The “Black Box”: How Students Use a Single Search Box to Search for Music Materials

Kirstin Dougan

ABSTRACT

Given the inherent challenges music materials present to systems and searchers (formats, title forms and languages, and the presence of additional metadata such as work numbers and keys), it is reasonable that those searching for music develop distinctive search habits compared to patrons in other subject areas. This study uses transaction log analysis of the music and performing arts module of a library’s federated discovery tool to determine how patrons search for music materials. It also makes a top-level comparison of searches done using other broadly defined subject disciplines’ modules in the same discovery tool. It seeks to determine, to the extent possible, whether users in each group have different search behaviors in this search environment. The study also looks more closely at searches in the music module to identify other search characteristics such as type of search conducted, use of advanced search techniques, and any other patterns of search behavior.

INTRODUCTION

Music materials have inherent qualities that present difficulties to the library systems that describe them and to the searchers who wish to find them. This can be exemplified in three main areas: formats, titles, and relationships. First, printed music comes in multiple formats such as full scores, vocal scores, study scores, and parts; and in multiple editions such as facsimiles, scholarly editions, performing editions (of various caliber); each format and edition serving a different purpose or need. Related to this, but less problematic, is the variety of sound recording formats available. Second, issues resulting from titling practices abound in music, ranging from frequent use of foreign terms, not just in descriptive titles (L’oiseau de feu = Zhar-ptitsa = The firebird = Feuervogel), but in generic titles as translated by various publishers from different countries (symphony=sinfonie). Additionally, musical works often have only generic genre titles enhanced by key and work number metadata, for example Symphony No. 1 in c minor. Third, music materials present a relationship issue best defined as “one-to-many.” Musical works often have multiple sections or songs in them (an aria in an opera or a movement in a symphony), and a CD or a score anthology may contain multiple pieces of music.

Given these three main challenges presented by music materials, it is possible that those searching for music develop distinctive search habits compared to patrons in other subject areas. This study uses transaction log analysis of the music and performing arts module of a library’s federated discovery tool to determine how patrons search for music materials. It also makes a top-level comparison of searches done using other broadly defined subject disciplines’ modules in the same discovery tool. It seeks to determine, to the extent possible, whether users in each group have different search behaviors in this search environment. The study also looks more closely at
searches in the music module to identify other search characteristics such as type of search conducted, use of advanced search techniques, and any other patterns of search behavior.

BACKGROUND

Since Fall 2007 the University of Illinois Library has had Easy Search (ES), a locally developed search tool designed to aid users in finding results from multiple catalog, A&I, and reference targets quickly and simultaneously. There is a “main” ES on the library’s main gateway page that searches a variety of cross-disciplinary tools (see figure 1).

**Figure 1.** Gateway Easy Search.

On the gateway, users have the option of selecting one of the format tabs to narrow their search to books, articles, journals, or media. When the data for this study was gathered, the journals tab was not present.

Starting in 2010 many of the subject and branch libraries in the University Library created their own ES modules with target resources specific to the disciplinary areas they serve. Search boxes for these ES subject modules are often displayed right on the branch library’s home page, but users can also select these subject module options from the dropdown in the main ES (see figure 2).
The MPAL ES interface as it appears on the MPAL homepage can be seen in figure 3—it was created in 2011.

**Figure 2.** Gateway dropdown subject choices.

**Figure 3.** MPAL Easy Search interface.
ES is a federated search tool and does not have a central index like most current discovery layer tools. Rather, it utilizes broadcast search technologies to target different tools and search them directly. While the Gateway ES now uses a “bento box” layout to display selected citations from each target, in the first iterations of the tool and still today in the subject modules, users are simply presented with a list of hit counts in each of the target tools (see figures 4 and 5).

**Figure 4.** MPAL Easy Search display screen part 1.

**Figure 5.** MPAL Easy Search results screen part 2.
Not shown in the screen captures are the results from various newspaper indexes and reference sources such as *Oxford Music Online* and the *International Encyclopedia of Dance*.

**LITERATURE REVIEW**

Many studies have examined patron search behavior using transaction log analysis and other methods over the past few decades. Since the appearance of Google in 1998, and its vast impact on individuals’ expectations and search behavior, recent studies have looked at user search behavior in tools that initially present a single search box. Additional studies have looked at discipline-specific searching behaviors.

**General Search Studies and Single Search Boxes**

Many advantages and disadvantages have been ascribed to tools with single search boxes (whether federated search tools or discovery layers), namely ease and convenience on the one hand, and the lack of precision possible in searching and overwhelming number of results on the other. Two companion articles, Boyd et al. and Dyas-Correia, written ten years apart, attempted to visit and revisit these issues.¹ Results and patron satisfaction can vary based on size of library and number of resources accessed by these tools. These types of tools will never be able to search and display everything and that problem is magnified by the number of resources a library has.

Holman, Porter, and Zimerman discovered in independent studies that undergraduates do not search very efficiently or effectively and find library tools difficult to use.² Avery and Tracy also found this true for the ES tool under discussion in this study:

> The generation of keywords by many students indicates they often struggled to identify alternative terminology that may have resulted in a more successful search . . . . Many students exhibited persistence in their searching, but the selection of search terms, sometimes compounded by spelling problems or problems in search string structure, likely did not yield the most relevant results.³

Asher, Duke, and Wilson state in their study comparing student search strategies and success across a variety of library search tools and Google that there were “strong patterns in the way students approached searches no matter which tool they used. . . . Students treated almost every search box like a Google search box, using simple keyword searches in 81.5 percent (679/829) of the searches observed.”⁴ Dempsey and Valenti note students’ infrequent use of limiters such as “peer-review” and “date” in EDS, the high non-use and misuse rates of quotation marks, relatively low instances of repeated uncorrected spelling errors, and variance patterns in keyword usage.⁵

