

Benchmark Case Studies of Two Digital Library Websites

Melina Zavala, Alex Kinnaman, Jacob Jett, Caroline Szylowicz, Timothy W. Cole

1 Executive Summary

This report presents the results from a pair of case studies conducted as part of the Exploring the benefits for users of Linked Open Data for digitized special collections project. Each case study was produced from a series of interviews with users of digital special collections. The case studies compare the Motley Collection of Theatre & Costume Design¹ (Motley) to the Harvard Theatre Collection² and the Kolb-Proust Archive for Research³ (KPA) to the Bovary Manuscript Archive⁴ respectively. Each of the users was a volunteer and was asked to compare to digital collection websites to one another during the course of completing a series of user tasks which included assessing the overall layout and utility of each digital collection's interface, searching for a specific resource, and characterizing how they might employ the collections in their research.

The primary purpose of the case studies is to establish a rough baseline (benchmark) of observations regarding how users interact with specific digital collection websites and the features and resources available through them. A second round of user interviews is pending that will record how users interact with the Motley and KPA websites after a variety of Linked Open Data improvements have been deployed on them. The sample size, both in terms of number of sites analyzed (2 pairs, so 4 in total) and in terms of numbers of users (5 for each pair, so 10 in total) is modest; establishing the relevance of findings to other sites and/or other classes of users would require further testing.

With regards to the benchmarks, a number of common themes emerged from the case studies which will be used for future comparisons. Among the themes that emerged from both case studies are:

- Users value clear sorting and filtering options on their search results pages;
- Users highly value facility for faceted browsing options;
- Users desire more comprehensive metadata to appear in the short descriptions for each item in the search results lists, especially with regards to:
 - The date the original item was created and
 - The creator of the original item;
- Users desire more prominent collection homepages to:

¹ <http://imagesearchnew.library.illinois.edu/cdm/search/collection/motley>

² <http://hcl.harvard.edu/libraries/houghton/collections/htc/>

³ <http://kolbproust.library.illinois.edu/proust/search>

⁴ <http://flaubert.univ-rouen.fr/bovary/>

- Differentiate the collection from other collections available through the search interface,
- Provide a richer explanation of the collection, and
- Provide clearer search examples.

With regards to the Motley website, users observed that providing the following improvements would be helpful:

- Options for changing number of items on a results page;
- Presenting a side-by-side view of item's image with its metadata;
- Having all advanced search fields available to streamline search process;
- Having thumbnail images of each item on the search results pages;
- Having a "Browse All" option;
- Adding the ability to search within results from a previous search; and
- Flattening the page hierarchies so that it takes fewer clicks to reach a full record of each item

With regards to the KPA website, users observed that providing the following improvements would be helpful:

- Having a more prominent search bar and access to "Advanced Search" options
- Having a summary of the amount of data that is available along with clearer instructions for how to access it.

The next round of user testing will determine which (if any) of these themes the deployment of LOD services and data addresses.

2 Introduction

The value of retrospectively digitizing the contents of library special collections has been well established through the impacts it has had on the manner in which scholars search for, access, and use special collections (Palmer et al. 2009, Tibbo 2003). However, much work remains to understand what more needs to be done to optimize these digital resources for accessibility, findability, and use in both pedagogical and analytical processes (Maron & Pickle 2013).

Linked Open Data (LOD) is among the many recent innovations in Web technologies designed specifically to ensure that data, documents, and metadata describing both are easily found, shared, and used. Multiple library-based research initiatives (see for instance, Marcondes 2016, Coyle et al. 2015, Torre-Bastida et al. 2015, and Allalouf et al. 2015, among others) and community-based incubators like Linked Data for Libraries (LD4L) and Linked Open Data in Libraries, Archives, and Museums (LODLAM) are creating a firm foundation upon which to build additional research efforts.

The Exploring the benefits for users of Linked Open Data for digitized special collections project builds on this foundation by specifically examining:

- Ways to enrich existing digital special collections' metadata with links to external resources,
- Methods for mapping idiosyncratic, legacy metadata into LOD-friendly vocabularies, and
- Enhancements that exploit LOD to improve user interactions and increase visibility of special collections' items.

This research is taking place in the context of three digitized special collections:

- The Motley Collection of Theatre & Costume Design,
- The Portraits of Actors (1720-1920) Collection,⁵ and
- The Kolb-Proust Archive for Research.

This case study describes a series of user interviews that provide a benchmark set of comparisons between two of the above special collections (the Motley Collection of Theatre & Costume Design and the Kolb-Proust Archive for Research) and two comparable digital special collections (The Harvard Digital Theatre Collection and the Bovary Archive respectively).

