Utah State University

DigitalCommons@USU

Library Faculty & Staff Publications   Libraries

2-17-2021

Missing the MARC: Utilization of MARC Fields in the Search Process

Liz Woolcott
Utah State University

Andrea Payant
Utah State University

Becky Skeen
Utah State University

Paul Daybell
Utah State University

Follow this and additional works at: https://digitalcommons.usu.edu/lib_pubs

Part of the Library and Information Science Commons

Recommended Citation

UtahState University
MERRILL-CAZIER LIBRARY

Missing the MARC : Utilization of MARC fields in the search process

Authors:

Liz Woolcott, Head of Cataloging and Metadata Services, Utah State University (liz.woolcott@usu.edu)

Andrea Payant, Metadata Librarian, Utah State University (andrea.payant@usu.edu)

Becky Skeen, Special Collections Cataloging Librarian, Utah State University (becky.skeen@usu.edu)

Paul Daybell, Archival Cataloging Librarian, Utah State University (paul.daybell@usu.edu)

Abstract

Utah State University Cataloging and Metadata Services (CMS) unit analyzed MARC record discoverability within the libraries' discovery layer, Encore, using web analytics, a web-scrapping tool, and a relational database to examine MARC records listed in users' search results. MARC records were identified, isolated, and coded to pinpoint where search terms appeared, determine whether they were present in full or in part, and ascertain prominent fields not present in records. Analysis of results showed that notes and summaries were important for record retrieval and that users interacted with authorized name fields more frequently than authorized subject fields.

Keywords: MARC, field usage, user search patterns, catalog, discovery layer, discoverability

Introduction

The Cataloging and Metadata Services (CMS) unit at Utah State University (USU) Libraries consistently evaluates its practices and workflows to identify areas of possible improvement. Past research projects have focused on areas such as the study of communication patterns between units, analysis of assessment tools for technical services, effective methods for outreach in technical services, and developing crowdsourcing methods for metadata creation. Recent discussions for improving cataloging services led to the decision to conduct an in-depth study of user search behavior connected with all areas of cataloging, including traditional MARC record creation, and non-MARC metadata creation (i.e. digital collections and archival finding aids). This work represents the first phase of research aimed at accomplishing the goal of gathering data about how library users interact with metadata provided and produced by CMS and using the data to drive work in the directions needed to improve discoverability of resources. The focal point of the first stage of investigation was traditional MARC record creation.

In the fall semester of 2019, the research team, comprised of the Head of Cataloging and Metadata Services, Metadata Librarian, Special Collections Cataloging Librarian, Archival Cataloging Librarian, and the EAD specialist, began the process of MARC record evaluation. The assessment sought to understand the correlation between user search terms, the placement of MARC records in search results lists, and the performance of individual MARC fields. The overall research questions were:

- What is the frequency and placement of MARC records in search result lists?

- Where are search terms located in MARC records?

With the support of unit staff and student workers, the research team was able to gather and code robust and informative data about user utilization of MARC records in the searching process within our online discovery layer. The data collected will be used to take definitive steps toward enhancing CMS services.

Literature Review

There have been relatively few studies specifically involving MARC records and discoverability in recent years. The trend in research focus has understandably shifted heavily toward emphasis on BIBFRAME and discoverability using the semantic web and linked open data. Although, there have been relevant works published in closely related areas of research that compliment MARC specific studies. They include the assessment of user search behaviors in library catalogs and the influence metadata has on the circulation of library materials. Findings from these areas of study are potentially impactful to cataloging practices and could have a beneficial influence on improving the discoverability of resources.

User-search behavior and library catalogs

There have been some valuable research contributions regarding user search behavior and library catalogs. For instance, research has been conducted to examine user search behavior using specific discovery tools, like the analysis of VuFind and Primo done by Niu, Zhang, and Chen.[1] Their results lend insights into typical search behaviors across both library search platforms as similarities arose, such as the dominance of keyword searching, prominence of two term or three term word searches, brief search sessions, and less use of facets compared to text searches. Their findings also showed some differences across platforms, which indicates that the framework of a discovery layer can influence user search behavior.

One research study, by Lown, Sierra, and Boyer, similarly analyzed library user search behavior, but with a focus on the ways people use a single search box.[2] Drawing from two semesters of data collected using custom logging software, the authors gathered information from over 700,000 searches. Their research led to the conclusion that some general lessons learned could be applied outside their institution. First, they found that the search box is used for more than just articles and to search the catalog. For example, users searched for things like journals and databases that were not part

of the system's target index. Second, they stress the value of continuous evaluation of user search behavior in improving search capabilities. Third, the ongoing evaluation of the most common searches may reveal an underlying desire for specific resources or offer implications as to whether search results meet user expectations. For instance, studying the source of retrieval, like the OPAC or a separate database, can indicate high or low usage of a particular resource. Or, the frequency of a search may indicate retrieval or non-retrieval if evaluation of user-search behavior shows modifications to search queries. And lastly, the authors emphasize the importance of understanding how best to integrate tools, like the single search box, with other tools and services and consider this a key consideration to improving the search experience for library users.

Other work, like the dissertation produced by Fredrick Lugya,[3] examined the search behavior of graduate students working from different levels of experience with library catalogs. Students outlined how they planned to search for resources using their institution's discovery layer (VuFind) and other databases like Google Scholar. The actual search behavior was then recorded and analyzed after completion of a series of information tasks or scenarios. Findings from this study included that: search behavior is exploratory, basic searches are used most often because of the low effort involved, advanced searches are rarely performed overall, but, when implemented, users seek to narrow large result lists, users believe search results will be highly relevant to their search, and ease of use is a major factor to consider when building a searchable database.

Cataloging Practices – Influences on Circulation

Gunnar Knutson published an early study on the correlations between catalog access points and circulation.[4] The goal of the work was to determine if there was an easy, quantitative way to improve access. The study focused on catalog records from different classifications, or levels, and ultimately showed some evidence that the presence of access points in catalog records may not influence higher circulation rates. The authors suggested it might not be cost effective to spend time producing access points if the underlying goal of the practice is to increase circulation. In contrast to these findings, research conducted by Gross and Taylor demonstrates the value of subject headings in bibliographic records.[5] The authors evaluated the relationship between keywords and their presence being found exclusively in the subject fields of catalog records. They determined that, if subject headings were not present in records, around 35.9% of

keyword searches performed would have been unsuccessful in producing relevant hits. Therefore, without subject headings, users would experience a considerable decrease in the number of relevant search results. A follow-up article in 2005 by Gross, Taylor, & Joudrey found that with the addition of table of contents, summary, and abstract notes the decrease in relevant search results as a consequence of subject headings not being present in records would change from 35.9% to 27% of searches.[6]

Cherie Madarash-Hill and J.B. Hill studied the impact of electronically enriched content in MARC fields on the usage of books.[7] The electronically enhanced bibliographic records for books emulate those found on the widely used website, Amazon.com, by including links to things like dust jackets, book reviews, and tables of content. Ultimately, research indicated that the use of searchable URL enhancements in catalog records can result in higher checkout rates for materials. Tosaka and Weng similarly researched content-enriched bibliographic records and their effect on usage.[8] They also studied online public access catalog (OPAC) searches for circulated items to assess the use of enriched metadata. The circulation and search data showed that enhanced records were linked to higher circulation rates. Furthermore, their findings suggest that formatted contents notes may be important factors in discovery and emphasize the continued inclusion of this data in records to increase the value of the catalog for users.