Students like federated search tools and discovery layers because of their convenience and ease, as found in studies by Armstrong, Belliston, and Williams et al.⁶ This is reiterated in Asher et al., “Despite the fact that they did not necessarily perform better on the research tasks in this study, students did prefer Google and Google-like discovery tools because they felt that they could get to full-text resources more quickly.”⁷ This one-box approach could hinder students, as described by Swanson and Green:

> The search box became an obstacle in other questions where it should not have been used. In some cases, the search box was viewed as an all-encompassing search of the entire site.
Several students searched for administrative information, research guides, and podcasts in this box.\textsuperscript{8}

Lown et al. also found that users hope to access a vast range of information via a single search box. “One lesson is that library search is about more than articles and the catalog. About 23 percent of use of QuickSearch took place outside either the catalog or articles modules, indicating that NCSU Library users attempt to access a wide range of information from the single search box.”\textsuperscript{9}

\textbf{Search and Library Use in Different Disciplines}

In their study comparing a discovery layer and subject-specific tools, Dahlen and Hanson found “subject-specific indexing and abstracting databases still play an important role for libraries that have adopted discovery layers. Discovery layers and subject-specific indexing and abstracting databases have different strengths and can complement each other within the suite of library resources.”\textsuperscript{10} They also observed things iterated by previous authors, chiefly that “not all students prefer discovery tools” and “the tools that students prefer may not be those that give them the best results.”\textsuperscript{11} In addition, they found that default configuration matters in terms of students’ success in and preference for a given tool. Fu and Thomes found that creating smaller discipline-specific subsets in discovery tools was beneficial to searchers by reducing results and increasing the results’ relevance.\textsuperscript{12}

Few studies investigate how music students search for music materials. Dougan found in her observational study of music students’ search behaviors that they have difficulty forming good searches; misuse quotation marks and other search elements; and at times struggle with finding music materials.\textsuperscript{13} Mayer noted upper-class music students’ frustration with using library tools to find specific works of music, going so far as to state, “the music students agreed that both the discovery layer and the catalog are not effective for music-related searching, for any format.”\textsuperscript{14} Clark and Yeager found that students had an easier time searching for media items than music scores, and frequently struggled with search strategy revisions.\textsuperscript{15}

There is more research on the larger information needs of disciplines and creating models for research behavior, and not necessarily specific search processes or constructions.\textsuperscript{16} Whitmire, in her 2002 pre-Google article, found that students majoring in the social sciences were engaged in information-seeking behaviors at a higher rate than students majoring in engineering.\textsuperscript{17} Chrzastowski and Joseph surveyed graduate students at the University of Illinois at Urbana–Champaign and found that those in the life sciences, physical sciences, and engineering visited the libraries less often than students in other academic disciplines.\textsuperscript{18} Students in the arts and humanities used the library more often than students in other disciplines. Collins and Stone report that in prior studies of users across different disciplines, arts and humanities users do not account for the biggest users of library materials, their survey found the opposite to be true.\textsuperscript{19} When looking at the various student populations in their study, musicians had the highest library usage in terms of items borrowed and almost the highest number of library visits. Music users in the study also showed high numbers of hours logged into the library e-resources and highest number of e-resources accessed compared to others in their discipline group (but not as much as other disciplines). However, they show a low number of PDF downloads and low number of e-resources accessed frequently.
METHODOLOGY

This study conducted quantitative analysis of Easy Search (ES) data as a whole and from a selection of the subject modules, including the Music and Performing Arts Library (MPAL) module, using data from the period June 20, 2014 through June 16, 2015. Additional quantitative and qualitative analysis was conducted only on the MPAL ES transaction log data.

Data from the following subject modules were included in comparative analyses:

- Funk Agricultural, Consumer and Environmental Sciences Library (ACES) ([http://www.library.illinois.edu/funkaces/](http://www.library.illinois.edu/funkaces/))
- Grainger Engineering Library ([http://search.grainger.illinois.edu/top/](http://search.grainger.illinois.edu/top/))
- History, Philosophy, and Newspaper Library (HPNL) ([http://www.library.illinois.edu/hpnl/](http://www.library.illinois.edu/hpnl/))
- Music and Performing Arts Library (MPAL) ([http://www.library.illinois.edu/mpal/](http://www.library.illinois.edu/mpal/))
- Social Science, Health, and Education Library (SSHEL) ([http://www.library.illinois.edu/sshel/](http://www.library.illinois.edu/sshel/))
- Undergraduate Library (UGL) ([http://www.library.illinois.edu/ugl/](http://www.library.illinois.edu/ugl/))

Each of these libraries has a search box for ES on its home page that is customized to the search targets identified as best for those subject areas by the subject librarians in that library.

Transaction log data on searches done in ES is continuously compiled in a SQL database and queries were written to determine certain quantitative measures. Searches done in these various subject modules were isolated by a variable in the SQL data that indicates whether the search was done in the main Gateway ES, in the main Gateway ES but using one of the subject dropdown choices, or from the subject ES box directly from that library's homepage. Searches in the six subject modules listed above and in the main ES were assessed for the average number of searches per session and the average words per search.

Further analysis of searches done in the MPAL ES module used 25,503 sessions conducted on MPAL public computers from March 21, 2014 to June 21, 2015, which is a slightly longer time-span than used for the comparative analysis between subject ES modules described above. To make this more manageable, only every tenth session was considered, meaning 2,550 sessions were analyzed out of the full set of MPAL data. Searches were sorted by session ID number, which is assigned to each session when a new session is begun. This method kept all strings from one session together, whereas simply sorting by date and string ID did not, since multiple sessions can occur simultaneously. A session is a series of user actions (searches and click-throughs) from the same workstation in which there is less than a twenty-minute pause between actions. If there are user actions from the same workstation after a twenty-minute pause, a new session is established, therefore, there is the possibility that some of the sequential sessions were from the same user, but there is no easy way to determine that.