3 Methods

This case study is based on interviews conducted with scholarly users of digital library services. Each interview adhered to a protocol designed by principal investigator Timothy Cole, project coordinator Ryan Dubnicek, and researcher Peter Organisciak. Two groups (Group A and Group B) of 5 scholarly users were asked to compare a pair of websites. User Group A compared the the Motley Collection of Theatre & Costume Design to comparable collections of theatre related content available through Harvard's Digital Collections portal. User Group B compared the Kolb-Proust Archive for Research to the Bovary Archive. Each individual interview was supplemented by notes taken by a note-taker present at the time of the interview. All participants agreed to an IRB-approved informed consent statement at the beginning of the interview sessions, all of which took place between October and December of 2016 in the University Library's user studies lab.

The case study analysis summarized below applied the experiential knowing method outlined by Stake (1995). Analysis within the two cases is based on naturalistic interpretations of the interviews, supplementary materials, and observations, followed by a categorical analysis of patterns across the cases. The main body of this report is organized into two primary sections: Site Descriptions and Findings. The Site Descriptions section provides a general overview of each of the websites that interviewees were asked to interact with. The Findings are presented in general categories that correspond to core website functionalities that interviewees interacted with. A final section summarizes the key findings of the two case studies.

⁵ <http://imagesearchnew.library.illinois.edu/cdm/search/collection/actors>

4 Site Descriptions

4.1 Motley Collection of Theatre and Costume Design

The Motley Collection of Theatre and Costume Design (Motley), held at the University of Illinois at Urbana-Champaign's Rare Book and Manuscript Library, is a collection of 4,085 items pertaining to the productions in which the Motley Group's designs were used. Motley designed costumes, set designs, and props for over 150 productions in England and the United States from 1932-1976 at nearly 60 different locations. Much of the Motley Group's work was for Shakespearean performances in England, most notably at the Royal Shakespeare Theatre, but the group also branched out into various Broadway and Metropolitan Opera productions.

Items included in this digital collection consist of scanned images of sketches, colored drawings and paintings, the occasional fabric sample, props and cast lists, storyboards, and black and white photographs. The bulk of the items are colored costume designs, often accompanied by a penciled-in note of the actor to play the character, followed by colored set designs. There is also a small collection of black and white photographs of stage settings, and various lists and sketches of props. The original metadata consists of general theatre, play, and performance information, as well as material information about the items' creation.

Motley items can be browsed in full or by theater with browse and search options available on the main page and browse papers. The collection contains a basic search function and an advanced search function. The advanced search can search by date, up to four different metadata fields, and by the exactness of the searched phrase. Results can be sorted by Type, Author/Composer, Subject, Theater, and Performance Title.

4.2 Harvard Digital Collections

The comparison site chosen for Motley was the Harvard Theatre Collection (HTC). Held at the Houghton Library at the Harvard College Library, the HTC is a comprehensive collection of documents relating to theatre history and evolution. There are 33,055 total items that extend beyond theatrical performances and plays and into dance, magic, music, pantomime and puppetry, circuses, and other festivities. HTC contains a much broader selection of item types than either Motley or Portraits, but still contain comparable material for digitized images of actors in their costumes and set designs.

The Harvard Theatre Collection is accessed through the Harvard Library Visual Information Access (VIA) portal, which is its own website and searches every digitized record for items own by Harvard and can be limited to the Harvard Theatre Collection. The search function can manage three search terms in three different metadata fields, by date, and results can be sorted by up to three metadata fields. Searches can be limited by records with digital images and by records with originals at Harvard. In addition to the VIA system, the HTC contains an array of pictures depicting actors in costume and posing in scenes from their current performances, many of which are Shakespearean, which creates a comparable collection and search system to Motley.

4.3 The Kolb-Proust Archive for Research

The Kolb-Proust Archive for Research,⁶ owned by the University of Illinois at Urbana-Champaign's Rare Book and Manuscript Library, is a collection of ca. 40,000 research notes compiled by Professor Philip Kolb (1907-1992) from primary and secondary documents, in the process of transcribing, editing, dating and publishing over 5,000 undated letters of French writer Marcel Proust (1871-1922). The research notes themselves are arranged in several distinct files,⁷ that, when combined, paint panorama of Proust's life and milieu, and of turn-of-the-century French intellectual life.⁸

To date, three files have been digitized, and two of them made available and searchable on the Kolb-Proust website:

- A bibliography of all writings by Proust published during his lifetime and posthumously, and all writings about Proust, i.e. criticism, during his lifetime (ZZZZZZ entries); and
- A detailed chronology of his life and times, compiled by Kolb as a tool to pinpoint the dates of many events, great and small, which were crucial in establishing the date of many letters of Proust (ca. 9,000 entries).

