

Examining the Richness of Browsing in Digital Libraries & Netflix

by

Jared Rixter

May 2021

Presented to the

Division of Science, Information Arts, and Technologies


University of Baltimore

In Partial Fulfillment


Of the Requirements for the Degree of

Master of Science

Approved by:

DocuSigned by:  
 7/2/2021  
A10D371CA4EC43F...

Kathryn Summers, Thesis Advisor

DocuSigned by:  
 7/2/2021  
F35B199DB3CE408...

Stephen Kiel, Committee Member

## Abstract

Search and browsing are essential to information retrieval in libraries. Digital library users will engage in these behaviors to find books they need in any digital library's interface. Search is supported through the search bar, which is an incredibly powerful tool for any website that holds vast amounts of information. The operation of search engines and the usability of search interfaces have improved greatly over the last two decades. In contrast, the mechanisms and tools for online browsing have not been as well developed. This paper seeks to understand user browsing behaviors on digital libraries and whether their experience suffered if they could not use the searching tool. This research was based heavily around Dana McKay's guidelines for digital browsing, and Marcia Bates' browsing behaviors. Sixteen participants completed a survey regarding their reading habits and library usage. They then conducted user testing with several digital libraries and Netflix. Netflix was chosen because its users rely heavily on browsing to find content, and Netflix supports browsing quite successfully. The digital library tasks required users to browse for books that fit a certain criteria and books that they would personally read. The task for Netflix required participants to find something to watch that they had never seen before. The results suggest that although digital libraries still have room for improvement, the gap between their browsing experiences and websites like Netflix may not be as far as initially anticipated. This paper compares the strengths and benefits of each digital library to illuminate some best practices for current and future digital repositories.

## Acknowledgements

I would like to thank my mom first and foremost. You have always been my biggest supporter. Words cannot describe how much I love and appreciate all you have done to help me get to this point. I would also like to thank Greg, Kathryn, and Sujan. Your guidance though my academic career has been truly invaluable.

Finally, I need to thank Dad and Auntie Pat. When I started this graduate school journey, I never would have imagined that I would finish without being able to celebrate with you. At every step in my life, I could always depend on the two of you for anything. I hope I have made you proud. I love you both.

List of Tables

Table of Contents

List of Tables ..... ii

Chapter 1: Introduction ..... 1

    Research questions ..... 3

Chapter 2: Literature..... 4

    Libraries..... 4

    Metadata..... 5

    Book Cover Analysis ..... 8

    Shelf Design..... 9

    Facets and Categories ..... 10

    User Behavior ..... 11

    Design Guidelines for Online Browsing..... 13

Chapter 3: Methodology ..... 19

    Methods..... 19

    Participants ..... 20

    Time ..... 21

    Tools and Tasks ..... 21

Chapter 4: Results..... 26

    Survey Results ..... 26

    User Patterns ..... 32

Chapter 5: Discussion..... 38

    Implications..... 42

Chapter 6: Conclusion ..... 46

    Limitations ..... 46

    Recommendations for future study..... 48

References ..... 50

Appendix A: User testing materials ..... 54

Appendix B: User Book Selections ..... 56

Appendix C: Researched libraries ..... 58

List of Tables

List of Tables

Table 1. *Participant demographics* ..... 20

Table 2. *Assigned Digital Libraries* ..... 23

Table 3. *Survey questions 3- 8* ..... 26

Table 4. *Survey questions continued* ..... 29

Table 5. *Selected Books for first digital library* ..... 56

Table 6. *Selected books for second digital library* ..... 56

## Chapter 1: Introduction

### Chapter 1: Introduction

Both search and browsing are essential behaviors that humans use to retrieve information, particularly when the information is stored in a large repository like a library. So why should we look to improve browsing? To find the answer we will first examine search. With the advent of internet search engines, people participate in numerous search queries daily. Search itself is ingrained in computer data retrieval because graphical screen space is valuable real estate for most website interfaces. However, search by itself does not give users a full understanding of a digital library's collection.

Whitelaw helps us understand the limitations of search through a museum analogy: “The visitor enters the building, whose collection she hasn’t encountered before. Instead of expansive exhibition halls, however, she finds a small, drab lobby with an attendant at a desk. The attendant asks the visitor to write her query about what she’d like to see on a small slip of paper. The visitor invents a query, and the attendant disappears for a moment before returning with a line of artworks on trolleys, which are paraded - ten at a time - through the lobby” (Whitelaw, 2012, p. 3).

The current information retrieval models usually only supply content based on keywords that the users themselves provided. If users are unfamiliar with the most relevant keywords, they may only receive a subset of relevant content; this effectively soft locks information behind an invisible wall that users do not know exists. Whitelaw’s example reflects a model that some libraries use called closed stack or shelves. Closed shelves are used more often in European libraries than they are in most American libraries, which use open shelves (Rovelstad, 1976). European libraries went through several evolutions shifting between open or closed based on their collection sizes, number of patrons, and patron needs (Rovelstad, 1976). Rovelstad sees open shelf libraries as systems that were influenced by the democratic concept of equal opportunity (Rovelstad, 1976). Rovelstad concludes her paper with the idea that combined closed and open shelf arrangements can help libraries meet their service goals (Rovelstad, 1976). I would suggest that if the closed shelves reflect modern search, then open shelves could be

## Chapter 1: Introduction

likened to browsing; and like Rovelstad, I suggest that a system that balances both would be ideal for modern digital libraries.

Since all digital libraries rely heavily on the search bar, this "soft lock" in front of information affects nearly every digital library on the web. This problem is even more relevant in the era of Covid-19. Due to this pandemic, libraries across the nation have needed to restrict users from using a library in person. Digital libraries are thus, even more frequently than before, the only way for users to access a collection. While digital libraries provide their patrons with access to literary content, users across the nation turned to another massive repository of content to occupy their time during quarantine: video streaming websites. Streaming sites always have content for users to explore, but on streaming sites like Netflix, the search bar is used far less often than while users are navigating digital libraries.

If search is one side of the coin that is information retrieval, then browsing is the other. Ellis states that "the concept of browsing has several different connotations, a variety of activities have been associated with it, and several different typologies of browsing have been put forward" (Ellis, 1989, p. 187). Ellis observed social scientists browse through contents pages of journals, checking periodicals held by the library, or simply browsing through the shelves (Ellis, 1989). He noted that the prerequisite for the activity to be effective was that there should be at least some collocation of like material (Ellis, 1989). The grouping and organization of similar materials was more important than the form of the material (Ellis, 1989). So browsing is a behavior that becomes more relevant when related items are grouped together, and exploring those relations can supersede the value of any of the single books in that area. In physical libraries this is what made shelves such a valuable tool for browsing (McKay 2018). This is also why the transition away from shelves has made browsing a new challenge to implement for digital libraries. Browsing can also serve the purpose of familiarizing users with the sources and material of an area (Ellis, 1989). There are two aspects to this activity, familiarization and differentiation (Ellis, 1989). Familiarization allows users to become aware of materials and dive deeper into research; whereas differentiation occurs when users develop a knowledge and appreciation of the differences in available material (Ellis, 1989).

These aspects help to identify what browsing is, and also identify some of the benefits that are lost when modern digital libraries do not support equivalent browsing behaviors.

## Chapter 1: Introduction

Browsing also allows for serendipity, which is the ability to unintentionally find new information. Waugh states that savvy library users used shelves as a part of their information seeking primarily on the understanding that they were likely to find information that they had not actively searched for (Waugh, 2017). Bell broaches this topic with an example about Marla Spivak. She is an insect researcher who made discoveries related to beehive diseases that helped prevent colony collapse (Bell, 2014). During a TED talk she revealed that her fascination with insects began with a book about bees that she found serendipitously at her local library (Bell, 2014). These are the types of interactions that allow browsing to complement search as the other side of information retrieval. This paper attempts to explore the current availability of serendipitous discovery through browsing as currently implemented in several digital libraries.

Thus, this paper will explore browsing between digital libraries and the streaming service Netflix. The literature review will discuss what scholars think about the current systems provided by digital libraries for browsing. The literature review will also discuss metadata, shelf design, book covers, browsing models, and several aspects of user browsing habits. The research methodology for this study will be a qualitative, exploratory experiment that is designed to illuminate how users navigate the architecture of three digital library websites without being reliant on keyword search, with some comparison with how they browse a popular streaming site.

### **Research questions**

This research revolves around three questions:

1. Will users still be able to search effectively on modern digital libraries without the search bar? What browsing behaviors will they use?
2. Will the tested digital libraries show any tool that has the potential to become as powerful as the search bar?
3. How do users' experiences in finding content change when browsing or searching on a digital streaming site such as Netflix? What do users enjoy about browsing on Netflix?



## Chapter 2: Literature

### Chapter 2: Literature

This paper will begin by examining how libraries operate, how digital libraries came about, and what scholars currently understand about both search and browsing. Browsing in particular may be one of the most difficult human behaviors to pin down, since there are so many strategies to analyze, and scholars' understanding of the browsing process has evolved gradually. Our current search systems have evolved from the library catalog systems that were initially designed to complement physical libraries.

#### Libraries

A library is a construct containing books, periodicals, films, and recorded music, and other forms of media. Why have we moved from physical to digital libraries to house so much of this information? We have had record-keeping for almost as long as we have had written language. Modern libraries have evolved from the archival collections of monasteries or other religious institutions (Foskett, et al. 2017). These early institutions were focused on preserving manuscripts and other important historical documents (Foskett, et al. 2017). Universities and colleges also created libraries in order to support their research and preserve its results (Foskett, et al. 2017). Monarchs and other elites also collected books to create their own personal libraries (Foskett, et al. 2017). Only groups with sufficient infrastructure or wealth could amass large collections of books, since books were expensive to reproduce by transcription. Later, with the invention of the printing press, came the ability to mass-produce books far more cheaply than through transcribing by hand, increasing the number of produced, reprinted, and copied books (Foskett, et al. 2017). From this point, the publishing industry grew exponentially as different kinds of books were written, published, and sold (Foskett, et al. 2017). In the United States, the Library of Congress was established to house documents, support and conduct research, and house the copyright office (Foskett, et al. 2017). Other national libraries were established in the same manner. In the nineteenth century public libraries became more widely accepted. Modern libraries have been collecting books ever since, and they have collected increasingly varied forms of media for many decades (Foskett, et al. 2017).