The MPAL data set was assessed using the following quantitative measures:

1) Average number of searches per session and whether session contained
   a) A single search
   b) Multiple searches for the same thing (either repeated exactly or varied)
c) Multiple strings searching for multiple things

2) Average number of search terms per search
3) Type of search by index (Title/Author/Keyword) or other advanced search
4) Use of Boolean, quotation marks, parentheses, etc.
5) Use of work or opus numbers or key indications
6) Search indicating format (score, CD, etc.)

**FINDINGS**

Comparing the data for searches done in the main ES to some of the subject modules (see table 1) shows that the UGL ES and the HPNL ES have the fewest average searches per session, and the MPAL ES has the third highest average number of searches per session. The sciences tend to have higher average words per search string values, while MPAL has the second lowest average number of words per search. This is not surprising given that the sciences tend to use a lot of journal literature and it is common for researchers to copy and paste such citations into ES. Whereas in music, as we will see later, keyword searches tend to focus on combinations of the composer’s name and words from the work title, occasionally with other terms added.

<table>
<thead>
<tr>
<th>Source</th>
<th>Sessions</th>
<th>Searches</th>
<th>Average searches per session</th>
<th>Average words per search</th>
</tr>
</thead>
<tbody>
<tr>
<td>All ES searches</td>
<td>599,482</td>
<td>1,340,159</td>
<td>2.121</td>
<td>5.082</td>
</tr>
<tr>
<td><strong>Gateway only</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gateway everything tab</td>
<td>382,040</td>
<td>757,862</td>
<td>1.983</td>
<td>5.255</td>
</tr>
<tr>
<td>Gateway books tab</td>
<td>71,007</td>
<td>136,724</td>
<td>1.925</td>
<td>4.048</td>
</tr>
<tr>
<td>Gateway articles tab</td>
<td>57,169</td>
<td>107,893</td>
<td>1.887</td>
<td>6.35</td>
</tr>
<tr>
<td>Gateway total</td>
<td>1,002,479</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>All Subject Modules</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Departmental searches (incl. those from Gateway dropdown)</td>
<td>75,035</td>
<td>214,364</td>
<td>1.928</td>
<td></td>
</tr>
<tr>
<td>Searches done directly from subject library pages</td>
<td>144,283</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Select Subject Modules</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agricultural, Consumer and Environmental Sciences Library</td>
<td>2,732</td>
<td>5,221</td>
<td>1.911</td>
<td>4.07</td>
</tr>
<tr>
<td>Engineering Library</td>
<td>32,018</td>
<td>68,146</td>
<td>2.128</td>
<td>5.092</td>
</tr>
<tr>
<td>History, Philosophy, and Newspaper Library</td>
<td>1,264</td>
<td>1,985</td>
<td>1.57</td>
<td>3.09</td>
</tr>
<tr>
<td><strong>Music and Performing Arts Library (MPAL)</strong></td>
<td>21,047</td>
<td>41,590</td>
<td>1.976</td>
<td>3.375</td>
</tr>
<tr>
<td>Social Science, Health, and Education Library</td>
<td>9,458</td>
<td>19,760</td>
<td>2.089</td>
<td>4.906</td>
</tr>
<tr>
<td>Undergraduate Library</td>
<td>26,988</td>
<td>44,588</td>
<td>1.65</td>
<td>3.909</td>
</tr>
</tbody>
</table>

**Table 1.** Comparative search data from June 20, 2014 to June 16, 2015 (unless otherwise noted).
Average Number (and Range) of Searches Per Session

In looking at the searches done directly from the MPAL homepage and from the Gateway dropdown from March 21, 2014 through June 21, 2015, there were 25,503 sessions conducted in the MPAL ES that contained a total of 49,702 searches, resulting in an average of 1.949 searches per session. Of the 2,550 MPAL search sessions in the study sample, the majority (63.2 percent) consisted of one search. This means the patron conducted one search and then left ES, presumably clicking into the library catalog or another tool that is a target in ES to complete their research. Sessions consisting of two to four searches account for 31 percent of sessions, while sessions involving five to nine searches only account for 5 percent of total sessions, and only 32 sessions, or fewer than 1 percent, consist of ten or more searches (see table 2).

<table>
<thead>
<tr>
<th>Searches per session</th>
<th>Number of sessions</th>
<th>Searches per session</th>
<th>Number of sessions</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1604</td>
<td>12</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>476</td>
<td>13</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>191</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>116</td>
<td>16</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>51</td>
<td>17</td>
<td>2</td>
</tr>
<tr>
<td>6</td>
<td>29</td>
<td>18</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>19</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>12</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>9</td>
<td>15</td>
<td>23</td>
<td>1</td>
</tr>
<tr>
<td>10</td>
<td>6</td>
<td>30</td>
<td>2</td>
</tr>
<tr>
<td>11</td>
<td>6</td>
<td>Total searches</td>
<td>2,550</td>
</tr>
</tbody>
</table>

Table 2. Searches per session.

Sessions with multiple searches (n= 946) were evaluated to see whether patrons were searching multiple times for the same thing (either with the same term[s] or with different terms), or whether they were searching for different things. Five sessions that were clearly not music-related were removed from the sample. Each session was categorized as “same/exact,” “same/different,” or “different.” At times, sessions might include several searches for the same thing using altered strings, in addition to searches for other things. Those sessions were coded as “different.” For example:

- crumb zodiac
- crumb
- georgy crumb
- georgy cromb
- korean music

There were 478 multi-search sessions (50.6 percent) in which patrons searched for different things within their session, 391 sessions (41.3 percent) in which patrons looked for the same thing with differing search strings, and 71 (7.5 percent) in which patrons reiterated the exact same search in each attempt. In the 71 sessions in which patrons used the same exact search
multiple times, they averaged 2.25 searches. Those sessions tagged as “same/exacts” provide an opportunity to try to determine why patrons repeat the same search. Common themes include: using too broad a search, searching in wrong place (non-performing-arts-related search), or repeatedly typing in the wrong info (e.g., typos or other errors) and not realizing their mistake.