The site contains about 20,000 digitized note cards, some with multiple entries on them and allows users to search the Kolb-Proust archive online. About half of the notecards have not yet been digitized. The different entries are transcriptions of Kolb's notecards marked up according to the TEI guidelines. Kolb's research notes do not conform to the several document types defined and described in TEI. Each file type, however, follows a stable structure and format, so that transcription templates were constructed from TEI elements for various data types.

The chronological entries were enhanced through detailed encoding of content. Titles were sorted in ten different genres (fiction, non-fiction, poetry, music, sculpture, etc.) Individuals, real or fictitious, were identified and given a unique identifier that was attached to each occurrence, making it possible to search for a specific person, regardless of the way in which s/he is cited: first name, last name, nickname, married name, pen name, allusion.

The main page of the website displays a menu that has an individual tab for the different ways the site can be searched. One can search by a keyword, advanced, or freeform search. At the bottom of each tab, there are examples of searches that can be done to show the viewer the proper way to type in information into the boxes. The advanced feature includes sections in which titles, genres, authors, dates, names and limitations can be selected to limit the search. Within the genre selection there are options to select from boxes that say prose non-fiction, prose fiction, poetry, theater, dance, variety show, graphic arts, opera, music and sculpture to limit the search. One can also search by entire text, given on the left hand side, which can search by specific phrases or words. In the freeform tab, one can search using Boolean operators, field names, and parentheses.

Once a search is done, the results are given in index card format that include the date, information on what is being said (who is writing to who if it is a letter, the name of the

⁶ <http://www.library.illinois.edu/kolbp/>

⁷ <http://www.library.illinois.edu/kolbp/tour/cardfiles.html>

⁸ <http://www.library.illinois.edu/kolbp/categories.html>

newspaper, etc.), and the citation of the source. On the left side of the screen, one can select the preferred names that the search term generated, the dates, and the type of result. All of these are selections that can be made to limit the results that are viewed on the page. One can also sort by relevance, title, author, publication date, or reverse date all of the results generated.

4.4 The Bovary Manuscript Archive

Chosen as a similar digital archive in French for comparison to the KPA, The Bovary Manuscript Archive (Atelier Bovary) is a digital archive of scholarly transcriptions of the working manuscripts of Gustave Flaubert (1821-1880) for his novel Madame Bovary. The collection includes 5,000 pages of transcriptions from Flaubert's notebooks. These transcriptions are arranged chapter by chapter, with multiple variants and drafts in each one. The different materials within the website support a study of the entire creative writing process, from start to finish of the writing of the book. Included within the website are supporting materials, such as excerpts from Flaubert's correspondence, an index of proper names, mappings of some of the locations mentioned in the manuscript, censored passages and notes made by the author. The material in this archive is smaller and much more focused than Proust, but still offers comparable content.

There are menus on both the middle of the page and on top of the page, and one can get to the search page from the top menu of the main page. Once there, one can search by words, a word with, phrases or Boolean operators. The specific search can be faceted by the whole corpus, by parts of the novel, plans or scenarios, everything but plans or scenarios, copy notes or edited text. The results given are placed in chunks on the screen with the word or phrase highlighted yellow and information on the location in the manuscript of the part selected.

5 Benchmark Case Study Discussion

Each of the following benchmark case studies are divided into three primary sections which detail general information about the interviewees, themes that emerged during their comparisons as they completed tasks on each of the two web-sites, and an explanation of factors that may have influenced their decisions and opinions. The emergent themes sub-sections are further divided into categories that emerged during the coding of the interviews. These categories are:

- Interface and Layout Impressions
- Search Feature Interactions
- Metadata Reactions
- Homepage Comments
- Relevance and Further Exploration

5.1 The Motley / HTC Benchmark Case Study

5.1.1 Summary of Participant Characteristics

Five participants were involved in the user study comparing Motley to Harvard's VIA. Their backgrounds were in the humanities, primarily in various aspects of British literature, English, and theatre. Collectively the participants were familiar with online databases for locating primary and secondary resources, including ECCO, EEBO, JSTOR, Project Muse, ProQuest, and EBSCO. Their primary source materials generally included rare books and manuscript

collections, academic journals, and bibliographies. Most of the participants had experience in locating and using digital sources, such as digitized versions, but most also were not as familiar with projects where the materials were purely digital. For the purposes of this document the participants will be referred to as “users.”