Laura N. Kirkland, in *The Relationship of Metadata to Item Circulation*, outlines a study on the correlation between OPAC searches and the circulation of materials during the same period of time.[9] The result analysis of their research showed that around 25% of searches could be measured a success due to the resulting checkout of materials searched for by the user. The study also revealed data on what information users found most helpful (e.g. title, contents, subject, and author information) for discovery and access of items, which closely related to the goal of this current study to determine the usefulness of MARC records for discoverability.

MARC Records and Discoverability

One of the most relevant publications to specifically evaluate MARC usage to inform more efficient and effective cataloging practice was a report produced in 2010 by the RLG (Research Library Group) Partnership MARC Tag Usage Working Group and OCLC Research.[10] This group's efforts built upon previous research, referenced in their report, conducted by William Moen as well as Karen Calhoun with Diane Cellentani. Moen analyzed Worldcat records for field

utilizations based on frequency counts to show evidence of catalogers' use of MARC content designation. Calhoun and Cellentani surveyed users to identify the most essential data elements of MARC and to figure out what data quality enhancements would be beneficial. The results of the survey showed that the most desired enhancements clearly impacted cataloging decisions. The most desired improvements concerned author, item details, and links to online or full text content. Changes to these record fields would be impacted by cataloging decisions because it would involve creating additional content for records. For example, adding more links to online content, subject headings, summaries or abstracts, and tables of content. As a result of survey findings, the MARC Tag Usage Working Group focused on the analysis of 5 areas: occurrences of MARC tags in Worldcat, MARC tags used for matching records, MARC tags indexed and aggregated in several aggregate databases, MARC fields represented at different encoding levels, and the use of form/genre designators and relator terms. Their findings applicable to end-users emphasized the importance of main entry fields (1XX), formatted content notes (505), summary, etc. (520), subject access (65X), and electronic access (856). There were other important key findings from their research that can have implications on cataloging practices and the group gives suggestions on ways to possibly make positive changes. For example, improving field consistency and accuracy as well as emphasizing response to local user needs. The group acknowledged an inability, due to lack of information, to show data about how often values are searched, which fields users are retrieving results from, and whether the results help users find a desired resource. Research conducted in this work seeks to address these points, with data gathered from a local online catalog and will provide some insights into how metadata practices can be enhanced to better serve the needs of end-users.

Methodology and Process

Utah State University is an R2 research institution serving a population of 27,000 students (24,000 of which are undergraduates) in 9 campuses around the state of Utah. Predominant fields of study for USU undergraduates include Communicative Disorders & Deaf Education, Economics, Psychology, Mechanical Engineering, Biology, Elementary Education, Human Movement Science, and Computer Science. USU Libraries uses Encore, a product of Innovative Interfaces, Inc. (III) as its discovery layer for library resources. Encore pulls together records from the library's catalog, Sierra, as well as journal articles from subscription databases into a single search interface to simplify the research

process for users. Sierra houses the library's roughly 2.5 million MARC-based records, including physical and electronic resource material. Fifty-one databases, many of them EBSCO databases, feed just over 3.6 million non-MARC records into Encore. While Encore includes this broad array of information sources, it is still the primary method for searching the library's catalog because it is the only search box presented to users on the library's homepage. While there is an additional link on the library homepage titled "Books & Media" that exclusively searches Sierra, the research team decided to examine only data from Encore because it represents the most common search path for library users from the main USU campus.

The USU Libraries' configuration for Encore on its homepage includes only a simple search box, without the option to browse search or conduct an advance search (although this option is available once the first search is performed.) Encore displays a results list that combines together Sierra's MARC records with records from some of the library's subscription databases into a single results list. From that list, users can click on titles to access the records in Sierra or the respective database from which the item is indexed. In many instances, Encore also provides links to full text copies directly on the search results page that bypass record pages. For physical materials, users can access call numbers and availability status of physical items from the search results list without visiting the record pages. Throughout this study, the research team will indicate where items showed up in search results lists and which items were viewed by the patron. However, while items that are viewed demonstrate some interest on the part of the user, they are not considered by the research team to be the sole mode of indicating interest in the item because the information needed to access the actual item (physically or electronically) are available without ever viewing the records.

In order to determine how MARC records interacted with the user search process, the research team examined the logs of URLs that were generated by Encore. This method allowed the researchers to examine user searches while preserving user anonymity. The research team chose three non-consecutive days in the last half of the Fall 2019 semester to get a sample of data that would represent user searches on different days of the week (Monday, Tuesday, and Thursday) as well as different weeks in the semester (weeks 12, 13, and 15 out of a 16-week semester.) Monday, Tuesday, and Thursday were chosen because they represented the days of the week with the highest amount of traffic and the weeks were chosen to minimize the amount of library instruction-related searches that typically takes place in the first half of the semester.

Each search session in Encore generates a combination of static and dynamic URLs. Static URLs referenced pages such as the homepage, time out page, blank advanced search page, or a permalink to a record page. The information in the static URL does not include search data. Dynamic URLs, however, point to pages such as a list of search results or a single bibliographic record clicked on from a search results list.   Each time a search is conducted in Encore, the system generates a dynamic URL that captures a user's search terms and any facets selected, advanced search categories used, additional search result pages accessed, or bibliographic record numbers for MARC record pages. As an example, the following URL "http://discover.lib.usu.edu/iii/encore/plus/C__S(selective attention) (sport)__P1__Orightresult__U__X0__W?lang=eng&suite=cobalt" demonstrates:

- Advanced search: "S(selective attention)(sport)" where the parentheses around the portions of the search string "selective attention" and "sport" show the user input two sets of keywords into separate keyword search indexes on the Advanced Search page with the Boolean operator of "AND"

- Use of Facets: "X0__W" which indicates a selection of the "Article" facet in Encore

- Additional search results pages visited: "P1" indicates that user went to the first page after the initial search results. For the initial search results page, no "P#" is shown. All subsequent pages of search results clicked on by the user include a P1, P2, P3, etc. in the URL.

If the user had also clicked on a listing in the results that led to a record in Sierra, an additional "/record/C__b######" would have preceded the search terms. An example of this is shown later on in this section. Every time an action is taken (for example when a facet is used or search term is modified) a new dynamic URL is generated under those changed parameters.

For each of the three days, the research team would use Google Analytics to pull a list of the Encore URLs that were generated outside the library staff's IP ranges during the identified 24-hour period. These URLs were pulled early the following day to enable the research team to recreate the user's search environment as closely as possible. Within Google Analytics, a special report was created that would time stamp the URLs down to the minute they were accessed. This report was downloaded, and dynamic URLs that led to a search results page were isolated from the main report and fed into Octoparse, a web scraping tool. Each dynamic URL was loaded by Octoparse and the resulting page scraped for the search terms used, the number of results on the page, the total number of results available to the user, and the title

and link of each item in the list of results presented to the user on that page. Below is an example of an Encore search results page created from one of the user searches pulled from Google Analytics:

Image 1: Encore search results page

In this example, the dynamic URL was

https://discover.lib.usu.edu/iii/encore/plus/C__Skenneth%20heineman%20campus%20wars_____Orightresult?lang=eng&suite=cobalt" From the resulting page, the webscraper would pull the search terms "Kenneth heineman campus wars" from the search box, the results statement "Results 1-25 of 623", and the text and hyperlink associated with each title in the list (note there are 25 results for each page, but this example only shows the first 6).