In the 391 sessions in which patrons spent their session searching for the same thing with different search strings, they did so with an average of 2.96 searches. Often the variation in the search string was a change in spelling or a minor change in the terms, but sometimes it involved the addition or subtraction of terms, such as starting with morley fitzwilliam virginalists and going to morley fitzwilliam. In another example, we see how music metadata can prove challenging for searchers to format, such as when a patron started with schumann op.68 (without the necessary space between op. and 68), then progressed to album for the young, and finally schumann album for the young.

In the 478 sessions in which patrons searched for completely different things within their session, they did so with an average of 4.08 searches per session. In many cases, although the searches were for different items, they were related in some way, either by genre, instrument, or some other element, such as in this example:

- microjazz
- color me jazz
- jamey aebersold play-along
- vandall jazz
- jazz piano pieces

But sometimes the searches were for very different things:

- debussy voiles
- composition as problem mart humal
- composition as problem
- debussy ursatz

**Average Number (and Range) of Terms Per Search**

In looking at the approximately 4,900 searches included in the sample of 2,550 MPAL sessions, without removing the small percentage of duplicate searches, two-term searches are the most common, followed by three-term searches— together accounting for more than half of the searches (55.3 percent). One- and four-term searches are the next most common, together accounting for 25.5 percent of searches (see table 3). In 2012, regular ES single-term searches were at almost 60 percent.23
<table>
<thead>
<tr>
<th>Number of terms in search string</th>
<th>Instances</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>605</td>
<td>12.4</td>
</tr>
<tr>
<td>2</td>
<td>1,559</td>
<td>31.8</td>
</tr>
<tr>
<td>3</td>
<td>1,149</td>
<td>23.5</td>
</tr>
<tr>
<td>4</td>
<td>642</td>
<td>13.1</td>
</tr>
<tr>
<td>5</td>
<td>400</td>
<td>8.2</td>
</tr>
<tr>
<td>6</td>
<td>196</td>
<td>4.0</td>
</tr>
<tr>
<td>7</td>
<td>100</td>
<td>2.0</td>
</tr>
<tr>
<td>8-15</td>
<td>216</td>
<td>4.5</td>
</tr>
<tr>
<td>16-57</td>
<td>29</td>
<td>.06</td>
</tr>
</tbody>
</table>

Table 3. Words per search string.

Longer search strings (8-15 terms) ranged from 74 to ten examples each, respectively, while searches with 16 to 20 terms ranged from 8 to 2 examples each, respectively. The following term counts each had only one example in the logs (25, 26, 31, 32, 36, 57).

**Single-Term Searches**

Types of single-term searches can be broken down into several categories (see table 4). Over half (58.4 percent) were searches for personal names or part/all of a work title. Some names and work titles are in fact so unique that a one-word search might in fact be successful (e.g., Beyoncé, Schwanengesang, Newsies, or Landowska). Over a fifth (22.2 percent) were classified as “Other or undetermined,” including publisher names, cities, or subject terms.

<table>
<thead>
<tr>
<th>Type of one-word search</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal name</td>
<td>260</td>
</tr>
<tr>
<td>Title</td>
<td>93</td>
</tr>
<tr>
<td>Instrument/genre</td>
<td>51</td>
</tr>
<tr>
<td>Tool/location/format</td>
<td>51</td>
</tr>
<tr>
<td>Call number/barcode/label number</td>
<td>15</td>
</tr>
<tr>
<td>Other/undetermined</td>
<td>135</td>
</tr>
</tbody>
</table>

Table 4. One-term search types.

In the Tool/Location/Format category patrons searched for things such as: albums, images, dissertation, RILM, WorldCat, JSTOR, and IMSLP. While RILM (Abstracts of Music Literature) and WorldCat can be found by a search in this tool because they will match on journal or database titles to which we subscribe, a search for IMSLP [International Music Score Library Project] only brings back mentions of IMSLP in RILM, etc. MPAL links to IMSLP on its webpage, but neither IMSLP nor the library’s website are targets in ES. When patrons only searched for a format, as in a session where a patron first searched for performances, then albums, and then audio cd [sic], it is difficult to know whether the patron expected to be led to a tool that only searched or listed recordings, whether they wanted a list of all of our recordings, or if some other logic was occurring. Searchers also used this technique in multi-word searches, such as in the example george gershwin articles.
Single-term searches in the “Other/undetermined” category were a mix of subject terms like solfege, tuning, and spectralism. The patron could be trying to find materials related to these topics, examples of them (in the case of solfege), or definitions for them. They also included publisher or label terms such as rubank and puntamayo [sic], and even, on more than one occasion, URLs and DOIs.

Two-Term Searches and Names
The largest segment of the MPAL data (31.8 percent) is comprised of two-term searches. The examples show that often a musical work can be easily sought based on the composer’s name and a word from the title, especially in cases where it is a common title but adding the composer’s name makes it unique (e.g., Ligeti Requiem). Sometimes the patron only knows the work’s characteristics and not its proper name (e.g., Lakme duet). Patrons do attempt to search for topical material using only two words, and that is not likely enough for a good topic search in most cases, such as in the example mahler dying. Sometimes phonetic spellings are employed such as woozy wick followed by woyzeck (which is both a play and a film with this spelling but could also potentially be a misspelling of Berg’s opera Wozzeck). Another example is image cartier followed by images quartier.