5.1.2 Emergent Themes from Interactions and Search Tasks

5.1.2.1 *Interface and Layout Impressions*

Users had mixed responses on the general layout and interfaces for both systems, with one user finding both of the interfaces “clean and pretty user-friendly” (MCP05) and another finding them overwhelming because “it was so heavy with text rather than images that it was just a lot to take on” (MCP02). Distinctions between the two websites became more distinguishable once the users had started their search tasks and found themselves interacting with search results pages.

With regards to Motley’s search results pages, most users particularly remarked on how the organization of the search results list was useful to them because they preferred having a flow-through list for a more comprehensive view of the material. One user mentioned that this organization provided a better “sense of the collection” and what the collection might contain (MCP02). Only one user found the organization of the results list confusing and was hindered by grouping by theatre rather than by date or performance title. This user did not find the sorting options for the results list at the top of the page readily apparent.

Similarly, users viewed the organization of search results pages for the HTC’s VIA system positively. On the results list page, users liked that there was an option for changing the number of results on a page, and that the images opened in a new tab when they clicked into an item so they did not have to click backward and could continue forward with their search. However, users also wanted the ability to sort the results chronologically which was not available using the HTC’s VIA system.

Despite reacting positively to how the search results pages were organized, users found the information presented for each of the Motley items to be problematic with one user observing, “I feel like it’s not quite giving me all of the information that I want about the initial results to be able to tell what to click on” (MCP05). Users most wanted to see the date and the affiliated company on the initial results list in addition to what information was already present so that they had a better sense of what items they needed. The absence of this information resulted in the users having to click through more items to determine which were relevant to their task. One user also would have preferred larger thumbnails, but did not see the option to change the layout at the top of the results list. The HTC results were viewed more positively than Motley results as the users collectively liked that the date appeared but, all users commented on wanting the theatre (also absent in HTC search results) listed in the results list metadata. After taking the time to click on several of the search results, most users felt that there was enough information provided on the pages for individual items to decide if an item was relevant to their task for both collection sites.

With regards to collection browsing activities, users seemed content with the Browse All option for both sites, but many commented on wanting the option to browse by specific facets, e.g. theatre, actor, director, and company. For Motley, the option to sort or narrow a search by theatre

is located both on the left side panel and above the search list, alongside several other browsing options, but these were not found by the users. There was a divided reaction in finding the “next” button on the results page, which half found easily and half did not. Ultimately browsing was easier in Motley but sorting options were more difficult to find, if they were found at all, and all users found both system’s use of thumbnails useful.

Many people were confused on the location of the search bar and how to access the Advanced Search option, with one user commenting that it was too small and “scrunched up towards the top” and therefore difficult to find. There were also comments on the color and size of the search bar and associated text. Conversely, the landing page the users began on for Harvard was comprised primarily of the search function so there was no comparison for finding the main search function.

Once users had clicked into an item in Motley, the primary comment was wanting the image and the metadata to be side by side rather than the metadata being below the image. One user commented on the utility to have “some of this information [on] Material/Techniques, watercolor, and pencil, you know, to have it side by side [for better connecting the information with the item]” (MCP02). This was reinforced by users’ positive interactions with Harvard’s side-by-side layout for metadata records. They also liked the image manipulation options in Harvard better than the zoom feature in Motley, as well as the access to high-resolution images, with one user noting that “they would be useful for including as image examples to papers” (MCP01).

Only two users found the “save to favorites” option in both systems and commented that they would use it for their research if they were using the digital collection. Combined with the difficulties in locating browsing options it would seem that not all site features are evident to users as they interact with digital collection websites.

5.1.2.2 Search Feature Interactions

Compared to the HTC’s VIA system, all of the users had various difficulties in efficiently using Motley’s search system. As previously discussed, there was initial confusion in finding the search bar and the text that said “Advanced Search” on the Motley site. In one case the user was not able to find the Advanced Search until after they had run several queries in the basic search function. Another user was uncertain where to start their search because it “looks like the search for the whole website” (MCP01). As previously mentioned, the users started on the search page rather than a collection description page when using the HTC’s VIA system so there were no issues locating the search feature for the HTC.

Users’ ability to search and sort by date in the Motley Collection was one of the largest problems they struggled with when using the Motley site. A few users did not find the “Search by date” option at all and, one commented on how unfortunate this lack of an option was to their search. At the end of one interview the user admitted “Oh look it does say ‘search by date’ right there. I missed that completely” (MCP05). Those who did locate that search option were then confused about how it functioned. For example, the following quotation is a user’s description of their reaction to the search by date option: “It’s a little confusing to me that now that I’ve added in a date, you know, a ten-year window, and I’ve looked at several of these costumes now and I know

they're all from 1954 so I'm wondering that, you know, why do I have it set from 1950 to 1960 and now I can search and I don't get anything" (MCP05). The majority of users found the HTC's VIA search system for Harvard mildly confusing but much more manageable than Motley's. Users often were looking for an advanced search option but the HTC's VIA is a single advanced search system with multiple options rather than presenting users with separate basic and advanced search features.