The output from Octoparse was a spreadsheet of these elements from each URL's resulting webpage. The researchers numbered the results to show the order in which the search results appeared, and placed the data into Airtable, a relational database, for coding and analysis. All totaled, the research team identified and scraped the data from 3,564 dynamically generated URLs for search results lists, totaling 68,022 search results in all. The 3,564 dynamically generated URLs represent 1,040 search sessions over three days. On average a search session included 3 dynamically generated URLs because users often refined or repeated their searches with new parameters. Search sessions range between 1 and 32 URLs, with a mode of 1 URL and, as noted, an average of 3 URLs in a search session. However, this paper will look predominantly at individual dynamic URLs and the MARC records that resulted instead of the search sessions because the primary research questions within this paper explore how and where the search terms were found in the records displayed to users.

Within Airtable, the research team linked each batch of item titles and hyperlinks (most typically 25 results shown per page) to their corresponding dynamic URL. The research team was heavily supported by the CMS unit members and student technicians, who provided quality control by spot checking the web scraped search result lists as well as all null outcomes, such as URLs that returned no results at all. When MARC records were part of the search results list, the link provided looked like this:

http://discover.lib.usu.edu/iii/encore/record/C__Rb2403851__Skenneth%20heineman%20campus%20wars__Orightresult__X2?lang=eng&suite=cobalt. The string "/record/" indicated a MARC record and the the string "/C__Rb2403851" is the bibliographic number assigned to the record in Sierra. Additionally, the search terms used to access the results list that led to this record (which show up as "Skenneth%20heineman%20campus%20wars") as well as any facets, advanced searches, or pages selected are also retained in the URL. Along with the time stamp, this information helped the research term aggregate search sessions together.   The bib record number was extracted from the URL and used to

generate a link to the MARC display. The link looked like:

http://tahoe.lib.usu.edu/search~S1?/.b2403851/.b2403851/1%2C1%2C1%2CB/marc~b2403851 and Image 2

demonstrates a typical display of the MARC record in Sierra's web display.

Image 2: Sierra MARC display

The item results from each search were then sorted into categories of "Sierra" or "other" to filter out non-MARC

records.   This left 13,312 results that originated from MARC records in Sierra.

Once the MARC records were identified and isolated from the other search results, the researchers and CMS unit used

the MARC display link to access and copy the full text of the MARC record into Airtable. The MARC record was then

analyzed and coded for the material type represented in the record and whether it was originally generated by library

personnel or was a vendor supplied record. Following this, each of the 13,312 records were coded for where the search

terms appeared in the record, whether the search terms were present in full or in part within those fields, and

prominent categories of fields that were not present in the record (such as author, subject, contents notes, or summary

notes).   For every instance where the search term appeared in the field, that field was copied into a separate column for

further analysis. As noted, search sessions averaged 3 dynamic URLs per session, with common changes between URLs being search term changes. These modifications in search parameters were often slight, so the same record could show up in more than one search results list. However, all records were retained even if they appeared more than once because shifts in URLs indicated an action on the part of the user and analysis of each step in the search query showed that users typically changed search terms rather than applied facets or used the advanced search to refine their search. Therefore, each time a record appeared in a results list, the search terms that summoned it had a high probability of being different from one another.

At the outset, the research team only planned to code a subset of records. However, due to extended work from home arrangements necessitated by the COVID-19 pandemic, the CMS unit and student technicians as a group participated in coding all 13,312 MARC records and provided quality control on each other's work. With unit members and student technicians working part time on the coding, the process took about 3 months to complete.

Finally, the search terms for each dynamic URL were analyzed by the research team for known item searching. Known item searching is a process where the user searches for an item they are already familiar with (such as a title or author) rather than using generic topical keywords. The search terms employed by the user were analyzed through a multi-step process that reran the same terms in a browser (typically Google), as well as Google Scholar and Microsoft Academic, to ascertain if the search terms reasonably matched the title or identifier of a known item. When found, the corresponding URLs were tagged as Known Items and coded for the media type (book, video, article, etc.) and keywords used (title, creator, DOI, etc.). Searches that looked for known items comprised roughly 54.8% of all search queries. Following this coding, each known item was double checked by a research team member to determine if the library provided access to it, either physically or in electronic format.

Analysis and Results

Research Question 1: What is the frequency and placement of MARC records in search results lists?

Analysis 1.1: How frequently are MARC records showing up in search results?

As previously noted, 68,022 results were displayed to users across 3 days in November and December 2019. As shown in *Table 1: Count of catalog vs. database records displayed to users*, the number of MARC records showing up in those

search results was 13,312, representing 19.58% of the total results shown to users during their search process. This indicated that 4 out of 5 records were generated from harvested database content and not MARC catalog records. This appears out of sync with the general proportion of 40% MARC and 60% non-MARC records in the discovery layer.

Table 1: Count of catalog vs. database records displayed to users

	Batch 1	Batch 2	Batch 3	Combined
MARC-based catalog records	5264	3299	4749	13312
Records from other platforms	20326	17560	16811	54697
Total Records	**25603**	**20859**	**21560**	**68022**
Percent MARC records	*20.56%*	*15.82%*	*22.03%*	*19.57%*

Of the 13,312 MARC records logged in search results lists, 418 records were "clicked on" by the patron to open the record, representing 3.14% of the total MARC records listed over the three data sessions gathered. Within the Encore system, users can obtain pertinent information from the search results list without opening the record (including title, author, and call numbers), therefore the lack of "clicking" on the record to view it further is not necessarily indicative of lack of importance to patrons. Conversely, opening a record is also not entirely a measure of value, but rather an indicator of interest. The authors will use this definition of "interest" in all further discussion of record views in this article.

Analysis 1.2: Is there a difference between locally created records and vendor supplied records in the frequency of listing in search results?

MARC records displayed in search results were coded for the record's creator. Six types of record creators contribute to the MARC records in the Sierra catalog; the main two record contributors were the Cataloging and Metadata Services Unit (CMS) and Vendors. Other contributors include library departments such as Acquisitions and Resource Sharing unit and the Patron Services (Circulation) unit as well as satellite libraries including the distance campus libraries and the natural history library. With the exception of CMS and Vendors, all of the groups listed below primarily contribute brief records and compose less than 4% of all records displayed to patrons. A complete breakdown is available in *Table 2: Records displayed in search results and records accessed, sorted by record creator.*

Table 2: Records displayed in search results and records accessed, sorted by record creator

Record Creator	# Records in results list	% Total records in results list	# Records accessed	% Total records accessed
Vendor	7,727	58.05%	163	39.00%
Cataloging and Metadata Services (CMS)	5,066	38.06%	239	57.18%
Distance Campus Libraries	410	3.08%	5	1.20%
Record Unavailable at time of coding	52	0.39%	2	0.48%
Patron Services, Library Media Collections, or Resource Sharing and Document Delivery	33	0.25%	8	1.91%
Acquisitions	16	0.12%	0	0.00%
Unknown	5	0.04%	1	0.24%
Natural History Library	3	0.02%	0	0.00%
Total	**13,312**		**418**	

In looking at the rate in which vendor or CMS created records appear in search results listing, vendor supplied records are appearing more often – with 58.05% of all MARC records listed in search results. Records crafted by CMS supply 38.06% of all MARC records found in search results lists. This proportion flips, however, when looking at only the records which were clicked on or accessed by patrons. In this case, vendor supplied records support 39% of patron-viewed records and CMS records supported 57.18% of patron views. The research team could not ascertain the total number of MARC records in Sierra created by vendors or CMS and therefore cannot contextualize these proportions.