Personal names are frequently seen in two-term search strings. Occasional use of foreign versions of names is observed, e.g., georgy crumb. It is difficult to know if these are typos or an artifact of our high international student population. As with any search that contains only a name, it is impossible to know whether the searcher was looking for materials by that individual or information about them. Additionally, when current faculty names are searched, it is difficult to know whether patrons are looking for contact information for them, or scores or recordings by them. Also observed in name searches is the phenomena of patrons repeating their search with a change in order of names, such as bryan gilliam and then gilliam bryan. This occurs with other two-word searches as well, such as a change from introitus gubaidulina to gubaidulina introitus. Switching the order of the words in a search no longer makes a difference in most search tools (although in some catalogs, of course, it was once required to formulate an author search as Last Name, First Name). There is still the occasional use of comma in LN, FN searches here.

Echoing the results of an earlier study that asked students what data points they used in searching, only occasionally did searches in this data set incorporate specific performers combined with a particular piece or composer: franck mutter, or for a particular edition: idomeneo barenreiter. Sometimes names/titles were combined with format, such as a session in which a patron searched for Hedwig images and then Hedwig photo. Here it is hard to tell if they are looking for pictures of a fictional owl or images from productions of Hedwig and the Angry Inch, or something else. Names are also frequently combined with work numbers instead of title words, such as mozart k.395 and moscheles op.73. Search strings in the “Other/undertermined” category sometimes included what appears to be an author/date search, perhaps for an article, such as mccord 2006.

Long Search Strings
On the other end of the spectrum, the vast majority of the ten-plus word string searches are for performing arts items, but some were in other subject areas. These long searches are often citations that have been copied and pasted, which can be discerned from the use of punctuation and capitalization, like “Welded in a single mass”: memory and community in London’s concert halls.
It is very common in general Gateway ES searches to see an entire citation pasted in, but less common in the MPAL module. Searches such as this are often truncated through iteration to make the search more generic (see table 5).

Given Easy Search’s DOI search recognition function, the longest version of this search would have worked had the DOI been correct, but the correct DOI number lacks the “.2” at the end (see table 6, query 1). The middle three searches (#s 2-4) failed because none of the A&I services that include this citation use hess, j. for the author’s name, but instead use her full first name (Juliet). Other examples showed that even when patrons use the exact citation, their search might not be successful if the citation formatting did not match that of the database(s) in which the article was indexed.

<table>
<thead>
<tr>
<th>Query</th>
<th>Query string</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>hess, j. radical musicking: towards a pedagogy of social change.</td>
</tr>
<tr>
<td>5</td>
<td>radical musicking:</td>
</tr>
</tbody>
</table>

Table 5. Search truncation.

In some instances, searches were long because the patron included additional information such as in this example: bernstein, leonard. arranger: jack mason. title: west side story-selections (for symphonic “full” orchestra piano-conductor score). edition/publisher: hal leonard corporation. It is hard to tell if this was a copy and paste from another source such as a publisher catalog, or if the patron was trying to be very precise. In any case, this search was not successful, but would have been had the searcher omitted extraneous information such as the terms “arranger” and “edition/publisher.”

**Type of Index Search—Title/Author/Keyword and Adding Subsets or Tools**

Easy Search does have an advanced search function with indexes for title and author, although it is rarely used by patrons. Including repeated searches, searches done selecting the “Title” index only numbered 207, or fewer than 10 percent of the sample. Searches done selecting “Author” were even scarcer, at 141 (5.5 percent). The remaining ~2,300 searches in the sample were conducted using the default keyword search. Occasionally there was a misuse of index searching, such as:

- ti: js bach english suite
- ti: scarlatti sonatas
- ti: haydn cello concerto D

In these examples, composer name is included in a title index search. It is unclear whether searchers do not realize that they have selected something other than a keyword search, or whether people inherently think of the composer’s name as part of the title. Later in this paper the phenomenon of searches using possessive name forms is discussed, which may be associated.
Patrons have the option to start from the Main Library gateway and perform a search in ES, and in the advanced search screen can choose other subject modules such as Arts and Humanities, Life Sciences, and so forth, and/or types of tools to cross-search (see figure 6). Patrons chose the music and performing arts tool subset in 161 sessions.

![Easy Search Advanced Search Screen](image)

**Figure 6.** Easy Search Advanced Search Screen.

The vast majority of the time (4,557 searches or 93 percent), patrons chose to start from the MPAL ES on the MPAL homepage and do a basic search there, but 179 times patrons started from the MPAL ES and chose other subsets through the advanced search. Given our large music education program, logically, some patrons made tool choices that included the music subset and the education and/or social science subsets. But sometimes patrons chose every or almost every option available across multiple unrelated subject areas, which likely made for a very unwieldy result set.
Use of Boolean Operators, Quotation Marks, Parentheses, Truncation, Etc.

As in most search tools, there are several ways in ES to conduct more sophisticated searches. However, patrons do not employ these techniques often, in part because they don’t always have to. In most older catalogs (including our classic Voyager OPAC), searchers had to use Boolean terms in capital letters, whereas in VuFind and WorldCat Boolean AND is now implied between terms. In the 159 examples of Boolean logic in the searches, AND is most common term used. Interestingly, some researchers used plus signs instead of AND (as they might in Google), not just between individual words, but in between multi-word segments of the string (without employing quotation marks). However, the + sign, like AND, is ignored/implied by ES.

berg + warm die lufte

progressive studies for trumpet

progressive studies for trumpet + john miller

progressive studies for trumpet (john miller)

new orleans + bossa nova

johnny alf + brazil

dick farney + brazil

dick farney + booker pittman

In some cases, the use of Boolean did not seem intentional, that is, the term “and” appears as part of a common phrase (especially for instrument combinations), such as in webern violin and piano.