Most of the users also suggested improvements while using the fields in the Advanced Search in Motley. Users liked that they could search by author, but wanted to also have the ability to sort by director. In general, the users wanted more options but could not provide specific examples as to what other sorting methods they wanted. A user commented that "while the advanced search was a good, it was like, there, I mean it has what you basically need, I felt like I'm used to seeing a little bit more options" (MCP02). Another user suggested having multiple fields open initially, rather than having the user manually "Add another field" because "it looks like Motley has those capabilities, I just didn't click through enough to get to them. So yeah, having those multiple fields available immediate I think makes it easier" (MCP05). Comparatively, the reaction to the HTC's VIA system's search function was mixed. Though users liked the option to enter multiple search terms they were unsure of how to narrow their searches, specifically to search by photographs or "artistic renderings" (MCP05, MCP04). Users spent several minutes of their time on this task experimenting with the search filters, specifically the Name, Subject, and Title filters. This was overwhelming to some, but others wanted more filter options. In the results section, one user also mentioned that "having the search section beyond this page with the access to add these Boolean phrases, to search for these images, to create the, um, the limitations of what you're looking for was extremely helpful" (MCP04).

5.1.2.3 Metadata Description Reactions

Users reacted positively to the amount and organization of the metadata provided in the Motley collection. When asked what additional information they would like to have for a collection like this, all users indicated that the metadata presented was everything they might want for research purposes. Most of the critical comments on metadata were in how it was displayed, as previously noted the users all preferred the HTC's side-by-side presentation. For the HTC, most of the metadata provided in the collection was useful to the users. However, they noted the lack of description for the items and one user said that they were "trying to figure out who this is, if this is in fact a photograph of Juliet. Right now it's saying who the photographer is but I'm not getting any sense" (MCP02) and had to rely on their knowledge of the play rather than the provided metadata. One user was also interested in having a description of the costumes, such as the color and fabric type (MCP01). Overall, the users found the metadata to be similar and mostly comprehensive in both systems.

5.1.2.4 Homepage Comments

Finally, Motley's homepage caused some confusion for users as they found it disorienting to start a search task on a page they did not locate themselves. Though there is a description of the collection on this page that explains the scope and content, half did not read it because the information is in small print. There were also mixed reactions to the image carousel on the right side of the page. One user said that it was a distraction (MCP05), and another that "if you are specifically searching for something, it's not terribly useful as a, like, informational page"

(MCP01), whereas another user described the carousel as providing “a library feeling, [because] you see the new books and you walk in and it makes you feel like there’s like an interactive, that this itself is like a library space” (MCP02). Comments on the HTC’s homepage are the same as the layout of the search feature described above as they are the same. It is interesting to note that there were no comments on the size of print on the HTC’s landing page as there was for Motley’s.

5.1.2.5 Relevance and Further Exploration

Towards the end of each interview, the users were asked if they could think of a search task that was relevant to their own research and to describe what they would do next if time permitted. Most of the users stated that the collection would be useful to their research and expressed interest in searching for other plays and designs with one user stating, “I’ll use this in the future, I did not know it existed” (MCP03). Further exploration users described included faceting the advanced search more and experimenting with different methods of searching, opening and saving images, and general exploration of the collection. One mentioned that it was not relevant to their research because of the timeframe of the collection (MCP04), but others also said that “Moving forward as a director I would love this. I do some directing as well to complement my academic work, and so being able to pull on past designs would be very beneficial” (MCP03). Similar comments were made regarding Harvard’s collection, especially because of its wider scope, but most users were interested in sketches and drawings rather than photographs.

5.1.3 Limitations

There are three factors that might have influenced the user’s understanding and reaction to the search task. First, the language of the search task was found confusing by several of the interviewees and the narrative describing the tasks needed to be repeated several times, and users often needed to read the task from the script in order to understand it. They were also unclear about the collections as part of a larger repository and often tried to change the settings in which the study was set, for example trying to search all of the University of Illinois’ digital collections instead of just within Motley, or trying to search for journal articles rather than staying within Harvard’s Theatre Collection.

Second, their lack of familiarity with the collections caused uncertainty in the search tasks, specifically in why the collections did not contain certain materials or materials from a certain time periods. The users were not supposed to be familiar with either of the collections, but their confusion created obstacles in how they searched. While exploring Motley for the same material the user found in the HTC, one user commented, “What did I do wrong? You’ve got to have pictures of *The Tempest*. Hm. I don’t believe you” (MCP01). The Motley Group never designed anything for a production of *The Tempest*, but the user assumed the collection would have a comprehensive collection of Shakespearean plays.