Analysis 1.3: How are MARC records ranked in the search results list?

When analyzing where in search results lists MARC records tended to appear, it became clear that MARC records did not appear in the top five results very often. As a default, each search results page shows a maximum of 25 results at a time. *Table 3: MARC record position number in results list ordered by frequency* shows that the most common position for MARC records in a search result set of 25 items, is position 4. MARC records appear in the top five search results 25.35% of the time.

Table 3: MARC record position number in results list ordered by frequency

Position Number	Total Records	% of total Records	CMS Records		Vendor Records		All other Records	
			# Records	% In this position	# Records	% In this position	# Records	% In this position
4	999	7.50%	315	31.53%	652	**65.27%**	32	3.20%
8	803	6.03%	256	31.88%	527	**65.63%**	20	2.49%
2	779	5.85%	278	35.69%	465	59.69%	36	4.62%
12	695	5.22%	236	33.96%	424	**61.01%**	35	5.04%
6	645	4.85%	269	41.71%	350	54.26%	26	4.03%
16	609	4.57%	231	37.93%	363	59.61%	15	2.46%
3	582	4.37%	223	38.32%	325	55.84%	34	5.84%
20	540	4.06%	203	37.59%	327	**60.56%**	10	1.85%
10	532	4.00%	204	38.35%	297	55.83%	31	5.83%
1	525	3.94%	244	46.48%	264	50.29%	17	3.24%
24	516	3.88%	194	37.60%	309	59.88%	13	2.52%
5	490	3.68%	202	41.22%	260	53.06%	28	5.71%
14	474	3.56%	184	38.82%	272	57.38%	18	3.80%
9	464	3.49%	208	44.83%	241	51.94%	15	3.23%
18	459	3.45%	180	39.22%	264	57.52%	15	3.27%
7	447	3.36%	189	42.28%	241	53.91%	17	3.80%
13	437	3.28%	149	34.10%	258	59.04%	30	6.86%
15	433	3.25%	189	43.65%	229	52.89%	15	3.46%
11	432	3.25%	205	47.45%	204	47.22%	23	5.32%
22	431	3.24%	183	42.46%	233	54.06%	15	3.48%
17	425	3.19%	160	37.65%	238	56.00%	27	6.35%
21	412	3.09%	173	41.99%	230	55.83%	9	2.18%
19	410	3.08%	140	34.15%	258	**62.93%**	12	2.93%
25	387	2.91%	125	32.30%	248	**64.08%**	14	3.62%
23	386	2.90%	126	32.64%	248	**64.25%**	12	3.11%
Total	**13312**		**5066**		**7727**		**519**	

Similar to the overall demographics of MARC records in the search results list, the range of Vendor records as a percentage of the total MARC records for all search result positions was between 47.22% and 65.63%. CMS records ranged between 31.53% and 44.83% of the records at any single position in the search results list. Position 11 is the only search results position where CMS records were more common than vendor records – but only by the slimmest of margins. Overall, vendor-supplied records represented a larger share of records in search results lists.

Analysis 1.4 Where do MARC records for known items rank in the search results list?

Narrowing down to look at just the known items, the picture changes. Known item searching accounted for 54.8% of searches in the discovery layer. The two most common types of known items users wanted were books and articles, representing 42.3% and 43.3% of all known item searches, respectively. The remaining types of known items represented 14.4% of known item searches and included videos, journals, dissertations and theses, plays, book chapters, newspaper and magazine articles, newspapers or magazine (whole title), reports, audio, assessment instrument, and book series. For simplicity, this article will refer to books, videos, journals, dissertations/theses, plays, newspaper or magazine (whole title), reports, audios, assessment instruments, and book series (which are cataloged at the individual level per library policy) as "whole objects" and articles (journal, magazine, or newspaper) and book chapters (which are rarely cataloged at the individual level by USU Libraries' CMS unit) to be "sectional objects." Therefore, all successful sectional object searches (for article level or book chapters) are expected to come from databases indexed by Encore and excluded from the data set.

As noted in the methodology, the researchers verified if the whole object known items were available through USU Libraries. Items which are not available would theoretically not show up in search results, and therefore the expectation that they should rank high in the results lists is void. Whole object known items verified to be available at the library represented 68.2% of the known items list and were most often found in the first four results listed. *Table 4: Available whole object known items by search results placement* demonstrates that distribution by known item type.

Table 4: Available whole object known items by search results placement

Number of Available Whole Objects = 210	Available Whole Objects Appearing in Search Results		Percentage of Times Available Whole Object Appeared in Search Results by Position Number									Available Whole Object Not Listed in Search Results		
	Total Number	Percent-age	Result 1	Result 2	Result 3	Result 4	Result 5	Results 6-10	Results 11-15	Results 16-20	Results 21-25	Not Listed	No results returned	Direct Link
Book	603	90.3%	17.9%	14.8%	9.5%	8.0%	6.0%	16.6%	10.8%	8.8%	5.0%	2.5%	0.2%	0.2%
Journal	29	4.3%	34.5%	34.5%	3.4%	0.0%	0.0%	0.0%	3.4%	3.4%	3.4%	17.2%	0.0%	0.0%
Video	18	2.7%	33.3%	22.2%	11.1%	5.6%	0.0%	5.6%	0.0%	5.6%	16.7%	0.0%	0.0%	0.0%
Play	10	1.5%	10.0%	20.0%	0.0%	0.0%	10.0%	30.0%	10.0%	10.0%	10.0%	0.0%	0.0%	0.0%
Dissertation/Thesis	4	0.6%	25.0%	0.0%	25.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	50.0%	0.0%	0.0%
Newspaper	2	0.3%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%
Music	1	0.1%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	100.0%	0.0%	0.0%
Technical/Scientific Reports	1	0.1%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Governmental Report	1	0.1%	0.0%	100.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Total Number	668	100%	125	107	61	49	37	104	67	56	35	25	1	1
Percentage in Result Category	n/a	n/a	18.7%	16.0%	9.1%	7.3%	5.5%	15.6%	10.0%	8.4%	5.2%	3.7%	0.1%	0.1%

Titles with multiple editions or variations, such as Oxford English Dictionary or the Diagnostic and Statistical Manual of Mental Disorders (DSM), often showed up in more than one position because these titles have more than one record that each reflect a different edition. Additionally, materials available in more than one format (for example a title available as an e-book and also a physical book) had more than one record. Books made up the largest portion of the category types for available whole objects and most commonly appeared in the top five search results. Books were the only category of materials where users had a direct link to the record item. These are noted at the end of the table because they would not show up in the search results but were counted among the known items. Journals were the most likely category of material to appear in the top 2 search results but also half as likely to not be listed at all in the search results. Dissertations and theses, newspapers, and music also had high rates of not showing up in the search results.