Only a handful of the Boolean searches included examples of OR and NOT, which seemed to stem from a class assignment designed by a professor, as the search strings are all very similar. One set is below:

Machaut NOT Mass
Machaut OR Mass
Machaut AND Mass
Machaut Mass Notre Dame
Machaut mass
Machaut AND mass
Commas were sometimes seen to stand in for Boolean operators in a sense, or at least to separate search concepts, like the plus signs above, but were not counted in the total uses of Boolean terms cited above. They are ignored by ES.

rachmaninoff, moment music
planet, holst
City noir, john adams
piazzolla, flute and marimba
Mussorgsky, Pictures at an Exhibition, Manfred Schandert

Searchers used quotation marks on occasion (n=162) to keep phrases together, and parentheses were also used in this manner eight times (although they are ignored by ES), such as in these examples:

Preludes and fugues (Well-tempered clavier) cohen
Chaconne (from Partita in D minor, BWV 1004)

In some cases, searchers did not seem to grasp the function of quotation marks, as in this example: “Snowforms” Raymond Murrey Schafer, which was also observed by Avery and Tracy.28

Truncation symbols can be another powerful tool in a searcher’s arsenal, but examples of their use in the transaction logs show that most searchers who attempt to use them do not understand them, such as in the examples Doctor Atomic?, Boethius music,* and:

orchestra* history
history of the orchestra
orchest* history
orchestr* history
orchestra history
orchestral history

In fact, the current library catalog assists users by automatically applying truncation logic so that “symphony” returns results for “symphonies” and vice versa. It is doubtful that this is generally known among users and likely functions in a manner transparent to most of them.

Work Numbers and Key Indications
Searching by music metadata elements such as work or opus numbers and key designations has always proved challenging in online search environments given that numbers and single letters can appear in other parts of the catalog record with different meanings (e.g., 1 part instead of symphony no. 1). Added to this is the difficulty of describing items that contain multiple works—the item’s title might be “Mozart’s Complete Symphonies” or “Beethoven Symphonies 1-6” without
complete work details provided. Nevertheless, 134 searches had some form of work number included, and 36 searches included a key indication. *Fantasie in f# minor presto* georg philipp Telemann and *concerto en ut mineur j.c. bach* are further examples of why a work’s key is hard to search by, one because of the use of the French solfege syllable “ut” and one because it includes a sharp symbol (#). The difficulties this can cause often led searchers to try various permutations of their search.

```plaintext
mozart concerto g major sam franko;
mozart concerto k 283 sam franko;
scores;
mozart violin concerto g major;
mozart violin concerto g major sam franko;
mozart violin concerto;

sonata g major flute cpe bach
sonata g major flute bach
hamburger sonata flute
cpe bach hamburger sonata
hamburger sonata
```

It is counterintuitive to searchers that including specific details in their search string might not help, but that is in fact the case in many online catalogs. Searchers often run into the question of how or if to include the work indicator (op., K., BWV, etc.), which can lead to a “misuse” of this extra data such as in *mozart k501* and *Mahler Symphony No.9* (no spaces).

Another observation includes the use of what the author calls musicians’ shorthand. That is, those familiar with classical repertoire will know that examples such as *Sibelius 1* and *Mahler 5* are searches for symphonies even though they do not say so, but it will be harder, if not impossible, for the catalog to interpret that, leaving the searcher to sort through many extra results. In addition is the long-standing issue of whether to enter the number as “1”, “1st”, or “first” and whether the system can interpret these against the form of the number present in the catalog record.

**Search by Format or Edition Type**

In forty-seven examples searchers used format terms in their searches, including *score, vocal score, full score, DVD, performance recordings, albums,* and *audio CD* as well as the following:

```plaintext
prokofiev romeo and juliet orchestra parts
orchestra excerpts prokofiev romeo and juliet viola
```
In fifteen examples searchers searched for edition types including *urtext, facsimile, critical edition, and complete works*. In the latter case they occasionally used the word “complete” and the composer's name, such as *complete Schumann* or *complete Webern*. Unfortunately, this approach will often not be successful, because even though the term “complete works” is used colloquially by musicians, the titles of such editions are often something else (and often in a foreign language, such as “opera omnia”).

**Other Observations on Formulation of Searches**

*Searching by Call Numbers and Recording Label Numbers*

While some catalogs allow call number searches, our current instance of VuFind does not have a call number index, and keyword searching for them only works in some instances. But while call number searching does not work well in VuFind (e.g., it has to be done as a keyword search and not a call number index search like in Voyager), it still works in ES because it is searching by keywords. There were thirty-two examples of searches in MPAL’s ES where patrons used entire call numbers or the first part of a classification number to find related materials:

- count basie biography
- count basie ml 410
- duke ellington ml 410
- duke ellington bibliography

It is also not unrealistic to think that patrons might want to search by a recording's label number, since most catalogs provide search options for ISBNs and ISSNs for print materials. Searchers attempted this in a handful of searches like *lpo-0014, 7.24356E+11, 777337-2*. Unfortunately this information is not usually reflected in MPAL’s catalog records.

**Common Descriptions, Natural Language Queries, Genre Queries, and Context Words**

As mentioned already with the examples *Mahler 1* and *complete works*, patrons regularly search with terms and phrases that make sense to them or that are used colloquially when discussing music and sources, which may or may not be in the bibliographic record. Additional examples in the data set include:

- handel messiah critical edition
- rodelinda in italian
- mamma mia! Book [for the text of a musical]
- grove encyclopedia [the title of this is in fact “dictionary” not “encyclopedia”]
- mgg sachteil [the abbreviation for *Musik in Geschichte und Gegenwart* and the name for a section of it]
dance collection

The last example in the list is particularly intriguing—somewhat like the earlier search examples of performances and albums, one wonders if the patron hoped to find everything in that category and then be able to browse, however it is hard to know what the searcher anticipated getting in return. Sometimes natural language queries appear, often in an attempt to find a smaller part of a larger work, such as the slow movement of Brahms's First symphony, anonymous chant from vespers for christmas day, and Chaconne (from Partita in D minor, BWV 1004); or for things other than musical works, such as in Reviews of Stravinsky Article by Robert Craft.