Finally, several users also commented on the anxiety caused by being watched perform the search tasks. Though the interviewers stated that the study was to test the system and not the users, as per the interview script, many felt it was necessary to complete the task successfully despite any workflow disruptions. This may not be an accurate representation of their search habits as they assumed the study would not ask them to find materials that did not exist, and

therefore they proceeded as though the task was a challenge with an end goal rather than a description of their search process.

5.2 The KPA / Bovary Benchmark Case Study

5.1.1 Summary of Participant Characteristics

Five participants were involved in the user study that compared Kolb-Proust Archive to the Bovary site. Their backgrounds were all in the humanities, for the most part in English literature or in a field dealing with the French language. As a whole, the participants were familiar with online databases for locating primary and secondary resources, such as JSTOR and other similar databases. Some participants relied more on online databases while others more on books to obtain both primary and secondary information. Most of the participants had experience in locating and using digital sources, such as digitized versions, but a few also were not as familiar with projects where the materials were purely digital. For the purposes of this document the participants will be referred to as “users.”

5.1.2 Emergent Themes from Interactions and Search Tasks

5.1.2.1 *Interface and Layout Impressions*

As a whole, the KPA interface was the one that was preferred by the users. Many of the users mentioned the site had a clear and neat look, with one even observing, “I liked that it felt like a traditional reference-you had the date right on the top, quick summary of where it’s from, and then the summary or snippet from the actual thing” (KP01). The website’s minimalism was for the most part a quality that made it the better site for the users, since the Bovary site had at times a debilitating page to the eye, with too much information to process at once. Many users had difficulty finding the main search page in the Bovary site, creating a challenging beginning to their searches. The KPA website gave easy access to the information for the users, with one user even saying that they preferred that one “for the research part, how to enter terms you are looking for” (KP05). The accessibility of the search bars made a big difference in how the users reacted to the sites.

Despite this general preference for the arrangement of the KPA site over the Bovary site, two of the five users noted how the search results were more visually accessible on the Bovary site. In particular, highlighting of the search keyword on the search results page was particularly helpful. It helped the users to hone in on the particulars of where exactly their search term was located in the context of each result and helped them to further investigate the source. One user did note of the Bovary site overwhelming amount of textual information though that “the words are put next to one another and I don’t understand the relationship” (KP05).

Some of the available content in the Bovary site, such as the thematic divisions given, was interesting to some of the users, and one user commented that, “It feels like there’s someone more in the back and directing me to what’s interesting... the thematic divisions here seem kind of interesting [and] for a student that would be kind of easy to get into” (KP03).

5.1.2.2 *Search Feature Interactions*

There were many comments on the search features for the KPA, but not many about those of the Bovary site. The KPA search function caused many issues for the users in the process of their

given task. The very particular way that one has to type in the dates in the search bar led many to confusion. There was also a lot of confusion over how the numbers were supposed to be entered in the Proust advanced search. For most participants, getting past how to begin to properly use the advanced search was the most difficult part of the experience. Many of the users did not realize till much later after starting the process of searching that there were examples of how to do correctly type in dates. One observed, “I didn’t see the examples first [of] how to conduct research... I found the example that was given here maybe was not enough if you are not familiar with a system like that” (KP05). Many had similar thoughts about the example given and would have liked more clearer and visible instructions.

The faceting options the users would choose also caused problems for some because it would give them a “Term limit exceeded” message if they had entered it wrong. This would then make some try to go back and specify the facet even more, which in most instances did not help and would give them the same result over again. One user thought that the error message meant that there was lot more results that they may have reached if they had searched properly. Another user attributed this to mean that the KPA website had a lot more to offer, saying “This one [KPA], I thought had a lot more because when I kept on trying to go through the advanced search and just find the certain timeline it kept on coming back as there’s too many results” (KP01). In contrast to this, users experienced less trouble in getting a result with the Bovary search features because for the most part, results easily appeared on the screen. They did, however, experience problems finding the Bovary’s search interface, sometimes having to make two attempts to locate it. One positive comment that did come up about the Bovary search features was that there were “so many more options for tabs and things [compared to the KPA]” (KP02).