One of the key issues discovered in verifying the availability of known items was the frequency with which book reviews saturated the results list. Records for books were frequently buried beneath an average of two book reviews with a display title similar or exactly the same as the item and poorly labeled as book reviews. Many of these reviews included hyperlinks to the electronic version of the reviews, which made the object appear to be available electronically even when it was not. This is an area where the library can potentially work with the vendor to clearly identify and downrank reviews.

Research Question #2: Where are search terms located in MARC records?

Analysis 2.1: What fields are used most in retrieving records?

To determine which fields played the most significant part in the retrieval of records, each catalog record that appeared in a search results list was analyzed and coded to indicate the fields in which the user's search terms were found. As shown in *Chart 1: MARC fields where search terms were located*, the top four fields used in record retrieval were the 245 (Title Statement), 505 (Formatted Contents Note), 650 (Subject Added Entry – Topical Term), and 520 fields (Summary) fields. The 245 field was particularly significant in record retrieval, appearing 80% more frequently than the 505, the next most frequently used field.
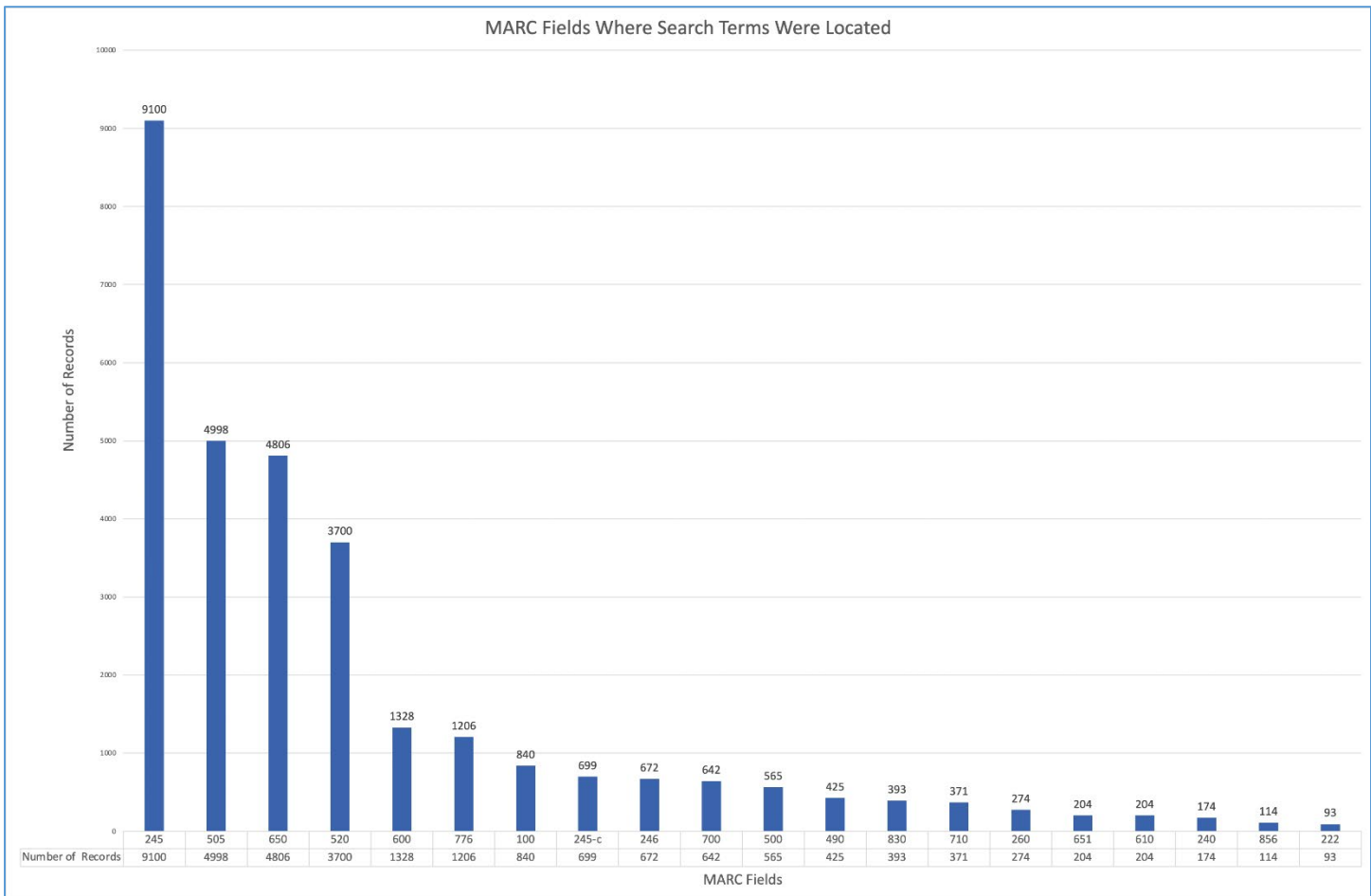
Chart 1: MARC fields where search terms were located

The CMS unit analyzed how often all terms were found in a single field. During the coding process, the CMS coding team indicated not only which field a term was located in, but also whether all terms within the search string were found in that field or if only part of the terms were found. The order in which the user put the terms was not a consideration during this process, the coders simply looked for the presence of each individual term within the field, with the exception of common words such as "the," "a," "an," "and," etc. Users' search strings ranged between 1 and 63 words, with an average of 5 terms and a mode of 2 terms. It was not expected that all search terms would be appropriate within a single field, but an analysis of how frequently they did occur within a single field and which fields this most commonly affected would illustrate a small portion of intersection of user searches and MARC records.

Table 5: Prevalence of user search terms, in full or in part, in the top 20 MARC fields

Rank	Field	Field Description	# Records	% All Terms in Field	% Part of Terms in Field
1	245	Title Statement	9100	79.02%	20.98%
2	505	Formatted Contents Note	4998	54.88%	45.12%
3	650	Subject Added Entry - Topical Term	4806	41.57%	58.43%
4	520	Summary, etc.	3700	50.81%	49.19%
5	600	Subject Added Entry - Personal Name	1328	86.67%	13.33%
6	776	Additional Physical Form Entry	1206	72.89%	27.11%
7	100	Main Entry - Personal Name	840	72.74%	27.26%
8	245-c	Statement of Responsibility	699	71.96%	28.04%
9	246	Varying Form of Title	672	62.35%	37.65%
10	700	Added Entry - Personal Name	642	59.03%	40.97%
11	500	General Note	565	44.78%	55.22%
12	490	Series Statement	425	16.47%	83.53%
13	830	Series Added Entry - Uniform Title	393	11.20%	88.80%
14	710	Added Entry - Corporate Name	371	31.00%	69.00%
15	260	Publication, Distribution, etc. (Imprint)	274	14.23%	85.77%
16	651	Subject Added Entry - Geographic Name	204	23.53%	76.47%
17	610	Subject Added Entry - Corporate Name	204	36.27%	63.73%
18	240	Uniform Title	174	60.92%	39.08%
19	856	Electronic Location and Access	114	57.89%	42.11%
20	222	Key Title	93	83.87%	16.13%