How do libraries organize all of this information? The answers are classifications and metadata—originally collected in lists or books, then eventually moving to card catalogs. North

## Chapter 2: Literature

American libraries commonly use either the Dewey decimal system or the Library of Congress classification system (Foskett, et al. 2017). Other classification systems include the universal decimal system, the bliss system, the colon system, and the Marxist system (Foskett, et al. 2017). Of these, the colon and Marxist systems are notable because they are used outside of the west, primarily in India and China respectively (Foskett, et al. 2017). The card catalog system became the primary tool for libraries to organize their collections before computers because it was easy to use and to expand. The transition to computers began in the 1970s as library collections continued to grow (Elrod. 1972). Libraries moved away from card catalogs because of several issues.

- **Size:** Library catalogs grow as collections grow which increases the difficulty of maintaining card catalogs. The physical space taken up by the cabinets that stored the cards was also an issue (Elrod. 1972).
- **Editorial condition:** The library card system was subject to human error since there were many steps involved in the process.
- **Physical condition:** Physical library cards are subject to wear as they are used and are costly to maintain.
- **Preservation:** Card catalogs were vulnerable to mutilation, damage, or loss of entries.
- **Staff considerations:** maintaining the card catalogs was an intricate procedure that requires staff to spend time performing repetitive or unappealing tasks (Elrod. 1972).

Card catalogs were a foundation for creating digital catalogs that are still used in libraries today (Elrod. 1972). These systems were extremely relevant for physical libraries and shelf design; however, in addition to the disadvantages listed above, card catalogs were unable to fully support the dynamic changes and flexible organizations that characterize digital libraries today.

### **Metadata**

The information that libraries collect and maintain about their holdings is called metadata. Metadata is simply information about information. For example, metadata can be information such as author, publisher, genre, language, page number, volume number, publishing

## Chapter 2: Literature

location, publishing date, awards won, semantic categories, related works, and other aspects that can be deemed relevant. Metadata is one of the most powerful tools for all library users and administrators. McKay and Cunningham discuss how metadata supports browsing in one of her papers on the Greenstone classifier system: “documents are classified at collection time according to their metadata, and browsing structures are pre-built ready for loading” (McKay & Cunningham, 2003, p. 331).

Both physical and digital libraries depend on metadata for both search and browsing, and as McKay explains, metadata systems shape the browsing structures that can be implemented readily. When physical libraries began to shift from card catalogs to digital catalogs during the 1970s and 1980s, their metadata systems mostly stayed the same. Digital libraries upload books, and the metadata keywords assigned to the books shape how every user will find that entity after that. As time passes, the metadata used by digital libraries has moved beyond the metadata contained in the old card catalogs.

When developing metadata, librarians need to consider quality, data checking, error correction, and the ongoing refinement of processes for error prevention in regards to metadata. Nichols has conducted research on metadata tools for institutional repositories (Nichols 2008). The tools he studied were the Metadata Analysis tool (MAT) and the Kiwi Research Information Service (KRIS) (Nichols 2008). These tools were developed to help pinpoint specific metadata errors and generate summary statistics. There are differences between what each tool specifically does for repository administrators: “One gives a holistic view of metadata while the other looks for specific problems; one seeks patterns in the data values while the other checks that those values conform to metadata standards” (Nichols 2008, p. 1). Nichols states that there can be circumstances where tradeoffs need to be made in order to reach other desirable goals for an institutional repository (Nichols 2008). For example, a digital library can temporarily sacrifice the quality of the metadata in order to improve the speed of its coverage across a collection (Nichols 2008). The quality of metadata in this case refers to the depth of metadata provided for each content item as well as how well the metadata aligns with national guidelines.

Selecting metadata for library content is an enormous responsibility placed upon librarians, who cannot possibly anticipate every current or future context of a content item, since discourse around many subjects can change in a very short time. It can be difficult to know what

## Chapter 2: Literature

metadata to prioritize. Users may place different values on different kinds of metadata based on their browsing behaviors. If a user is conducting an author search, metadata about that author and their work may be more relevant to them than metadata regarding publishing date or location.

McKay and Cunningham's research with the Greenstone library software was intended to observe the capabilities of a new browsing system (McKay & Cunningham, 2003). The capabilities were allowing users to combine search and browse, allowing users to choose the metadata by which they browse, allowing users to browse by more than one type of metadata at a time, and restricting the amount of information on any one screen (McKay & Cunningham, 2003). Users in their research appreciated these capabilities in the new system, suggesting that metadata in other digital libraries should provide users with these same capabilities.

Metadata tools to support browsing can be anything from lists of collections to interactive graphic displays. However, choosing how to display sets of metadata can be a challenge. Nichols has explored visualizations to aid metadata harvesting software (Nichols 2008). He gives developers of tools for reviewing metadata quality the advice that although visualizations were considered useful, repository managers expressed strong preferences for textual and statistical approaches. This suggests that both forms of presentation should be available (Nichols 2008, p. 14). What tools are effective can also change based on users' level of prior experience. Library users in academia may be more familiar with or understand different levels of organization than non-academic users. Frias-Martinez and Chen studied the satisfaction of users with digital library interfaces (Frias-Martinez & Chen, 2005). They classified users into four groups based on their own perception of their expertise: never used the system, novice, medium, and expert (Frias-Martinez & Chen, 2005). The type of collection and the type of audience served by the digital library also affect user experience (Frias-Martinez & Chen, 2005). Digital library administrators should consider what browsing tools would be best paired with both the collection and the users. For example, some digital libraries use interactive maps to illustrate where certain publications originate from or where certain publications are being checked out the most. (These examples can be seen in at the world digital library and the Howard digital library that are referenced in Appendix C.) The full value of new browsing tools cannot be known or evaluated until we see what users think when they see these library interfaces and what strategies they employ while using them.

## Chapter 2: Literature

### **Book Cover Analysis**

Although metadata is the primary tool that libraries use to organize books into their interfaces, users also pay attention to visual aspects of content such as book covers. There are other physical affordances inherent to physical books such as size and the cover or spine, which are attributes that relate to the feel of the book as well as its appearance. These traits help us focus on or dismiss books from our selection process in ways that may not reflect the actual relevance of the information.

So what is happening when users look at the arrangement of books in these interfaces in a digital library? Gudinaivicius and Suminas explore user preferences when they look at the cover of a book online. Their study observed correlations between attraction to warm and cool colors between men and women, in addition to studying their reading habits (Gudinaivicius & Suminas, 2017). Gudinaivicius and Suminas were only interested in how visual design affected choices, so their research attempted to obscure book titles as much as possible in order to not influence a user into making assumptions about the content of the book. They specifically tried to prevent users from gleaning data clues such as "known author," "seen in advertisement," and "friend recommendation" when making their selection decisions (Gudinaivicius & Suminas, 2017). For the purposes of this thesis, Gudinaivicius and Suminas' study suggests that users' high value on these excluded factors confirms that these elements are important pieces of information that a digital library should convey to users.

Gudinaivicius and Suminas' work confirmed that the variables that impact the readers' interest in a new book release include the reputation of the author or publisher, the attractiveness of the cover, the degree to which the cover represents the content of the book, and the type of book. Thus, book covers convey both direct and indirect information (Gudinaivicius & Suminas, 2017). Although this particular study focused on book preferences as affected by gender, this research is valuable because it highlighted key information elements that should be supported for browsing. The user interfaces that each digital library will provide alongside its "galleries" will need to support retrieving these kinds of information in order to be successful.

Most of the information in Gudinaivicius and Suminas' variables is successfully conveyed by streaming video sites. For example, when users mouse over a thumbnail on Netflix, a trailer is played alongside an overview. Text, images, words, and narrative development are all invoked

## Chapter 2: Literature

(although Netflix recently stopped providing release date in its mouseover information). Most digital libraries could likely incorporate some form of information package that similarly provides users with multiple quick, convenient options to learn about each item in a collection.

Digital libraries do not control book titles and thus do not have any control of how well book titles can attract users while in their collections. However, digital libraries can control their interfaces and practices. Bell discusses ways libraries foster what he calls collisions (Bell, 2014). Collisions are moments where users randomly encounter a book that can lead to them briefly scanning, moderately reading, or having a life changing experience with the book they find (Bell, 2014). Bell discusses how university libraries create potential collisions by placing book collections and shelves in unexpected or unique locations on campus. For instance, some university libraries placed shelves of graphic novels in computer labs, or a new books rack near a café on campus (Bell, 2014). Creative solutions like this show what libraries are already doing to improve the serendipity of physical spaces. These innovations can be used as inspiration for digital libraries to further support a serendipitous browsing experience, even without physical shelves.

### **Shelf Design**

Shelf design has had considerable influence on digital libraries' interfaces because library shelves are where most browsing occurs in physical libraries. McKay has explored this. She outlines several key points in shelf design: display a large range of books for browsing, enable multiple points of access to the collection, support zooming capability, provide seamless transitions, support place marking, provide visual alternatives during triage, and enable easy access to detailed book information (McKay, 2017). Library shelves have been described as a "near-perfect browsing engine" (Kleiner et al., 2013). However, considering physical shelves a perfect browsing engine ignores both the limitations of shelves in supporting browsing and the fact that people always adjust their behavior to make the best of imperfect systems. For example, physical shelves can have books that are checked out or misplaced (McKay et al., 2017). Children primarily browse through shelves that are at eye level (McKay et al., 2017). Some physical affordances are inherent to physical books like size and the cover or spine, which are both attributes that relate to the feel of the book as well as its appearance. These traits help us focus on or dismiss books from our selection process in ways that may not reflect the actual

## Chapter 2: Literature

relevance of the information. As digital libraries attempt to support browsing, they should be aware of both the strengths and weaknesses connected to browsing physical shelves.