Another variation on natural language or colloquial searches is the use of the possessive form of composer names. Although not common (23 examples), patrons do this when searching for composer and title of a work, e.g., verdi's requiem. It seems unlikely that people do this when searching for books or other works, but musicians make works possessive to the composer, such as in the examples mendelssohn’s violin concerto, to differentiate between pieces with the same form/generic title. In rare cases searchers used the term “by,” such as Jeptha by Carrissimi.

Genre searches such as South Indian Vocal music and hindustani classical music show that people may want to search the way they might in Pandora or iTunes, although it is possible this person was looking for secondary materials and not recordings or scores:

- pop
- female pop
- women pop
- contemporary pop

Searchers also exhibit a desire to find things by genre and instrument or voice type, such as soprano arias [which is 'high voice’ in the LC subject heading], mozart satb sanctus, and baroque arias for medium voice. Other examples include marimba literature, organ literature, and organ techniques. Catalogs do not necessarily aid in these types of searches, even though they are natural constructions for users. Sometimes searchers add context words to their search like they would in Google in a way that will not necessarily help them in the catalog, such as daniel read composer.

DISCUSSION

Even given the difficulties of searching for music materials, MPAL patrons have embraced ES—its module has almost as many searches as the Undergraduate Library’s, which serves a much larger population. It also has twice as many searches as the Social Science, Health, and Education Library module, which also serves a much larger population than MPAL. One of the possible reasons for this is the fact that MPAL was an early adopter of developing an ES subject module that could be searched from our homepage, which means our patrons have had longer to grow accustomed to using it.

MPAL has lower average words-per-search ratio (3.375 or 3.349 depending on data set) than most other ES modules, likely because there are more composer plus title keyword searches for musical works and not as many pasted article citation searches, which tend to be longer. This is supported by the comparison of the average number of words in searches done in the Gateway books tab.
(4.048) vs. the Gateway articles tab (6.35). In addition, although two- and three-word searches are most common, MPAL has a significant number of single-word searches (12.4 percent). Such searches can work in music, when there are unique titles like *Turandot* and *Treemonisha* that are unlikely to appear for more than one composer or as terms in other disciplines. For this same reason, single- or even two-word searches are unlikely to be effective in most other disciplines. At around seven words per search a transition in search patterns occurs. Eight word and longer search strings are almost always some version of a title of a book, article, chapter or dissertation, etc. and strings with six words and fewer tend to be topical searches or combination composer/piece searches.

Other transaction log studies of ES have shown that “title searching and results display—of journal titles, article titles, and book titles—is being heavily employed by users.” However, in music, where title alone may not be sufficient to identify and retrieve a musical work, searches with a combination of composer name and elements of the title and/or additional information will always be most prevalent.

**Search Location Appropriateness and Context**

Even though discovery layers and federated search tools help with minimizing the number of silos and places in which scholars need to search, there are still issues with patrons attempting to use the ES box to find things it is not designed to find. Searchers see a box and search, without always understanding the context. This can happen on multiple levels. The MPAL page clearly states that the MPAL ES box searches for arts-related things, but obviously patrons do not always see or comprehend this, even after they type in many queries that do not provide (good) results. This is likely related to the number of visitors to MPAL from other disciplines who do not realize that there are various differently scoped versions of ES.

The following example could be a theatre set construction related search, which would work only moderately well in our tool. Or, it may have been conducted by an architecture or structural engineering student, who would have better luck using a different ES module.

- light weigh [sic] structures in architecture
- building research
- the evolving design vocabulary of fabric structure
- the engineering discipline of tent structures
- building research jan/feb 1972:22

It would be ideal if the system was smart enough to make suggestions: “you appear to need architecture resources—if you are not finding what you need, might we suggest tool X, Y, or Z?” While ES does this to an extent when it can in the generic ES, it does not do so in the subject modules, and in reality, can only go so far. It raises the question of whether we are we doing patrons a disservice by offering pre-defined subject modules. While this approach has some benefits for most users, it also creates different challenges for some. MPAL’s ES does not target all available relevant online tools and neither does the general ES, so interdisciplinary researchers still need to be cautious of silos, even well-intentioned ones created by librarians or traditional
ones created by vendors. It is difficult to inform patrons of this in one-box search settings—they see the box and are eager to get started without first having to read a lengthy set of instructions.

Search location context is also important when patrons use ES to try to find things that are described or linked on our website and not in ES, such as for any of our named special collections. Patrons also use ES to find tools such as Naxos, jstor, worldcat, and librarysource, some of which are targeted by ES and some of which are not. ES will at least provide a link to a tool, however (see figure 7).

Figure 7. Easy Search Post-Search Suggestion.

These particular tools are all also linked from the MPAL website (in fact, Naxos is linked further down the home page from the ES box) and we also have a separate tool that enables one to search for databases and online journals by name.

On some occasions, searchers used ES to look for help using library tools, such as in the following example:

rilm retrieval
rilm
using rilm

The library website, not the discovery layer, is a better tool for finding instructions, since help information is currently delivered via various LibGuides. However, this is not intuitive to patrons. On a related note, it is interesting to consider whether patrons searching for specific tools such as IMSLP expect to find results from non-library resources in our search layers, or if they simply do not differentiate in their minds what is an open tool and what is a library subscription tool.

Patron Knowledge Level

Many of the observations of this study are related to known-item searching, since a large percentage of people looking for music materials are looking for specific pieces of music. Earlier studies show that it is difficult to search for something if you do not know what it is. This can be seen in examples like ombramaifu handel (should be Ombra mai fu) or the interworkings of tennis (which was followed by the correct inner game of tennis).