Table 5: Prevalence of user search terms, in full or in part, in the top 20 MARC fields shows that for all records where search terms were found in the 245 Title statement, 79% of the time it included all terms within the search string. Less so, the next most utilized field, the 505 Formatted Content Notes, contained the all search terms 54.88% of the time. The 650 Subject Added Entry – Topic Term field, which was the third most utilized field, showed that over half of the time (58.43%) search terms were only found in part. With the exception of the 600 Subject Added Entry Personal Name field, all 6xx subject fields were more likely to include only part of the terms in the search string rather than the whole. Interestingly, all fields with proper names (600 Personal Name, 100 Main Entry Personal Name, 245 subfield c Statement of Responsibility, and 700 Added Entry Personal Name) were more likely to hold all terms used in the search string rather than just part of terms within the search string. Results of this analysis were further supported by coding of known items. As noted in the methodology, roughly 54.2% of all search query sessions included searches for known

items, which were searched for predominantly using title or author terms. Ninety-five percent of searches for known

items included title and 21% included the creator, with 19% including both.

When a user clicked on a hyperlinked name or subject field, the system generated a new URL with the field terms as the

search term.   While this did not happen very frequently (only 35 dynamic URLs in 12 search sessions included either

name or subject fields), the most commonly occurring authorized headings appearing as search terms were names,

accounting for 65.7% of the dynamic URLs and 10 of the 12 search sessions. Further analysis also showed that only one

of the two search sessions that used authorized subject headings was generated from clicking on the heading. The other

was directly input into the subject index search in an advanced search page. In this search session, the user input three

subject terms into the subject index on the advanced search page "cancer -- history -- seventeenth century," "cancer --

history -- 17th century," "cancer -- history." These subject headings never appeared in a search results list or a record

prior to their time stamps. The authors assume this was a search session done by another librarian or an advanced user

who understood subject heading construction or had copied it previously and pasted it into the subject index on the

advanced search page. All other authorized headings that appeared as search terms were present in search results lists

prior to their dynamic URL time stamp. It should be noted that subject fields are not well displayed in the user interface

for Encore. To access authorized subject headings, the user must first open the record for viewing and then "click on"

the hyperlinked text, which appears toward the bottom of the record. See *Image 3: Encore Record Page Example*.

My Temporary List (0 items) | My Account | Logout

Utah State University
UNIVERSITY LIBRARIES
Request Materials

‹ Back to results

Campus wars : the peace movement at American state universities in the Vietnam era / Kenneth J. Heineman
Heineman, Kenneth J., 1962-

Book | New York University Press | 1993

Available at Available Merrill-Cazier Books (2nd Floor South) (Call number: DS 559.62 .U6 H45 1993)

Request it

Additional actions:

Other Sources

View in Library Catalog

Items

Location	Call No.	Status
Merrill-Cazier Books (2nd Floor South)	DS 559.62 .U6 H45 1993	Available

Details

Description	xvi, 348 pages ; 24 cm.
Content	text
Carrier	volume
Bibliography	Includes bibliographical references (pages 315-325) and index.
Subject	Vietnam War, 1961-1975 -- Protest movements -- United States.
	Peace movements -- United States -- History -- 20th century.
ISBN	0814734901 (alk. paper) : $40.00

Google Preview

Bookmark This Record

Persistent link to this record: https://discover.lib.usu.edu/iii/encore/record/C__Rb2403851

↑ Back to top

Recently Viewed

• Campus wars : the peace movement at American state universities in the Vietnam era / Kenneth J. Heineman. by Heineman, Kenneth J., 1962-

Search Tools

Advanced Search
Encore Home
My Temporary List (0 items)
My Account
Logout

Library Links

Utah State University
UNIVERSITY LIBRARIES

Image 3: Encore Record Page Example

Analysis 2.2: For records accessed by the patron, is there a difference in where search terms are located?

Table 5: Prevalence of user search terms, in full or in part, in the top 20 MARC fields looked at all of the records displayed to a patron during the search process, and therefore are more indicative of the retrieval algorithm. Narrowing down the view to only records which were "clicked on", the picture changes subtly. For the purposes of this exercise, the authors are considering any record that was viewed by the searcher to be "of interest" and worth taking a closer look at, but do not want to indicate that this is the sole measure of record usefulness.

Table 6: Prevalence of search terms, in full or in part, in the top 20 MARC fields in records viewed by patrons

Rank	Field	Field Description	# Records	% All Terms in Field	% Part of Terms in Field
1	245	Title Statement	325	86.77%	13.23%
2	650	Subject Added Entry - Topical Term	198	33.33%	66.67%
3	505	Formatted Contents Note	108	36.11%	63.89%
4	520	Summary, etc.	105	46.67%	53.33%
5	245-c	Statement of Responsibility	54	85.19%	14.81%
6	100	Main Entry - Personal Name	47	63.83%	36.17%
7	776	Additional Physical Form Entry	44	88.64%	11.36%
8	490	Series Statement	28	10.71%	89.29%
9	830	Series Added Entry - Uniform Title	26	11.54%	88.46%
10	246	Varying Form of Title	23	47.83%	52.17%
11	600	Subject Added Entry - Personal Name	21	57.14%	42.86%
12	700	Added Entry - Personal Name	18	55.56%	44.44%
13	260	Publication, Distribution, etc. (Imprint)	17	11.76%	88.24%
14	500	General Note	17	29.41%	70.59%
15	710	Added Entry - Corporate Name	14	7.14%	92.86%
16	780	Preceding Entry	11	45.45%	54.55%
17	264	Production, Publication, Distribution, Manufacture, and Copyright Notice	7	0.00%	100.00%
18	222	Key Title	7	100.00%	0.00%
19	240	Uniform Title	6	50.00%	50.00%
20	651	Subject Added Entry - Geographic Name	6	0.00%	100.00%

When looking at just the record pages that were accessed during the search process, the distribution of terms, in full or in part shifted in order of frequency. The 245 Title statement still retained clear superiority, appearing 64% more often

than the next most utilized field. However, instead of the 505 Formatted Contents Note holding the second place, the

650 Subject Added Entry – Topical Term is the next most commonly used field. Even so, the percentage of times only

part of the search terms are found in the 650 Subject Added Entry – Topical Term increased to 66.6% of the time. The

505 Formatted Content Note and the 520 Summary fields both retain a spot in the top four fields utilized, but also shift

from all search terms appearing most of the time to only part of the search terms appearing in the field most of the

time. Similar to *Table 5*, which showed all records in a results list, *Table 6* (showing records accessed by the patron)

demonstrates that fields that reflect personal names continue to find the all terms from the search string more

frequently than only part of the terms from the search string.

Analysis 2.3: For locally created records and vendor-supplied records, is there a difference in where search terms are located?