Many digital libraries arrange their collections into galleries on every possible page, which may be an attempt to capitalize on the search behaviors supported by physical shelves. Galleried views, rich search results, categories, and browsing interfaces are usefully versatile, since they can change dynamically based on user needs and on their devices. Such structures have already been widely adopted in other forms of digital media and live streaming services like Netflix. The enhancements to gallery views exemplified by Netflix, discussed later in this thesis, could potentially be capitalized on by digital libraries to onboard users more quickly.

### **Facets and Categories**

Both physical and digital libraries attempt to support access to their content by organizing it into categories. These categories are often organized alphabetically or in some other enumerative scheme. The general idea is that if every subject is on a list, all users have to do is check the list to find whatever they need. The problem with this approach is that subjects and categories have many nuances and change over time. Broad categories like math and science cannot guarantee users with results that are relevant to their query. Information categories will continue to be refined and expanded, which will add new topics and perspectives that need to be considered when responding to queries or supporting browsing.

Facets were developed for libraries when S.R. Ranganathan published his *Colon Classification* in 1928. Facets were adapted for online search by Marti Hearst and her team in the early 2000s. Metadata about facets can supplement more rigid categorical structures for organizing information. Faceted classification is where instead of classifying individual items in an enumerative scheme, items are described with rich detail that allows them to be organized into dynamically generated combinations in order to filter content in readily specified ways. A faceted classification system supports more flexible search and browsing. For example, a digital library that focuses on military history could be organized by wars: WWI, WWII, the Korean War, the Vietnam War, etc., in a well-developed enumerative categorization system. Using facets, a digital library focused on recipes could be organized by styles of cooking like French or Korean, or it could be organized by the ingredients used in each recipe. Facets have been

## Chapter 2: Literature

aggressively deployed in ecommerce platforms, but are sometimes underutilized in modern digital libraries.

Facets are used in both search and browsing. While searching, facets allow users to refine what they are looking for. While browsing, facets allow users to take alternative paths to the results that are presented to them.

Yeh and Liu explicitly compared the value of faceted searching and browsing interfaces compared to standard enumerative classifications. In their work, they discovered that a faceted browser did increase their users' success in finding relevant information as well as their satisfaction while using a digital library. They used the FLAMENCO faceted search interface. This was an open-source search interface framework that utilized faceted navigation for browsing and searching (Yeh et al., 2011). The website of the FLAMENCO interface claims that “it uses category metadata to guide the user toward possible choices, and to organize the results of keyword searches. FLAMENCO also uses hierarchical faceted metadata in a manner that allows users to both refine and expand the current query. This use of metadata is integrated with free-text search, allowing the user to follow links, then add search terms, then follow more links without interrupting the interaction flow” (“Flamenco search,” n.d.).

### **User Behavior**

One of the earliest descriptions of human information search behavior came from Marcia Bates. Bates talks about the context of berry-picking search (Bates, 1989). In this model, users operate in two spaces while searching: the universe of knowledge and the universe of interest. The universe of knowledge is where users will likely develop their keywords. They will match what they know about subjects to the tools they have available in the library interface and to the results their queries provide. The universe of interest is more involved with the results of the user's queries. After developing a keyword from what they know, users will begin searching or browsing through the interface until they have found some document or information that matches that keyword, or found a new keyword to try (Bates, 1989). During this process, both the universe of knowledge and the universe of interest are likely to evolve as users refine their results until they decide that what they think they are looking for and the results from the interface are aligned.



## Chapter 2: Literature

Bates discusses six search strategies that users employ (Bates, 1989). Footnote chasing is where users check footnotes, articles of interest, or other reference lists. Citation searching involves finding a citation and then finding out who cites it by looking it up in a citation index. In a Journal Run, the user identifies a central journal in a topic area, then locates the run of volumes of the journal and searches straight through relevant volume years (Bates, 1989). Area scanning involves browsing the materials that are physically collocated with materials located earlier in a search. Subject searches in bibliographies and indexing (A & I services) are where users search through bibliographies for a subject to pursue elsewhere. Author searching is when a user searches an author's name to find other work on the same topic (Bates, 1989). Not all digital libraries support these search strategies. In particular, some digital libraries do not provide access to footnote or citation tools.

Bates discusses four levels of search activities: moves, tactics, stratagems, and strategies (Bates, 1990). "A move is the lowest unit of search activities, like entering a query term or selecting a certain document. Tactics are described as a combination of several moves, like the selection of a broader search term or breaking down complex search queries into sub-problems" (Carevic et al. 2018, p. 1). Bates defines a stratagem as a "complex number of moves and/or tactics that generally involves both a particular identified information search domain anticipated to be productive by the searcher and a mode of tackling the particular file organization of that domain (Bates, 1990, p. 6)."

Each of the subsequent levels incorporates the prior levels but with increasing complexity. For this research, moves were considered as actions that can be undertaken in one click, tactics as any point where a user chooses or changes a search target, which then leads to a new tree/set of moves, and stratagems were considered as the culmination of tactics and moves that users develop while browsing during each task. Strategies from Bates' levels were not included in this study because they are a mixture of all the lower levels and subsequently cover the "entire information-seeking process." This level was too broad to focus on in a short, guided session.

These search activity levels help define users' actions. Every step that users take while they browse a digital library can be translated into one of the above. In addition, many of these behaviors can also be observed on Netflix and other streaming video websites. Categorizing

## Chapter 2: Literature

these behaviors will allow a comparison between the usage of each search activity level between the digital libraries and Netflix. Carevic later makes notes from another study that investigated a faceted interface. In his second study, there are three defined areas of interest: the search results page, facets for filtering results, and the queries leading to results (Carevic et al. 2017). Carevic's study revealed that users spent an average of 50 seconds analyzing the search results per task, followed by choosing among facets at 25 seconds, and building queries at 6 seconds (Carevic et al. 2017). Carevic's results reflect the importance of each area with participants in this thesis project, except for queries since keyword search was excluded during this research. Although this thesis focused on browsing behaviors, Carevic's work helps to categorize these browsing behaviors using stratagems. This structure of moves, tactics, and stratagems will also be used to compare and contrast browsing patterns on digital libraries versus browsing patterns on Netflix.

### **Design Guidelines for Online Browsing**

McKay has also identified a set of interface requirements that she sees as essential for online searching and browsing:

1. Search is simply the act of looking for something or otherwise seeking. Search in this regard should be considered as looking for a singular book, whereas browsing has no singular target. Search can still be involved in the browsing process (McKay, 2018).

2. Interleaved Search and Browse means that "information seekers must be able to interleave search and browse, with each interaction seeding further exploration in the other mode" (McKay, 2018, p. 349). This means that users need to be able to switch between the two behaviors without being impeded. Browsing should generate results and facets or categories to search for, and search results should be presented in a way that encourages browsing. Ideally, this behavioral loop will help users churn through what information is useful (McKay, 2018).

3. Search-Free Browsing means browsing without search (not meant to conflict with feature #1). This guideline is in place to ensure that browsing can still occur without relying on a searching mechanism. Some users will not use catalogs or search bars in order to find books. Regardless of why users do not search, they still need to be able to browse effectively (McKay, 2018).

## Chapter 2: Literature

4. Book Set Filtering is a feature that allows users to group books based on their preferences. This guideline is akin to users grouping books together based on bibliographic or non-bibliographic features in a physical library. Digital libraries will generally focus on bibliographic features because these are all easily translated into metadata. These features include observations like date of publication, publisher, author, genre, and other types of metadata (McKay, 2018).

5. Place marking is equivalent to actions like "running one's fingers along shelves as they look" in a physical library (McKay 2018, p. 350). In a digital library, place-marking can be represented by any breadcrumb navigation tool that allows users to know at what stage they are in during a browsing process (McKay, 2018).

6. Mark-as-Read is a feature which allows users to avoid unknowingly backtracking into books they have already discovered. Being able to return deliberately to books they have previously found is a good feature. However, accidentally stumbling into the same book a second time can lead to frustration or confusion (McKay, 2018).

7. Rapid- Seamless Zooming specifies that users need to be able to move in between a state where they have an overview of a topic down to a particular book (McKay, 2018).

8. Overview enables users to get an overall sense of the information space. This is particularly helpful to users with loosely defined needs. It is important to help such users refine their needs in order to begin a more in-depth browsing process. Overviews also support users' agency, by allowing them to make a mindful choice about any subsequent area of focus. The overview needs to be both "useful and neutral." This means that the overview should not obstruct or confine a user's browsing process (McKay, 2018).

9. Ability to Jump within a hierarchy means that users should be able to move easily between different levels within their query (McKay, 2018, p. 350).

10. Display of a (Very) Large Number of Books is a particular challenge in digital libraries. McKay notes that in other studies, users "glanced through hundreds of books while navigating to a target or search area, or choosing a target or search area (McKay, 2018)."

## Chapter 2: Literature

11. Rearranging Shelves means changing what books are displayed to users dynamically. Instead of physical shelves with limited space, digital libraries have results pages, galleries, interactive shelves, digital shelves, and other categories that display book covers. All of these tools need to be able to change what books are displayed. Rearranging shelves allows digital libraries to optimize screen space usage and present the most relevant options to users during information retrieval (McKay, 2018).

12. Visible Alternatives is when readers can “examine one candidate book while retaining a visual sense of alternative options.” Essentially, users should not be in a position where they do not have readily available options to continue browsing (McKay, 2018, p. 351).