Topical searches can be especially difficult in any subject when the patron does not quite know how to put what they want into words (or literally does not know the right words, especially in the case of our many patrons for whom English is not their first language).
qualitize musical tension
spell change click: kw:qualitize musical tension
quantize musical tension
quantitative musical tension
music motive similarity
surveying musical form through melodic-motivic similarities
a paradigmatic approach to extract the melodic structure of a musical piece
inding subse- quences of melodies in musical pieces
spell change click: kw:finding subsequences melodies musical pieces
similarity measures for melodies
measures of musical tension
measuring musical tension

This echoes Head and Eisenberg’s 2009 findings and Dempsey and Valenti’s 2016 findings.37

**Shortcomings of the Easy Search Tool**

This study helped illuminate some shortcomings in ES. Sometimes the search formulation changes from ES to the target, for example *cramer preludes* in ES becomes *all(cramer preludes)* [a bound phrase] in one target, resulting in many fewer results than if the search had been done in the native interface. Patrons may not realize this as they are searching. In another case there were no results for *Danças folclóricas brasileiras e suas aplicações educativas* but removing the diacritics retrieves this title in our catalog, so it appears that diacritics do not function in ES (at least when VuFind is the target)—something that may not be apparent to searchers and hopefully can be addressed in the code.

**Further Research**

Additional analysis could be done on this data set, including assessing whether searches were for known items or topics, and more specifically whether for articles, books, scores, or recordings. However, in many cases it is difficult to tell if a patron is looking for a score, recording, or information about a piece or composer. Other research on ES shows over half of searches (just over 58 percent in 2015) in the main ES are for known items.38 This percentage is likely to be much higher in MPAL’s ES. With an enhanced data set it would also be possible to identify which target tools searchers are choosing most often.

**CONCLUSION**

While many patrons (and librarians) are eager for a tool that can truly search everything, we are not there yet. Some have tried to make music-specific interfaces for library catalogs, but this work is not widespread.39 Perhaps because music students are often searching for things other than articles it would be better to have one tool that searches the catalog and streaming media tools
and one that only searches article indexes. Some schools have taken this approach—configuring their discovery layer indexes to include article content but not the local catalog. There were several observations in this data of patron search behavior are not fully supported by library systems in all cases, but perhaps should be (e.g., use of + signs, searching by record label numbers or genre names/types of music/formats). In some cases, this is an issue with the metadata standards in use and in others it is about needing more flexible search options based on the metadata that we already have. Newcomer et al. discuss this in their article outlining music discovery requirements.40

Tools like Easy Search and discovery layers solve some problems for users but can create others. Dedicated library catalogs are still generally the best tools for finding scores and recordings in our physical (and some online) collections, but not all libraries offer that tool anymore, instead offering a discovery layer as the primary search tool. In those cases, serious consideration needs to be given to facets, the ability to limit by format, and especially the FRBRization of items, which is particularly problematic for music. Additionally, there is a continued need for targeted instruction for music library users, because not only are the tools used in libraries less than perfect, the inherent challenges in searching for music because of its formats and titles are aggravated by musicians’ use of shorthand and colloquialisms to describe music materials.

ENDNOTES


7 Asher et al., “Paths of Discovery,” 476.


10 Sarah Dahlen and Kathlene Hanson, “Preference vs. Authority: A Comparison of Student Searching in a Subject-Specific Indexing and Abstracting Database and a Customized Discovery Layer,” *College & Research Libraries* 78, no. 7 (November 2017), 892, https://doi.org/10.5860/crl.78.7.878.

11 Ibid.


“Information for Inspiration: Understanding Architects’ Information Seeking and Use Behaviors to Inform Design,” *Journal of the American Society for Information Science & Technology* 61, no. 9 (September 2010): 1,745-770, [https://doi.org/10.1002/asi.21338](https://doi.org/10.1002/asi.21338);


18 Tina Chrzastowski and Lura Joseph, “Surveying Graduate and Professional Students' Perspectives on Library Services, Facilities and Collections at the University of Illinois at Urbana-Champaign: Does Subject Discipline Continue to Influence Library Use?,” *Issues in Science & Technology Librarianship* 45, no. 1 (Winter 2006), [https://doi.org/10.5062/F4DZ068](https://doi.org/10.5062/F4DZ068).

19 Ellen Collins and Graham Stone, “Understanding Patterns of Library Use Among Undergraduate Students from Different Disciplines,” *Evidence Based Library and Information Practice* 9 (September 2014): 51–67, [https://doi.org/10.18438/B8930K](https://doi.org/10.18438/B8930K).

20 This is up from the 4.33 average reported by Mischo in 2012 (164).

21 Including direct from departmental webpage and via Gateway ES dropdown choices.

22 In Mischo’s 2012 analysis of Easy Search logs, 52 percent of sessions had one string and 48 percent had two or more. By 2015, single-query sessions had risen to 57 percent (William Mischo, et al., ”The Bento Approach to Library Discovery: Web-Scale and Beyond,” *Internet Librarian International*, October 21, 2015).


27 This echoes earlier research that shows most searchers use default settings and keyword searches.


30 We still have to use our older Voyager OPAC or the staff-side of Voyager to effectively search by call number until we get a newer version of VuFind.

31 Symphony no. 4 in E flat “Romantic” by Anton Bruckner, Klaus Tennstedt (Conductor), London Philharmonic Orchestra. (Performer).

32 This is Mozart, “Clarinet Concerto in A, K. 622,” Meyer/Berlin Philharmonic/Abbado EMI Classics 57128; 7.24356E+11.

33 This is REICH: Sextet / Piano Phase / Eight Lines (Griffiths Kevin/ London Steve Reich Ensemble/ The/ Stephen Wallace) (Cpo: 777337-2)).


35 This reinforces what Lown and Asher et al. found as cited in the literature review above.