When looking at the top 20 fields that were utilized in the search retrieval process, the CMS unit analyzed how often

those fields appeared in records created by CMS or vendor supplied records. As noted above, vendor records comprised

58.05% of all MARC records in the results list, so a similar distribution in the individual fields where search terms are

found should be expected, within range. Unexpected outliers included the 505 Formatted Content Note, 520 Summary

field, 776 Additional Physical Form, and 856 Electronic Location and Access fields, which were significantly different from

CMS records. This is consistent with the larger number of vendor-supplied electronic resources housed in the library

catalog. For fields that were more often utilized in CMS records, 6xx subject fields and controlled name and title fields

played a bigger role in record retrieval.

Table 7: Frequency and percentage of fields used in record retrieval, listed by CMS and vendor-supplied MARC records

Field	Field Description	Frequencies			Percentage of Records	
		# Total Records	# CMS Records	# Vendor Records	CMS % of Records	Vendor % of Records
245	Title Statement	9100	3986	4699	43.80%	51.64%
505	Formatted Contents Note	4998	1406	3481	28.13%	69.65%
650	Subject Added Entry - Topical Term	4806	1965	2719	40.89%	56.58%
520	Summary, etc.	3700	866	2813	23.41%	76.03%
600	Subject Added Entry - Personal Name	1328	796	434	59.94%	32.68%
776	Additional Physical Form Entry	1206	84	1122	6.97%	93.03%
100	Main Entry - Personal Name	840	503	276	59.88%	32.86%
245-c	Statement of Responsibility	699	385	277	55.08%	39.63%
246	Varying Form of Title	672	325	327	48.36%	48.66%
700	Added Entry - Personal Name	642	313	305	48.75%	47.51%
500	General Note	565	278	271	49.20%	47.96%
490	Series Statement	425	153	265	36.00%	62.35%
830	Series Added Entry - Uniform Title	393	142	247	36.13%	62.85%
710	Added Entry - Corporate Name	371	151	215	40.70%	57.95%
260	Publication, Distribution, etc. (Imprint)	274	149	116	54.38%	42.34%
651	Subject Added Entry - Geographic Name	204	69	123	33.82%	60.29%
610	Subject Added Entry - Corporate Name	204	127	76	62.25%	37.25%
240	Uniform Title	174	100	65	57.47%	37.36%
856	Electronic Location and Access	114	8	106	7.02%	92.98%

Analysis 2.4 What fields are not present in the records?

As noted previously, the MARC Tag Usage Working Group[11] identified five groups of tags that were important to users: main entry fields (1XX), 505 Formatted Content Notes, 520 Summary, etc., subject fields (65X), and electronic access (856). Using this grouping, but with a few modifications, the CMS unit coded for fields that were not present in the records. The CMS unit expanded the definitions of the groupings slightly, with 65x coding changed to any 6xx to indicate that there were no 6xx fields in the record. Additionally, the 1xx was refocused to code for all creator fields (1xx or 7xx) that were not present in in the records. In this case, the record would be coded as not having a creator or author present if it lacked a 100, 110, 111, 700, 710, or 711. Last, but not least, the unit did not code for 856 electronic access field because it is not applicable to all resource types. Results in *Table 8: Fields not present in the MARC record*

showed that Vendor records were more likely to have all fields present than CMS records. Of the records that did not

have prominent fields included, vendor records were slightly more likely to not have an author and subject fields than

CMS. However, vendor records were significantly more likely to have 505 (Formatted Content Notes) and 520 (Summary

Notes) than CMS records, although neither record creator exceed 55% of records for either note field. Given that both

505 and 520 notes play a significant role in search retrieval, addition of these fields could improve search results ranking.

Table 8: Fields not present in the MARC record

	CMS		Vendor	
	Not Present	**Present**	**Not Present**	**Present**
Author (both 1xx and 7xx)	0.75%	99.25%	1.18%	98.82%
Subject (Any authorized)	4.46%	95.54%	6.73%	93.27%
505-Formatted Contents Note	63.96%	36.04%	45.54%	54.46%
520-Summary Note	75.60%	24.40%	50.45%	49.55%
All Categories Present	14.86%		33.26%	

Analysis 2.5: Which fields would make the greatest impact if not included in the record?

Taking an approach similar to Gross and Taylor, but broadening out beyond subject fields, the researchers isolated all

records where any search terms (regardless of the full search string or only part) were found in only one MARC tag.[12]

Lack of these fields would prevent the record from being listed in search results at all. Records where search terms were

only found in one MARC field comprised 3,714 records, or 27.9% of all MARC records appearing in the search results.

Table 9: Frequency of records where search terms matched only one MARC field

Field	Field Description	# Records	% All single MARC field records	% All MARC records (n=13,312)
505	Formatted Contents Note	1602	43.13%	12.03%
245	Title Statement	1294	34.84%	9.72%
520	Summary, etc.	643	17.31%	4.83%
650	Subject Added Entry - Topical Term	73	1.97%	0.55%
245-c	Statement of Responsibility	35	0.94%	0.26%
500	General Note	20	0.54%	0.15%
600	Subject Added Entry - Personal Name	14	0.38%	0.11%
490	Series Statement	7	0.19%	0.05%
100	Main Entry - Personal Name	6	0.16%	0.05%
246	Varying Form of Title	5	0.13%	0.04%
700	Added Entry - Personal Name	3	0.08%	0.02%
545	Biographical or Historical Data	2	0.05%	0.02%
630	Subject Added Entry – Uniform Title	2	0.05%	0.02%
856	Electronic Location and Access	2	0.05%	0.02%
20	International Standard Book Number	1	0.03%	0.01%
24	Other Standard Identified	1	0.03%	0.01%
130	Main Entry – Uniform Title	1	0.03%	0.01%
511	Participant or Performer Note	1	0.03%	0.01%
597	Local Note	1	0.03%	0.01%
653	Index Term - Uncontrolled	1	0.03%	0.01%
Total		**3714**		27.90%

As is demonstrated by *Table 9*, the top four fields are once again the 505 Formatted Content Notes, 245 Title Statement, 520 Summary, etc., and 650 Added Entry – Topical Term, albeit in a different order this time. Without the presence of the 505 or 520 note fields, roughly 16.86% of all the records that appeared in the results list would not have shown up. In contrast, without the two most prominent subject fields (650 Subject Added Entry – Topical Term and 600 Subject Added Entry – Personal Name), only 0.66% of the records would not have appeared in the search results list. This shows a significance difference in impact between the note fields and the subject fields.

Table 10: Frequency of records where search terms matched only one MARC field in records viewed by patrons

Field		# Records	% of all single MARC field records viewed by patron	% of all MARC records viewed by patron
245	Title Statement	52	71.23%	12.44%
505	Formatted Contents Note	17	23.29%	4.07%
520	Summary, etc.	2	2.74%	0.48%
650	Subject Added Entry - Topical Term	1	1.37%	0.24%
020	International Standard Book Number	1	1.37%	0.24%
Total		**73**		**17.46%**

When looking exclusively at the records which patrons viewed, 73 of the 418 records had only one field include any of the search terms input by the user.   As demonstrated in *Table 10: Frequency of records where search terms matched only one MARC field in records viewed by patrons*, the most pivotal field is the 245 Title Statement, representing 71.23% of those 73 records and 12.44% of all of the records viewed by a patron that would not have been present without the search terms in the field. This is a switch from Table 9 where 505 Formatted Content Notes played the largest role in matching search terms. The 505 Formatted Content Notes field, representing 23.29% of the single field records and 4.07% of all records viewed by the patron, still plays an outsized role compared to the 520 Summary, etc., 650 – Subject Added Entry – Topical Term, and 020 – International Standard Book Number fields.