13. Convenience in a digital library is less about the actual interface and more about accessibility and ubiquity. Digital libraries need to be convenient in ways that make them available for the widest audience possible. Digital libraries should avoid systems that require them to rely only on limited access points or obscure devices (McKay, 2018).

These are the most concrete interaction design guidelines regarding browsing digital libraries. The research of this paper will explore the success of several digital libraries in meeting the third guideline, “Search-free Browsing.”

McKay also discusses a variety of current browsing interfaces, including search result lists, Amazon recommenders, slider-based recommenders, bohemian bookshelves, next-on-the-shelf, and blended virtual reality. Search results and Amazon recommender systems are the most common interfaces for modern digital libraries.

The selected digital libraries for this research were Barnes & Noble, Book Depository.com, the Open Library, and Project Gutenberg. The former two are commercial digital libraries, and the latter two are nonprofit organizations.

Barnes & Noble’s is a high-volume commercial website, with nearly 25 million visitors per day (SimilarWeb, 6/29/21). This website embodies guidelines eight, ten and eleven. Barnes & Noble provides categories in the header of the website. This header then provides an overview of content at every stage of browsing. Unless users have selected a single book, Barnes & Noble fills as much screen space as possible with book covers, banners, carousel categories, and

## Chapter 2: Literature

advertisements for books. Users can choose whether query results are displayed in a grid or a list, and how many are shown are on a page. Books can be displayed alphabetically or not. The website also provides filters that can change how users receive their results.

Bookdepository.com is also a commercial library that uses both an Amazon recommender system and search results list. This library has almost 9.5 million visitors per day (SimilarWeb, 6/29/21), and uses categories and subcategories to address almost every McKay guideline. Like Barnes & Noble, Book Depository uses its header to provide overview functionality while browsing. The library provides rapid seamless zooming via facets and a dynamic set of subcategories relevant to the particular set of search results that are present on search results pages and book pages. When a user views an individual book, the library provides them with links to related categories in case the user wants to pursue a related, non-identical search (an information behavior identified by Marcia Bates). Search results can also be filtered by categories, and these category filters are easy to add or subtract. Both types of filters (facets and categories) also allow each search results page to rapidly change the displayed content as user queries evolve. With an economic incentive to make good use of its screen space, the Book Depository optimizes screen space similarly to Barnes & Noble, thus adhering to McKay's guideline ten. The one feature that Bookdepository.com does not support is book set filtering. There is no tool for users to group selections based on their observations.

Project Gutenberg is a non-commercial site with about 5.6 million visitors per day (SimilarWeb, 6/29/21). This site has features that align with guidelines ten and twelve. For guideline ten, Project Gutenberg does can show several pages worth of books in a search results page. This library also has several top one hundred lists that can display books that are organized by site-wide popularity or by author as measured over the last seven or thirty days. Guideline twelve is shown on pages for individual books. On each page there are multiple subject categories as well as a "readers also downloaded" link that provides recommended books that other users downloaded while browsing. This library only uses the search result list as interface, instead of a mixed system like the commercial libraries.

The Open Library is a noncommercial site with the smallest number of daily visitors—about 2.79 million (SimilarWeb, 6/29/21). This site has a unique way to tackle guideline four, book set filtering. This digital library allows users to create lists and view lists created by other

## Chapter 2: Literature

users. The closest alternative tools observed were “recently viewed” categories, wish lists, and carts from commercial digital libraries, although the latter may interrupt a query. The Open Library is developed and maintained by grants and by volunteers, and relies heavily upon the search result list browsing interface. Using search to access content is the original online browsing interface for information collections (McKay, 2018, p. 351). Users are familiar with search systems, and search systems can be easy to implement and maintain for digital libraries, although the best implementations of search take more time and effort.

In contrast to non-commercial digital libraries, commercial digital libraries like Barnes & Noble tend to follow the Amazon recommender systems, largely because these systems are effective at not only selling books but collecting user data in order to determine what books to suggest to users.

The remaining types of interfaces were not explored in my research but are noteworthy. The next-on-shelf system is a graphical one to one correspondence to how books would be displayed or organized on a physical library shelf (McKay, 2018). Slider-based recommenders ask their readers to indicate their preferences on a range using sliders (McKay, 2018). The next-on-shelf system is only currently available at [knowledge.exlibrisgroup.com](http://knowledge.exlibrisgroup.com). The Bohemian bookshelf is only available at the University of Calgary; however, the tool can be viewed in use online at [alicethudt.de/BohemianBookshelf/](http://alicethudt.de/BohemianBookshelf/). Other unique systems like these are available and will be discussed in the limitations section of this paper. Blended virtual reality systems are the only systems I could not find examples of on the web that were suitable for user testing over Zoom.

The most popular tools that appeared on most sites were search results lists and Amazon recommender systems. These systems are commonly seen together in digital libraries. Because some tools are ubiquitous across digital libraries, users become familiar with them, making them easier to use. Search results and Amazon recommenders are both popular examples of McKay’s guidelines.

The Bohemian bookshelf that McKay recommends is potentially another good candidate for ubiquitous use across libraries; however, the difficulty of implementing and maintaining this or similar systems in libraries with larger collections seems to have limited its prevalence. The

## Chapter 2: Literature

reason the Bohemian bookshelf stands out from the other interfaces is because it focuses on data visualization to support browsing. This interface has visualization tools that arrange a collection in multiple ways, such as by book cover color, keywords, alphabetically, number of pages, and publication date. At the beginning of this research, I believed it might be valuable to find a visualization tool that could rival the search bar in order to provide users with useful content. The Bohemian bookshelf has an extremely varied set of tools to support browsing. However, the usefulness of some of these tools to actual users has not yet been established. The keyword and alphabetical visualizations may be less usable in this format than the search results list, which has the benefit of being familiar to users. The number of pages of a book may be useful to would-be readers in a few contexts, but it certainly does not give users any information about the book's subject matter. The publication date may or may not be relevant in some types of search. For example, the Open Library, a digital library tested in this study, also contained a publishing date tool embedded into the search page, and this option was not utilized by any of the study participants.

To review, digital libraries have several tools at their disposal to support search and browsing. Since their inception, libraries have relied on metadata to organize their perpetually increasing collections. Metadata—especially if supplemented by book covers!—can be used to create categories and identify facets that will be valuable if they can anticipate user needs. Once McKay's tools have been implemented, librarians need to observe what strategies users employ while navigating their digital libraries. Finally, creative alternatives to these standard tools should also be explored and then evaluated.

## Chapter 3: Methodology

### Methods

For this study, I initially hypothesized that the most fruitful strategies to translate to modern digital libraries would be journal run, area scanning, subject searches, and author searching. Although the other search forms are extremely valuable, untrained or casual users are unlikely to spend extended amounts of time going through footnotes or citations (Bates, 1989). The first step in this study was to explore the digital libraries that were accessible online. Thirty-five different digital libraries were examined. Each was evaluated based on their collections and their browsing mechanics. The site's search filters were also recorded because they are an indicator of the metadata used to organize each digital library. Every digital library was checked for the ability to retrieve a specific book without the use of keyword search (McKay, 2018). The results of this exploration can be seen in Appendix C.

Second, four digital libraries were selected for observational user testing. Two commercial and two non-commercial sites were selected. The sites had varying levels of daily traffic, but all four sites had large book collections.

After a pilot test of the observational user testing protocol, it was decided to include a browsing task on Netflix for comparison. The reason this comparison was added was in order to understand if participants' browsing behaviors would change if they were browsing for something other than books, or if their browsing behaviors would work better on Netflix than on the digital library sites—thus suggesting that expectations for online browsing may be well-adapted for video streaming sites but less well-adapted to digital libraries. Ellis stated that all that was needed to support online browsing was to make the metadata held about objects browsable (Ellis, 1989). Netflix is a comparable repository of objects that is organized by metadata, accessed through browsing, and has more than 2.4 billion daily visits (SimilarWeb, 6/29/21).

For Netflix, a brand-new account was created without any recommendations that would potentially skew what categories or shows Netflix would present to users first. The next step was to create a Google forms questionnaire that would ask users about their browsing habits, along with a protocol for user testing. The results of this testing are provided and discussed in chapters 4 and 5.



### Chapter 3: Methodology

#### Participants

Table 1.  
*Participant demographics*

	Gender	Age	Education level
Participant 1	m	26	Some College
Participant 2	m	26	Bachelor's
Participant 3	m	39	Graduate
Participant 4	m	30	Bachelor's
Participant 5	f	64	Bachelor's
Participant 6	m	31	Bachelor's
Participant 7	m	58	Graduate
Participant 8	m	22	Some College
Participant 9	m	27	High school
Participant 10	f	23	Bachelor's
Participant 11	f	26	Bachelor's
Participant 12	m	26	High school
Participant 13	f	41	Graduate
Participant 14	f	28	Bachelor's
Participant 15	f	25	Bachelor's
Participant 16	f	25	Graduate

Sixteen adult participants participated in the user testing. All 16 of these participants also completed the surveys, either before or after the test session. Children were excluded because many digital libraries are made for college-level research instead of grade school education and to simplify review by the university's Institution Review Board. The only adult candidates that were excluded were users that were especially familiar with libraries, such as librarians, or users that would be especially familiar with research methods or interaction design principles, such as user experience designers. Of the 16 participants, nine were male and seven were female. Participant ages ranged from 22 to 64. Educational attainment varied from high school graduates to recipients of graduate degrees. In contrast, research done by McKay and Carevic has focused on college educated participants (Carevic et al. 2017; McKay, et al. 2004). Participants were recruited by advertising to friends, family, and acquaintances about the study on social media sites like Facebook and Instagram.

## Chapter 3: Methodology

### **Time**

The first stage of this project, research into digital libraries, began in August of 2020. User testing took place in February of 2021. Surveys were administered at the same time users participated in testing.