Overall, the keyword search for the KPA worked better than the advanced search. As one user put it, “It was a little odd that the keyword search ended up working better than the advanced search. I started trying to get more specific but it rejected that and when I just searched his name, I actually got results” (KP01). Many of the users tried very hard to get the KPA’s advanced search to work, but in the end did not manage to do so. One user gave an interesting idea of why the system tends to behave this way. They said that it was because of the initial audience it tried to serve; “Kolb-Proust is great in terms of information, but it seems to be targeting researchers who know how to handle a search engine that doesn’t offer, that doesn’t really take you by the hand” (KP03). Most users also agreed that in terms of information, the KPA site provided much more than the Bovary site. However, even after recognizing the abundance of information on the KPA site, users still had problems accessing any of it, lead that same user to say, “in its simplicity, the Kolb-Proust archive is actually a little cold. It can give you everything, which is great, but it’s a little hard to use” (KP03).

5.1.2.3 Metadata Description Reactions

One of the shared comments that both websites obtained was that more links should be added to the material that is already available. The users wanted to have more information directly linked to what they saw on the screen. Many commented on the desire to be able to see the original source to determine how it fit in with the results in the KPA. One typical observation was, “It’d be cool if you could see where it is in the context of the letter in addition to just these little snippets” (KP04). Better organization of the data available was also an issue in both systems. In

the KPA's case, it became an issue of what was more important for one user: "I think sometimes it gives you a lot of information, but it puts all the information in the same level...That's the thing about Kolb—it's interesting once you manage to find your way to the information you're looking for, it has a lot of stuff going on, lots of entries, but it can be a little bit difficult to [search]- that information is not as categorized" (KP03). This user saw the abundance of information available but desired a more narrow approach at how it was presented. The lack of a good search term would leave many of the users to not find what they were looking for, or for them to miss pertinent information.

Labeling was an issue for two users in both systems; trying to decipher what the results were and how they fit in with the search term was a recurring issue. One user had an idea for a change that could be done to the KPA's site, saying, "It'd be nice to have a way to easily sort them from these are the newspaper things, these are the letters, these are the letters about him" (KP04). A lot of the time the users spent was trying to figure out what the results were and how they fit in with the search term they were given. Many users also overlooked the fact that the Bovary site was a manuscript archive and believed that the results they saw were other material embedded within the data available, creating further issues. This led many users to being confused; both in regards to what they were looking at, as well as from what time period the resource was from or about. In particular, one user observed, "It's kind of hard to decide here, just in the snippets if it's exactly like from Madame Bovary or if it's from something else... Ok, but I can't find [dates] really labeled anywhere" (KP01). The Bovary site did not have the information very clearly dated like the KPA's items, and only after some digging did one user find a chart that could possibly aid in the categorization of the material they were viewing.

5.1.2.4 Homepage Comments

The KPA's homepage was generally thought of as the clearer one of the two website homepages. One user commented that it would have been nice to have a page that explained more about what the KPA collection held and how to actually search it. In point of fact, a page like this currently exists, but the interviews did not start on the information page of the project and instead began on the KPA's search page. One user commented, "...that it's very minimal. That's what struck me" (KP05). This was also both a plus and a negative for that user, who further commented that in its formality, the KPA had become "old fashioned" (KP05). Overall, users found the search page layout for the KPA to be very simple and easy to access, while they found that the search bar for the Bovary site was much harder to find.

Many users took a long time just finding the search bar in the front page of the Bovary site. The homepage of the Bovary site has a rather large menu section which spans the top and both sides of the screen with a large picture dominating the middle of the page. The excessive links confused some of the users and the search bar was hidden among all of the links. In regards to the search link being on top of the screen, one user said, "Just in terms of the look of it, I somehow felt that this was a subsection of the site and my brain just went whatever is on top is probably the site above, and this is the site for Bovary, and whatever is above is not" (KP03). A disconnect between what was part of the website or not made this user take unusually longer in finding the search link, though users that did read through all of the links on the page usually went there at some point.

5.1.2.5 Relevance and Further Exploration

During the interview, each user was asked to come up with a search task that was relevant to their own research and to describe what they would do next if time permitted. Most of the users preferred the KPA and showed an interest in getting more historical documents from it. One user also said that it “would be interesting is looking at the relationship between authors” (KP05). A lot of names are in the archive, and creating a relationships among them interested this user. Another user expressed interest in analyzing specific expressions used within the documents. Two users expressed interest in using historical documents for analyze. One user said, “Probably whatever research I’d be doing would be historical context. So I’d be searching for historical names” (KP02). Most of the users expressed interest in analyzing the texts through a historical analysis of some type.

5.1.3 Limitations

There are two main limitations on the task flow that most of the users faced. The first was the wording of the research task was confusing to several of the users. Explanation of the tasks needed to be repeated several times throughout the interview. Clarity for who was who (Kolb vs. Proust vs. Bovary vs. Flaubert) needed to be reestablished throughout as well. The names were at times confused from switching from one archive to the other, especially since the tasks were to look for the same author in both.