Discussion

The purpose of this research project was to understand the interplay between user search terms, placement of MARC records, and the performance of individual MARC fields. The CMS unit at Utah State University plans to use this information to structure local practices and procedures based on user search habits and needs. The data shown in Analysis 1.1 demonstrated that 80.42% of all records displayed in search results came from non-MARC records, even though non-MARC records only comprise 60% of the database. Additionally, Analysis 1.3 showed that MARC records only placed in the top 5 results 25.35% of the time, leaving non-MARC records to dominate the search results top 5 list 74.62% of the time. Both numbers point to a significant advantage to non-MARC records and lead the research team to wonder if full text might play a leading role in search retrieval.

Given that MARC records do not typically include the full text of an item, the next best approach for the CMS unit to make MARC records more applicable to user searches could be the inclusion of 505 Formatted Content Notes and where possible, the 520 Summary, etc. fields to help increase the relevancy of record in a way that is useful to the searcher without too much bloat. From the data shown in Analysis 2.1, the 505 and 520 fields were consistently in the top four fields that retrieved a record. They were also shown to contain the search terms at least half the time. The 505 Formatted Contents Note, in particular, showed a significant role in the retrieval process, representing 12.03% of the total records examined in this project that would not have been displayed to patrons had it not been included in the bibliographic record and 4.03% of individual records viewed by patrons. Vendor created records appeared more often in the search results lists than CMS created records, according to Analysis 2.3. Vendor records were also more likely to include the 505 and 520 fields, while containing roughly the same frequency of author and subject fields as CMS records, as show in Analysis 2.4.

Title fields out-represented any other field in search results matches. This is appropriate given the predominance of known item searching, as discussed in Analysis 1.4. However, it is interesting that 505 Formatted Contents Note ranked higher than 245 Title field for records where search terms matched only one field and was second only to Title in records which were viewed by patrons. This indicates that 505 Formatted Contents Note plays a bigger role than was previously understood and could potentially increase the likelihood that MARC records are displayed in search results if they were employed more frequently.

Results from subjects were mixed, with the 650 (Subject Added Entry – Topical Term) playing the third most important role in overall search term matching for general search results lists and 2nd most important role for MARC records that were viewed by users. Subject fields included all search terms input by a user only a small amount of the time as compared with other top performing fields, with the exception of the 600 (Subject Added Entry – Personal Name). Likewise, only 0.55% of records would not have appeared without the 650 field, compared to the 12.03% loss from the lack of 505 Formatted Contents Note field. There was only one instance of a hyperlinked subject fields being "clicked on" to connect the user to similar works, but subject links suffer from the disadvantage of not appearing in the main search results display (as is evident from Image 1) nor being highly visible in the full record page (see Image 3).

While representing only a small subset of the dynamic URLs generated, the 100 – Main Entry -Personal Name field was the only authorized field that appeared to be used (or "clicked on") directly by patrons. However, the 100 field is also the only authorized field that Encore regularly displayed in search results list, which provided greater access to the functionality of MARC records, even without opening the record to view in full. Users showed evidence of clicking on these headings from the search results list to display works associated with the creator name. Due to this, continued efforts in creating and maintaining personal names will be supported by the unit, with the hope of increasing those efforts if time or opportunity allow.

Conclusion

Data driven decision making is a key component of the CMS unit at Utah State University Libraries. The benefits of having the unit work together to code and analyze how search terms interact with MARC metadata is the opportunity to collectively and concretely reflect on how user searches connected (or did not connect) with records produced or provided by the unit. Reviewing this information allowed the unit to see that efforts needed to be focused on maximizing the number of MARC records displayed in the search results lists. From the results of this study, adding the 505 Formatted Contents Note appeared to be the most likely method for increasing the number of MARC records in results lists. Additionally, due to their visibility in the search results display as well as evidence of direct user interaction with the field, local practice will continue to support creation and maintenance of the 100 – Main Entry - Personal Name field and increase efforts in this direction as time or opportunity provides. However, subject fields may not have the same impact, whether due to system display issues or lack of matches to search terms employed by the user. Given the utility of the 650 Subject Added entry – Topical Term in records directly viewed or accessed by patrons, though, time spent assigning subject terms will remain part of the local practice, even if it is de-emphasized as a broader means of access.

References

1. Xi Niu, Tao Zhang, and Hsin liang Chen, "Study of User Search Activities With Two Discovery Tools at an Academic Library," *International Journal of Human-Computer Interaction* 30, no. 5 (2014): 422–33, https://doi.org/10.1080/10447318.2013.873281.

2. Cory Lown, Tito Sierra, and Josh Boyer, "How Users Search the Library from a Single Search Box," *College and Research Libraries* 74, no. 3 (2013): 227–41, https://doi.org/10.5860/crl-321.

3. Fredrick Kiwuwa Lugya, "Next Generation Catalogues: An Analysis of User Search Strategies and Behavior" (University of Illinois at Urbana-Champaign, 2016).

4. Gunnar Knutson, "Does the Catalog Record Make a Difference? Access Points and Book Use," *College & Research Libraries* 47, no. 5 (1986): 460–69, https://doi.org/10.5860/crl_47_05_460.

5. Tina Gross and Arlene G. Taylor, "What Have We Got to Lose? The Effect of Controlled Vocabulary on Keyword Searching Results," *College and Research Libraries* 66, no. 3 (2005): 212–30, https://doi.org/10.5860/crl.66.3.212.

6 Tina Gross, Arlene G. Taylor, and Daniel N. Joudrey, "Still a Lot to Lose: The Role of Controlled Vocabulary in Keyword Searching," Cataloging & Classification Quarterly 53, no. 1 (January 2015): 1-39, https://doi.org/10.1080/01639374.2014.917447.

7. Cherie Madarash-Hill and J B Hill, "Electronically Enriched Enhancements in Catalog Records: A Use Study of Books Described on Records with URL Enhancements versus Those Without," *Technical Services Quarterly* 23, no. 2 (February 16, 2005): 19–31, https://doi.org/10.1300/J124v23n02_02.

8. Yuji Tosaka and Cathy Weng, "Reexamining Content-Enriched Access: Its Effect on Usage and Discovery," *College & Research Libraries; Vol 72, No 5 (2011): September*, 2011, https://doi.org/10.5860/crl-137.

9. Laura N. Kirkland, "The Relationship of Metadata to Item Circulation," *Cataloging and Classification Quarterly* 51, no. 5 (2013): 510–31, https://doi.org/10.1080/01639374.2012.762963.

10. Karen Smith-Yoshimura et al., *Implications of MARC Tag Usage on Library Metadata Practices*, *Implications of MARC Tag Usage on Library Metadata Practices*, 2010, http://www.wip.oclc.org/content/dam/research/publications/library/2010/2010-06.pdf.

11. Smith-Yoshimura et al.

12. Gross and Taylor, 2005.