### **Tools and Tasks**

Test sessions were held via Zoom as video conferences. Participants were directed to each of the four online digital libraries that were selected for testing. The digital libraries that were tested are as follows: Barnes & Noble, BookDepository.com, Project Gutenberg, and The Open Library. The split between commercial digital libraries and non-commercial was intentional. This approach adds to new knowledge because it will compare several different digital libraries currently available online.

All of the tested repositories shared several qualities that made them comparable for the user testing. Each of the digital libraries was selected based on the ability to retrieve a book during the pilot tests, and adherence to more than one McKay guideline. How each digital library addresses McKay's guidelines is discussed in the discussion section. Each library had abundant metadata that was available while searching or browsing. All of the digital libraries tested used search results lists; the commercial libraries used these alongside Amazon recommender systems (McKay, 2018). Netflix's interface is largely based around galleried carousels but also uses recommendations and lists of search results. The McKay guidelines can also be applied to Netflix's browsable interface.

Barnes & Noble has a wide variety of categories and sub-categories. Since this is a well-funded commercial site that is widely used, it has a vast collection of books and books that are currently popular. Barnes & Noble provides users with an overview of the book, product details about the publisher and seller, information about the author, and an excerpt from the selected book. User-submitted data about the book is also included at the bottom of the page, including reviews and ratings. Finally, users are provided with books that they have recently viewed at the bottom of the page.

Book Depository is also a commercial site, and also allows users to browse by category. It offers top categories, more categories, top authors, bestselling series, and books by language.

### Chapter 3: Methodology

When viewing a single book, the Book Depository provides users with the book's associated categories at the top of the page, to make it easier to find related books. In addition, the site provides Amazon-style recommendations and bestsellers in related subjects below the book information. This page also shows background information about the author and provides product ratings.

Project Gutenberg is a nonprofit that offers free e-books. This site offers bookshelves where topics are presented alphabetically. Books in different languages are also noted on the bookshelves. Under the search and browse tab, users can view the top 100 e-books over the last one, seven, or thirty days. When viewing a singular book, each user can see similar books by related subject or based on what readers also downloaded.

The Open Library, another nonprofit, allows users to create lists and view lists created by other users. There is also an option to be taken to a page for a random book. Below the top subjects, there are more specific categories below, such as classic books, books we love, recently returned, romance, kids, and more. The Open library allows users to browse by subject on a carousel on the front page. There is also a list of subjects that are arranged alphabetically.

Netflix's content was largely arranged into rows of carousel categories, each with its own unique title. Netflix also has a header menu with the options for home, TV shows, movies, new and popular, and my list. My list was ignored for this study because the Netflix account was deliberately unaltered by preferences so as not to influence users' decisions. When any of the site's category header buttons is clicked, the content on the page will change. Selecting home will change each set of categories to Netflix's usual assortment of seemingly random collections. The categories are organized into creative titles like "binge-worthy," "get in on the action," and "because you watched," in addition to more standard categories like comedy, horror, or documentaries. Selecting TV shows or Movies will change all of the carousel categories to hold only that type of content. Furthermore, shows or movies can be organized by genre with a new option that appears when either TV shows or Movies has been clicked. These sub-categories are similar to the categories that support browsing in other digital libraries, but they are handled differently. Clicking on New and popular limits the content on the main page to shows organized in categories like New on Netflix, Coming This Week, Top 10 in the US Today, and Coming Next Week. Each carousel category houses video thumbnails that, when hovered over, start to

### Chapter 3: Methodology

play trailers for each show. When users click on shows, they are given an overview of the show, a list of the cast, a list of the show's genres, and several words to describe the show. Users can also choose to see episodes with descriptions for each episode. Underneath episodes is a “more like this” section that provides recommendations of similar shows.

Table 2.  
*Assigned Digital Libraries*

	Digital Library 1	Digital library 2
Participant 1	B&N	Open Library
Participant 2	Book Depository	Project Gutenberg
Participant 3	Book Depository	Open Library
Participant 4	B & N	Project Gutenberg
Participant 5	Book Depository	Open Library
Participant 6	Project Gutenberg	Open Library
Participant 7	Project Gutenberg	Book Depository
Participant 8	Open Library	Project Gutenberg
Participant 9	Book Depository	Project Gutenberg
Participant 10	B & N	Open Library
Participant 11	Project Gutenberg	Book Depository
Participant 12	Book Depository	Open Library
Participant 13	B & N	Book Depository
Participant 14	B & N	Open Library
Participant 15	B & N	Project Gutenberg
Participant 16	Book Depository	Open Library

Table 2 shows the digital library sites used by each participant. During test sessions, each participant performed the same two tasks on two of the previously mentioned digital libraries. The two libraries assigned to each participant followed a Latin square order. For the first task, users were given a scenario in which they were a teaching assistant who needed to find two books about groups of people who changed the world. A group of people was the chosen topic in order to reduce the chance of users selecting the first biography of a world leader they see on the front page. This task was intended to simulate users coming to the digital library with an

### Chapter 3: Methodology

intended need or a goal to satisfy. Users were prohibited from using any form of keyword search. However, any other tool provided by the digital library was acceptable for use.

The second task had the same constraints as the first task; however, participants were asked to browse for books that they would choose to read in their own time. This task was meant to simulate users coming back to or browsing further in the digital library out of curiosity. Participants performed these tasks on two different digital libraries, with no time limit. The last task that was given to participants was for the digital streaming site Netflix. For this task, users were asked to find two shows they were interested in watching that they had not seen before, without using the search bar. Because Netflix is known to support browsing well (more movies are selected through browsing than through direct search), it was useful to compare and contrast user behavior between the different digital libraries and Netflix.

Users were prohibited from using the search bar in order to focus their behavior on browsing. This also directly focused the tasks on McKay's third guideline of supporting search free browsing (McKay, 2018). Participants were asked to speak their train of thought aloud to help the researcher understand what techniques users were employing at different stages of information retrieval. Think-aloud protocols are often used in qualitative research (Van den Haak, 2004). During each task I observed their activities and noted what site tools they used, and browsing behaviors from Bates' list (Bates, 1989). Analyzing participants this way is reflective of Carevic's methodology for user testing (Carevic et al. 2017). I required this data in order to shed light on why users had positive experiences with certain tasks, on certain digital libraries, or with specific tools, as well as why they disliked or failed in particular interactions.

The survey that accompanied the user testing session was created with fifteen questions. Several questions collected demographic data irrelevant to the research. Other questions asked participants about their past use of digital libraries. This reflects similar surveys conducted by Carevic (Carevic et al. 2017). The fifteen survey questions and the reasoning for asking each question were as follows:

- Q1: What is your first name? This information was collected only for my data organization purposes.
- Q2: Year of Birth?

### Chapter 3: Methodology

- Q3: When was the last time you borrowed a book from a physical library?
- Q4: When was the last time you borrowed a book from a digital library?
- Q5: When was the last time you purchased an E-book or an audiobook?
- Q6: Do you prefer to buy books or to borrow them?
- Q7: If you prefer to buy books, what websites do you purchase from?
- Q8: Please tell me why you use this site or sites.
- Q9: Do you prefer to read physical books or E-books?
- Q10: Why are either of these your preference?
- Q11: What electronic devices do you use to read (Select all that apply)?
- Q12: What genres of books do you prefer to read?
- Q13: Do you prefer to browse in person in a physical library or online through a digital library?
- Q14: What is your highest completed level of education?
- Q15: If you are employed, does your position require you to read books regularly?
- Q16: How many books do you read in an average year?

## Chapter 4: Results

**Chapter 4: Results**

This section of the paper will discuss the data submitted by users in answer to the survey and the observed user patterns from the user testing sessions. All sixteen participants were able to complete each of the tasks. Users discussed the difficulty of refraining from using the search bar at first but were quickly able to find the tools they needed on each website to find books. One factor that was not anticipated that may have influenced the results was that user testing was conducted during February, which is also black history month. Several users decided to choose African American civil rights groups or other prominent black figures as the target group for their task. However, several of the digital libraries featured African American texts on their homepage for the same reason, making these texts much easier to find.

**Survey Results**

Q2: Year of birth: This question was asked to gauge participants' familiarity with library organization conventions. The average age of participants was 32, with the ages ranging between 22 and 64. A majority of the participants tested were in their late twenties.

Q14: Education level: Half of the participants had completed up to a bachelor's degree. A quarter of participants had completed a graduate degree, and the last quarter was split between some college and a high school diploma. The two participants whose highest degree was a high school diploma were also military veterans.

Q16: Average number of books read in one year: Six participants responded that they read less than five books per year. Six participants also responded that they read between five and ten books per year. The remaining participants answered that they read between ten and twenty books per year.