Not all of the tasks were perfectly repeatable on both websites. Specifically, within the KPA part of the interview, a date limit was set but not during the Bovary section. Many users carried on the belief from the KPA part that they were supposed to find mentions of Proust during specific time frames in the Bovary section, but this was not the case. One of the follow up questions did ask if a determination of what the date of the results was, but it was not part of the original task given to the user.

The second barrier that the users faced was the difficulty of the advanced search feature in the KPA. Most users instantly went to the advanced search when they were told their task and then spent much of their time on the KPA task trying to get results from it. Most of the users would move on after not succeeding and then would go back to the keyword search, which would then tend to display results. A lot of the users commented that the amount of information in the KPA was immense, yet the lack of the advanced search feature’s usability made it very difficult to access the specific information that they were looking for.

6 Emergent Themes and Conclusions

6.1 The Motley / HTC Benchmark Case Study

Ultimately, users were split on overall preferences with two preferring Motley, two preferring the HTC, and one who stated they “[did not] feel comfortable saying only because in one of them I was searching for photographs and one of them I was searching for renderings” (MCP04). The following is a list of aspects from each system the users experienced that they responded positively to, commented on the utility of, or mentioned wanting the option to have in regards to their search task. These were:

- Clear layout

- Prominent search bar and access to Advanced Search option
- Clear sorting/filtering options on results list page
- Option to change number of items on a results page
- Faceted browsing
- Side-by-side view of image with metadata in a record
- Having all advanced search fields available to streamline search process
- More comprehensive metadata on results list
- More prominent collection homepage to differentiate it from other collections
 - Readily available information on the scope of the collection
 - Direct link back to the collection's landing page

There were several aspects of Motley in particular that users responded positively to, including:

- List-view for results page
- Thumbnails on results page
- Browse All option
- Ability to search within results from a previous search
- Fewer clicks – Motley contains few levels to move through to access a full record
- Comprehensive information in a full record

6.2 The KPA / Bovary Benchmark Case Study

In the end, three of the users preferred the KPA, one the Bovary site and one was unsure of which of the two was better. All of the ones who selected the KPA as their preferred site said that in the end, it was a bit easier to access than the Bovary site. The Bovary site was for most, more difficult to understand what the information given was actually about, which created difficulties in determining whether their task was actually complete. The following is a list of aspects from each system the users experienced that they responded positively to, commented on the utility of, or mentioned wanting the option to have in regards to their search task for Proust.

- Clear layout
 - Prominent search bar and access to Advanced Search option
 - Clear sorting/filtering options on results list page
- Faceted browsing
- Having the date and source clearly displayed on individual results
- More prominent collection homepage with various information parts
 - Explain what the collection is and how to search it
 - Clearer search examples.
- The amount of data that seemed available if site was searched properly.

7 References

- Allalouf, M., Artyomov, A., and Mendelsson, D. (2015). LODLI: Linked open data back-end systems for libraries interoperability. Paper presented at 7th International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Management (IC3K 2015), Lisbon, Portugal, 12-14 November 2015.

- Coyle, K., Silvello, G., and Tammaro, A. M. (2015). Comparing methodologies: Linked open data and digital libraries. Paper presented at 3rd AIUCD Annual Conference on Humanities and Their Methods in Digital Ecosystem (AIUCD 2014), Bologna, Italy, 18-19 September 2015.
- Marcondes, C. H. (2016). Interoperability between digital collections in archives, libraries and museums: Potentialities of linked open data technologies. *Perspectivas em Ciencia da Informacao* 21(2), 61-83.
- Maron, N. L. and Pickle, S. (2013). Appraising our Digital Investment: Sustainability of Digitized Special Collections in ARL Libraries, Association of Research Libraries and Ithaka S+R. Retrieved from:
<http://sr.ithaka.org/researchpublications/appraising-our-digital-investment>
- Palmer, C. L., Tefteau, L. C., and Pirmann, C. M. (2009). Scholarly Information Practices in the Online Environment: Themes from the Literature and Implications for Library Service Development. Dublin, OH: OCLC Research and Programs. Retrieved from:
<http://www.oclc.org/content/dam/research/publications/library/2009/2009-02.pdf>
- Stake, R. E. (1995). *The Art of Case Study Research*. New York: SAGE Publishing.
- Tibbo, H. (2003). Primarily history in America: How U.S. historians search for primary materials at the dawn of the digital age. *The American Archivist* 66(1), 9–50.
- Torre-Bastida, A.-I., González-Rodríguez, M., and Villar-Rodríguez, E. (2015). Linked open data (LOD) and its implementation in libraries: Initiatives and technologies. *Profesional de la Informacion* 24(2), 113-20.