools and supporting materials should facilitate and accommodate those before other concerns. Second, users' most common strategies (i.e., adjusting search terms, then applying facets) should drive interventions the library may plan. Again, the researchers do not yet know to what degree users accommodate their behaviors to the tool in front of them, but it is safe to assume that today's users will generally expect similar features from the next iteration of a search tool. Along with data on these dual considerations, this research systematizes at least a partial framework for measuring the effectiveness of searches and tools. In essence, are the resources that users want 1) in the collection, 2) discoverable, and 3) accessible? Next steps might be summarized as planning interventions—in the current configuration, user support activities, and assessing potential tools—and then measuring the improvements achieved through those interventions.

—Copyright 2021 Robert Heaton and Liz Woolcott