Table 3.  
*Survey questions 3- 8*

	Q.3 Borrowed from physical library	Q.4 Borrowed from digital library	Q.5 Purchased an e-book or Audiobook	Q.6 Buy vs borrow	Q. 7 Preferred website to purchase from	Q.8 Why this site?
Participant 1	Greater than five years	Greater than five years	Greater than five years	Buy	Amazon	Usually have the books I'm looking for
Participant 2	Greater than five years	Greater than five years	Less than six months	Buy	Amazon	Everybody puts their books on Amazon so it's a wide selection
Participant 3	About two years	Less than six months	Less than six months	Buy	Amazon	Convenience

## Chapter 4: Results

Participant 4	About a year	Never	About three years	Buy	Amazon, eBay	Ease of selection, option of new or used books based on books ordered
Participant 5	Greater than five years	Less than six months	About a year	Buy	Amazon	It's easy and has a lot in stock
Participant 6	Greater than five years	Never	Less than six months	Buy	Amazon	Supports my Kindle Paperwhite
Participant 7	About a year	Greater than five years	Greater than five years	Borrow	Amazon	Convenience
Participant 8	Less than six months	Less than six months	Less than six months	Buy	Kindle	I use Libby, Rbdigital, Kindle
Participant 9	Greater than five years	Less than six months	Less than six months	Borrow	School website i.e. GSU.edu	To buy books and things for the class
Participant 10	About two years	Never	About three years	Borrow	I love borrowing and buying. I may be an unusual subject here, though, because really all books I read are non-digital. I support the local bookstore around the corner from my house, called "Unabridged Bookstore." I share hard covers/paperbacks w/neighbors and friends, and purchase a lot of my books. Articles that I read online are Vanity Fair, and NYT.	I know this sounds strange - but I am a trainer in Health Information Technology, teaching / training clinical teams on how to use the electronic medical record. I'm in Community Health. I also pull data from the EMR and work with teams on quality improvement projects. It is because I am starting at a couple of monitors all day, I just prefer the physical books
Participant 11	About two years	About three years	Greater than five years	Borrow		
Participant 12	Greater than five years	Less than six months	Never	Buy	Amazon	it's easy
Participant 13	About two years	Never	About three years	Buy	Amazon	Ease and price
Participant 14	Less than six months	Less than six months	Less than six months	Buy	Amazon, Barnes & Noble	Convenience, Selection
Participant 15	Greater than five years	Less than six months	Less than six months	Buy	Amazon	Quick, easy, can get an e-book or physical
Participant 16	Greater than five years	Never	Less than six months	Buy	Audible	More accessible

This table shows participants borrowing habits and how recently they purchased an e-book. Most participants had not borrowed from a physical library in at least two years. This could indicate that participants did not borrow from physical libraries regularly and that the pandemic likely did not affect their borrowing habits. Participants preferred to buy their books overall and Amazon was the primary website they used for book purchasing. Only one of the participants that said that they preferred borrowing books claimed that they also bought books from Amazon; however it had been more than five years since this participant had purchased a book from Amazon. This participant said that he preferred to buy books from his local physical bookstore. Therefore, this participant may be less familiar with digital library tools than some of the other participants.



## Chapter 4: Results

Q3: Date of last book borrowed from physical library: This question was asked to gauge participants' familiarity with library organization conventions. Seventy-five percent of participants reported that they had borrowed a book from a physical library up to two or more years ago. The remaining 4 participants reported that they had borrowed books either six months or one year prior.

Q4: Date of last book borrowed from digital library: This information was intended to understand users' familiarity with digital libraries. For this question, five participants answered never. Seven participants answered less than six months. Three participants responded greater than five years. One participant answered about three years.

Q5: Date of last E-book or audio book purchase: Project Gutenberg, one digital library that was tested, was focused on E-Books. This question was to learn if users regularly used E-books and what patterns they may have developed regarding them. Fifty percent of participants claimed that they had purchased an E-book or audiobook in the last six months. So at least half of the users were recently familiar with a commercial digital library.

Q6: Prefer to buy or borrow books: This question was intended to complement aspects of other questions in the survey. The survey revealed that most users prefer to purchase books rather than borrow books.

Q7: Where digital books are purchased: This question was intended to learn if participants had used any of the commercial digital libraries that I had chosen for my user testing. It was also intended to learn if there were common websites that users frequented that I had not included in my initial research. The most popular site that participants recorded on the survey to buy books from was Amazon.com. Amazon was avoided as a candidate for the research because users' familiarity would likely influence their browsing behavior. Amazon is also a much larger retail website that no longer focuses on books, so it was also disqualified in order to avoid distracting users with advertisements for products unrelated to books and to the assigned tasks.

Q8: Tell us why you use these websites to purchase E-books: This question was designed to elicit qualitative information about participants' answer to question seven. Participants' most common reasons for using their website of choice were convenience and the variety in the

## Chapter 4: Results

selection. Some users stated that their use of a website was partly determined by its compatibility with a device they owned, such as a kindle.

Table 4.  
*Survey questions continued*

	Q.9 Prefer to read physical books or E-books	Q.10 Why is this your preference	Q.11 Electronic devices	Q.12 Favorite Genres	Q.13 Prefer to browse in a physical library or through a digital library	Q. 15 Reads for work
Participant 1	Physical	better concentration and more likely to read when I have a physical book	laptop, smartpho ne	Manga, sci-fi, dystopian worlds	digital library	No
Participant 2	E-Books	E-books have audio features and don't need physical storage	Laptop	non-fiction	digital library	No
Participant 3	Physical	Easier to use for reference at home	laptop, smart phone, tablet	Sci fi, Historical, Psychology, Sociology, Black Experience Autobiography	Physical library	Yes
Participant 4	Physical	I don't have to worry about power	Desktop, Laptop, Smart phone, Tablet	Science fiction	Digital library	N/A
Participant 5	Physical	I look at too many screens, paper is easy on the eyes		Mystery	Digital library	Yes
Participant 6	E-books	I LOVE reading physical books. However, as I have gotten older, physical books are difficult to read: the font is too small, the contrast between ink and paper too low, and they are heavy. E-books allow me to indulge my passion for reading while allowing me to adjust for these variables	Laptop, Smart phone, Kindle	Science fiction	Physical library	No
Participant 7	Physical	I like to turn the pages	Laptop, Kindle	Anything except romance	digital library	N/A
Participant 8	E-Books	eBook	Laptop Smart Phone	Military history, Sci-fi	digital library	No
Participant 9	E-Books	Feels better to pick up and get going	Desktop, Laptop, Smart Phone	History, Psychological Fiction, Political fiction	Digital library	N/A
Participant 10	Physical	I like the physical aspect of turning the pages and the satisfaction of knowing I'm getting close to the end by just looking at the book. I also don't really like the idea of spending that much time looking at a scree	Laptop Smart phone	Romance, mystery	Physical library	No
Participant 11	Physical	Above statement: Am so tired of looking at a screen. Also, when travelling, I find that a small paperback can be more comfortable to have in your hands. There is a great Vanity Fair article from years ago talking about how fun it was, pre- e-books, to look around if you were on public transportation and see what people were reading. Now, w/e-books, you can't do that anymore	Smart phone	Informative	Physical library	Yes

## Chapter 4: Results

Participant 12	Physical	I like to have physical copies of my stuff	Desktop, Laptop, Smart phone	History	digital library	N/A
Participant 13	Physical	When I read it is often to take a break from a screen (I'd be reading on an iPad) and I don't retain the information as well when I'm reading on a device. I'm more likely to skim that way.	Tablet	Mysteries	Physical library	No
Participant 14	E-Books	More portable	Laptop, Tablet	Romance, non fiction	Digital library	Yes
Participant 15	Physical	Forces me to actually read and analyze what I am reading. But e-books are quicker	Laptop, Smart Phone	Fantasy, Poetry	Digital library	No
Participant 16	E-Books	more accessible	Smart Phone	Fiction	Physical library	No

This table covers information about user preferences regarding digital reading. There were no noticeable trends in the types of devices that they preferred to read with. However, despite the preferred type book being predominantly physical, participants answered that they preferred to browse in a digital library. This was an unexpected finding, and nothing in the usability testing sessions was focused on explaining this. Do participants spend less time in physical libraries because they don't find the physical library experience to be sufficiently engaging, or because they have fewer opportunities to visit physical libraries? Are they more likely to browse in digital libraries because they prefer this browsing experience, or because it is more convenient? Participants that read for their occupation were split evenly based on whether they preferred a physical or digital library. Although, if these participants preferred a digital library their preferred type of book was often an e-book with the exception of one participant; If they preferred physical books their preference was toward physical books.

Q9: Preference for physical or E-books: Unlike the previous questions that tied reading habits to ownership of books or what digital libraries they visited, this question asked purely for the preference of reading experience. Despite the reliance on digital libraries to acquire their books, most participants claimed they prefer to read physical books.

Q10: Reason behind physical/E-book preference: Again, this follow up question allowed participants to elaborate on their experiences with both kinds of books. Most users responded that they preferred to read physical books. There were several common themes in their answers. Participants enjoy the physical experience of books like turning pages, physical books were comfortable to hold, or the lack of a screen that strains their eyes. Participants who responded

## Chapter 4: Results

that they enjoyed E-books were more concerned with the physical storage of multiple books or valued the access they could have to multiple e-books through devices.

Q11: What electronic devices are used to read: As we have learned through the history of libraries, more and more books will be published. As this happens, the number of devices people use to read and store these books will increase as well. The available options in the survey were Desktop, Laptop, Smart Phone, Tablet, None, or Other (which they were able to fill in). Out of the digital devices that participants used to read, laptops and smartphones were selected the most frequently (13 laptops, 12 smartphones). Desktop and tablet were each selected four times. Every user answered that they used some manner of device. The only response that was filled in for “Other” was for the Kindle, which was filled in twice.

Q12: Preferred genres: This question was intended to understand whether user tastes might influence what books they browse for and ultimately select. Although some users’ selection during the test sessions aligned with their reported tastes, there was no correlation overall.

Q13: Preference for in-person or digital browsing: This question attempted to understand if participants preferred using library shelves over a digital interface. It was included to understand if users might have a better experience browsing a digital library with tools that resemble a physical shelf in some way. Ten participants answered that they prefer to browse through a digital library rather than inside a physical library. The remaining six answered that they prefer physical.

Q15: Does your job require reading: This question was intended to understand whether users might be required to read multiple books outside of their leisure or interests. Only a quarter of participants were required to read books regularly for their work; another quarter of participants replied no. The remaining half of users selected not applicable. When asked during the user sessions, the participants who indicated they were required to read multiple books for their job were employed in education and psychology positions. Although these careers would likely use a digital library more often than an average person, it is unlikely they would use a digital library daily like librarians. I did not feel that these participants needed to be excluded from the study.

## Chapter 4: Results

### User Patterns

Users were very comfortable using the Barnes & Noble site because some initial categories and book cover thumbnails were prominently displayed on the web page. As previously mentioned, some categories were titled and had content aligned with holidays and current events, specifically black history month. Several tools were popular in this digital library. Pre-defined categories were used by most of the participants while browsing through categories and results pages. Barnes & Noble provided users with a set of initial categories (based on book status rather than subject matter; I.e., top 100 bestsellers, new releases, coming soon, signed edition). Once the user has clicked on a category, a broad selection of subcategory filters are provided on the left side of the page. However, these subcategory filters are not at all consistent across categories. It also became apparent that the sorting options provided by the digital library (best sellers, newest to oldest, oldest to newest, price – low to high, price - high to low, title – A to Z, Title – Z to A) did not meet user needs. For example, one user attempted to sort a set of results by the author but could not do so with the available features. Her workaround-stratagem was to sort the books alphabetically by title, change from a grid view to a list view, and cycle through the individual pages of results until she found the specific book she wanted. This stratagem was successful because the participant knew what book she wanted. There was no apparent way to “browse” books by a particular author without using search.

Barnes & Noble's website navigation was used heavily by every user. The Books tab was used to begin browsing sessions more frequently than any other tool on the main page. Using the Books option in the site navigation leads to a set of subcategories based on book subject matter and genre.

Tools that were used almost as frequently as “Books” were “Best Sellers” and “New Releases.” Users often made use of these when looking for books for the second task. They felt that these categories gave them a sense of what is currently relevant. Barnes & Noble was also well suited for the second task because users could easily see book prices and whether books were hardback versus paperback in order to make moves based on their personal preferences. While users were browsing on individual book pages, the tools that they preferred to use were the books that were recommended with “customers also bought.” Several users mentioned that

## Chapter 4: Results

they felt that they needed to "judge a book by its cover" because there are no overviews on the results pages. Overviews are only presented when a book has been selected. Participants avoided spending unnecessary moves or tactics on books they could not quickly identify as relevant. This was true for all of the digital library sites, in contrast to Netflix, which provides movie overviews upon mouseover. Two users only selected books and categories advertised with large banners during these user sessions.

Book Depository had the most categories and subcategories of any of the selected digital libraries. Almost every user began browsing by clicking on the categories tab on the home page and then selecting a category. The history and archaeology category was the most browsed during each session, due to the assigned task (find a historically important group of people). Once users had selected this or another category, Book Depository provides a rich set of subcategories on the left side of the results page. On the pages for individual books, a set of related categories is provided underneath two rows of suggested book thumbnails. Users primarily used the subcategory links while on the results page. Only two users interacted with subcategories provided on the individual book pages. Although most of the users appreciated the abundance of categories, some felt confused by categories that they found unconventional. For example, a user searching for books on ancient Egypt selected the category "earliest times to present day." This category allowed her to refine her search by selecting time periods with the suffix CE or Common Era. She was unfamiliar with this particular term and what particular span of time ancient Egypt had existed in relation to the Common Era. This same user was also determined to find books concerning the Black Panther Party. She had attempted to find this material on Project Gutenberg but eventually decided to change the target group she was looking for while using that library. With this new tactic, she finally came across books about the Panthers on the Book Depository site when she used the "people also bought" category to progressively move towards content related to the panthers. Multiple users browsed through the "people also bought" section as a tactic to quickly find related content or to find content by the same author. Several users noted that they appreciated how much easier it was to browse by the author on this library than others. Some users also searched the "people also bought" category for related authors. Others used the Top Authors category to search for specific authors they had difficulty finding on other digital libraries. This website's abundance of author information and

## Chapter 4: Results

categories allowed users to develop varied tactics and stratagems that were not as easily executed on the other digital libraries.

On Project Gutenberg, most users began navigating their websites via the bookshelves feature, located about halfway down the homepage, or hidden in the second level of the site navigation. (For every main navigation item, there is a list of secondary navigational items that can be opened by clicking on an arrow on the right side of the page, but these arrows are not visually prominent and are far away from the navigational menu text.) Unfortunately, clicking directly on the main navigation takes the user directly to landing pages that do not provide access to the second level navigational pages, making it easy for some of the participants to miss the second-level pages completely. The location where users found the bookshelves tool could influence whether they began browsing within a few moves or after trying several tactics.

The Bookshelves tool lists several main categories above an alphabetical list of all the categories in Project Gutenberg. The latest books category was also used fairly frequently during the test sessions, probably due in part to its placement. It was placed center screen underneath a carousel of book covers near the top of the homepage. Users found that some of the category labels were misleading. For example, after searching through history and going deeper into the archaeology, category one user clicked on a "current history" category. She was presented with books that she felt were periodicals. One user questioned the relevancy of books in the Journal of Negro History since African American was a more modern term (this participant may not have realized that most books in Project Gutenberg were published at least 95 years ago). Another group of books under "Frequently downloaded" provided users with choices based on the last one, seven, or thirty days and then the top 100 e-books over the last twenty-four hours.

Project Gutenberg has three category options available as links on the book search page accessible under the "search and browse" tab in the site menu. (The "bookshelves" were also available under the "search and browse" tab.) One of the category options on the "search" page was "Popular books." Some users did not find this category, even though it was also reachable by a hyperlink on the main page.

Some of the information provided by Project Gutenberg was confusing for participants because it was unlike the information provided on other digital library sites that participants had

## Chapter 4: Results

previously encountered. For example, some participants mentioned confusion about the “downloads” number provided on the results page after clicking a bookshelf category. It was not clear to these users whether this number stood for the number of available e-books for download or the number of times users have downloaded books from each subject. It did in fact represent the number of times books in that category had been downloaded. Once the user has clicked into a subcategory, information is provided about how many times each individual book in that category has been downloaded. In contrast to the other digital libraries, Project Gutenberg does not provide book overviews and summaries on many of the individual book pages. Project Gutenberg has apparently chosen to invest the limited time of its mostly volunteer workforce in making more books available rather than in providing summaries to books that are already freely available, and thus do not “need” to be promoted. However, some users preferred the overviews they saw on other libraries and did not spend time reading through the prefaces or introductions of books they happened upon while browsing.

In another decision that works against the way that users have been shown to make book choices (Gudinavicius & Suminas, 2017), some pages, such as “Emily’s List,” list books by title and name of other, but without providing thumbnails of the book covers. As a result, participants were not sure what relation these books had to each other or if the items listed were books at all. A few participants clicked on a few of the titles to find that they were books. Others turned away from these pages in favor of the bookshelves tool that leads to results pages with book covers. In a similar issue related to book covers, several participants noted that some of the book covers seemed strange. Book covers seem to be procedurally generated when the website does not have a cover for a specific E-book. Thus, the Gutenberg Project tried to support a variety of browsing approaches, but the browsing experience suffered from inconsistent interactions, inconsistent paths, inconsistent portrayal of results, and the decision not to provide book overviews.

For the Open Library, users predominantly began browsing with the subject/genre categories provided at the top of the main page in a carousel. This carousel, called browse by subject, includes categories for art, science fiction, fantasy, bibliographies, recipes, romance, textbooks, children, history, medicine, religion, mystery and detective stories, plays, music, and science. Some participants also made use of the browse button at the very top of the main page. They discovered that this route would provide them with several sub-categories, where they



## Chapter 4: Results

needed to make an additional selection. In contrast, the browse by subject carousel would take participants immediately to a results page containing a carousel of books. Below the carousel, the page displays a graph showing the number of books published by century in this category, a dense collection of links for related categories, places, people, and times, followed by a long vertical list of “prolific” authors in this category, in no particular order.

Unfortunately, most participants were confused or unsatisfied with the books presented in the carousel. Each participant mentioned that the book titles were hard to read or that the books presented made them feel like they were in the wrong section. For example, most users began browsing in either the history or archaeology sections; however, they routinely were shown books like *Huckleberry Finn* or *Macbeth*, which might conceivably be seen as historical literary pieces but were not seen this way by the study participants. Participants also struggled with the publishing history graph. Its size and placement on the results page made it seem important, and many of the participants expected it to be interactive in some way, although it is not. Next, participants tried to use some of the dense links in the “Related” section that provides users with four sub-categories: subjects, places, people, and times. Some participants were confused by these subcategories; others did use them to navigate to another results page. The participants that used these subcategories developed tactics where they would navigate back and forth between several results pages while browsing. The “Prolific Authors” section also elicited mixed reactions. Some participants appreciated the ability to explore by author. However, this section was only valuable when the participants recognized the author; they did not use it to explore works by unknown authors. Some works also seemed to be titled in other languages, which dissuaded users from browsing books by a particular author.

When participants reached the tasks for Netflix, they scrolled down to see more featured categories more often than searching for any type of navigation. This allowed participants to find shows without spending excess moves and tactics on categories, facets, and filters. Since the task was to find shows that each user had not seen before, a common resource was the category for new releases. Netflix advertises its newest shows at the top of the screen. One show is always advertised and takes up the top half of the header’s screen real estate. Then Netflix begins with its iconic rows of carousel categories with thumbnails. Some of the most common categories to find in the first few categories are “popular on Netflix,” “trending now,” and “top 10 in the US

## Chapter 4: Results

today.” Almost every user used tactics involving these initial categories at some point during their browsing session. The popularity of certain shows and friends' recommendations played a significant role when users selected certain shows. An unexpectedly common secondary tactic that was used was browsing under the tab "more like this." Several users spent time clicking on the page for a particular show and then browsing through these recommended shows rather than returning to the main page for related content. However, the primary method of browsing for nearly all participants was to scroll down continuously on the main page and sort through the additional categories that Netflix would generate until they found a category or a thumbnail that sparked their interest.

## Chapter 5: Discussion

The research shows that participants could browse all of the digital libraries selected without using the search bar. The browsing behaviors that they exhibited were in line with the browsing behaviors identified by Bates (1989), including area scanning, subject searches, and author searching. Only one participant exhibited the journal run search behavior. The digital libraries displayed various tools, but there were no tools that were as successful as the search bar on their own. While users browsed Netflix, their experiences became more focused around the covers/titles of shows and the reputation of actors and other creators. Perhaps the major advantage streaming websites like Netflix and others have over digital libraries is their ability to convey the nature of their content within seconds in multiple ways. Video thumbnails can present users with screenshots of familiar stars or imagery that allows users to make split-second inferences. Mouseovers can trigger previews to play or otherwise reward even minor curiosity by providing additional information for almost no additional effort. Users enjoyed scrolling to see the content without needing to spend extra effort to filter sets of content or learn about individual shows. Digital libraries provide users with browsing experiences that meet their needs with varying levels of adequacy; however, each of the digital libraries tested in this study still lags behind Netflix in the execution of McKay's guidelines.

Barnes & Noble supported many of McKay's guidelines. Barnes & Noble's has an excellent overview while browsing. It provides participants with subject categories in the header under books, and on the left-hand side of the search results page. Participants browsed through the subject categories by hovering over the books tab. The consistent use of category headers helped to orient participants, while the local navigation with subcategories was also relatively helpful. Participants did have trouble backing out of subcategories, however.

Barnes & Noble shows many books at every stage of browsing, particularly in the results pages. It is a solid example of flexibly rearranging shelves because participants were able to change the layout of the results page and the subject content. For mark as read, Barnes & Noble remembers the books that participants had recently viewed and displays these "recently viewed" books at the bottom of individual book pages. Barnes & Noble does well at place-marking and is able to support rapid seamless zooming but not in ways that were consistently discoverable

## Chapter 5: Discussion

across participants. Barnes & Noble, as well as other commercial libraries, could usefully focus on improving the visibility of its overview and rapid seamless zooming tools because streamlining the browsing or search process will likely lead to increased sales.

Since Book Depository is also a commercial library it shares many similarities with Barnes & Noble. The overview generally follows the same formula, where the header is static and one tab can be hovered over to browse through categories. That leads to results pages with categories, a popular authors section, and other filters. This library does the best job of following the guideline for rapid seamless zooming—once participants discovered the category links in the main navigation—because it provides users with unique categories at every stage of browsing. These categories allow users to quickly pinpoint a specific subcategory very early in the browsing process, while also making it easy to back out of a category. This digital library also boasts the best example of rearranging shelves and successfully displaying a lot of books. It has rows of carousels that allow users to scroll through multiple categories or click to rearrange a particular category to browse further. There are buttons in between each carousel that are titled after sub-categories that can change the selected subject. Clicking one of these will change all of the content on the page to reflect the new sub-category. These buttons are also on individual book pages allowing the ability to jump within the hierarchy as well. Participants were overall pleased with these categories and their placements. This website tackles place marking by labeling what subject participants are browsing under at the top of the page. Book depository does not have a feature that directly covers mark as read, although it supports the other guidelines well enough that it had a minimal impact on participant sessions. Users encountered so many new books and categories that they never mentioned running into books repeatedly. Overall, this digital library performed the best regarding the McKay guidelines.

The Open Library held up several of McKay's guidelines, although participants were less comfortable on this website than they were with the commercial digital libraries. The participants' experience on this site demonstrates that the McKay guidelines are insufficient by themselves to guarantee a positive user experience. In many cases, the site supported a guideline, but its implementation fell short. For example, it did not display large numbers of books outside of the home page. Once participants reached a results page there was only one carousel that was present to browse through. However, the site does integrate book set filtering and visible

## Chapter 5: Discussion

alternatives into its search pages. The links that are located under the “Related” section also provided numerous opportunities to try new tactics. The categories of subject, places, people, and times allow readers to filter sets of books. However, the sheer density of these links often caused participants to spend excess time going through them; some users deliberately avoided these links to save time or energy. The prolific authors section was appreciated by several participants that liked to search by author, fulfilling their need for visible alternatives. Rapid seamless zooming on the website suffered because of this. Lists are also unique sets of information that can be provided to the user while they browse. The lists on this website were unique among the selected digital libraries.

The lists on this site seemed like a potentially useful tool. Lists like these could complement book set filtering by allowing users to group relevant discoveries they have made in digital libraries. Some participants used the lists made by other users to find relevant books, but no participants used this tool to organize the books they found themselves. The Open Library’s overview does confine users into relying on certain tools, making the feature less flexible. The Open Library was sufficient but not outstanding when it comes to place-marking. Participants understood where they were in the information retrieval process, but participants often spent time deciphering the results page when they first saw it. Mark as read was not supported well on this digital library. The related subjects on the results page would often be recycled as participants clicked on new links. There was no indicator anywhere on the website that let users know that they had already discovered a book. The Open Library had the weakest example of the rearranging shelves guideline. This was due to the lack of a large number of books on the results page. There is only one carousel provided, and to find more books participants needed to select a new category, list, or author which would take them to a new results page. The architecture described essentially makes each results page one shelf. This behavior is a stark contrast to the commercial libraries that had pages of books with multiple rows, categories, and carousels that could dynamically change based on selected filters or facets.

Project Gutenberg did not support the McKay guidelines well. Users relied on the bookshelves tool and this limited their usage of other features of the website. Although the website supported multiple ways to view large numbers of books, the lack of book covers on tools like top one hundred lists discouraged users from experimenting with different tools. The

## Chapter 5: Discussion

same can be said for visual alternatives. When users view individual books there are options for them to continue browsing for related books. However, these tools are displayed only as text or a hyperlink instead of as book covers or carousels which were used regularly on the other digital libraries. The book search tab addresses the overview and interleaved search and browse guidelines because it provides users with a way to search and browse using every tool except the top one hundred lists. However, this overview does confine users browsing process to a degree. Under this tab, users gravitated towards the browsing options tool which produced results that were organized by language, alphabetically, or by author alphabetically. However, when users interacted with this tool the results they were presented with had no book covers, which made them unsure if they had found relevant books. So even though Project Gutenberg does follow certain guidelines, since they are not recognizable, users interact with the site as if the guideline is not implemented. The guidelines that project Gutenberg outright does not address are rapid seamless zooming, book set filtering, and place marking. There were no ways for users understand where they were in the browsing process without relying on the back and forward buttons in their web browsers. This behavior ties into all three of these guidelines and users' reliance on bookshelves instead of the book search overview page. Mark as read was only addressed by hyperlink text color changing to purple. The website's home page also uses a majority of screen space on aspects of their organization instead of books. As an organization that relies on volunteer help they should attempt to advertise their need and keep regular users informed. However, other digital libraries use their home page to display as many book covers and categories as possible. Participants on this site would occasionally explore certain informative links first and then backtrack to the main page until they found a browsing tool. Their need to rely on volunteers may show that this organization may not have as many resources to update their collection as other libraries.

Netflix has implemented all of McKay's guidelines the most effectively. Netflix interleaves the search and browsing process with its galleried views. Browsing happens while users scan through categories vertically and search happens when they scroll through horizontal carousels. Netflix allows book set filtering through the filters that change the content on the page. Place marking can be observed when users click on or mouseover a title. Either action will provide users with trailer and description of the selected show. The thumbnail of the show expands to a larger size than others in its category, and although the selected show is visibly the

## Chapter 5: Discussion

largest other shows would still be visible. This takes into consideration the guideline of visible alternatives. Netflix marks shows as read at multiple levels while browsing. Shows that are unselected but have been watched have a progression bar over the title thumbnail. When a show has been selected text changes from play to resume. Rapid seamless zooming is one of the guidelines that Netflix excels at. Transitioning between categories in the home view, or to individual shows only requires users to scroll and make one move. Netflix provides an overview through its galleried main page as well as the header that allows users to change between mediums like TV or movies. Not only can users jump through these hierarchies quickly but Netflix also minimizes the number of levels available so users do not get confused. Netflix's home page gallery covers the entire screen with title thumbnails. Netflix also has some of the most effective rearranging shelves. Scrolling down can change all of the content that users see whereas moving through a carousel can change one category that users are interested in.

### Survey Results

The survey showed that the tested users were fairly diverse. As a group they were relatively young and educated. There were nine male and seven female users. Despite the data showing that all of the users were at least somewhat familiar with digital libraries, users reported that they had never visited any of the selected digital libraries; with the exception of Barnes & Noble. All of this data can be used to compare similar users who visit each digital library regularly; this will give a full picture of whether the patterns that users developed would continue to be effective as they become familiar with each library.

### Implications

This research was initially intended as a search for a tool that does for browsing what the search bar has done for search. A particular goal was to find a tool that would allow users to browse through large amounts of content. This study has confirmed that search tools, though imperfect, continue to be more successful than browsing tools. Thus, browsing is a behavior that needs to be fostered and nurtured throughout an entire website.

Earlier in the paper, the usage of data visualization to improve the effectiveness of metadata was discussed in helping to organize query results for users. However, during the research sessions, there were times when visualizations and graphics either hindered or did not help users in any observable way. For example, theopenlibrary.org presented users with

## Chapter 5: Discussion

publishing history for books within a certain category, presented in a chronological graph in the middle of a query results page. Most users did not know what to do with this graph and found it unresponsive when they clicked on it. This data was almost unanimously cast aside by all users as a waste of time. This example is not to say that data visualization could not help, but it has no value for its own sake. Rather, it should be used only when necessary or deemed valuable by a digital library's user base.

Book covers were a core feature of the most prominent tools on each digital library, but the difficulties and frustration experienced by participants on the Project Gutenberg site when book covers were missing or flawed really underscored their import. Project Gutenberg fulfilled McKay's guideline of displaying a large number of books; however, on the results page, some covers were procedurally generated, thereby withholding key information from the user. There were no book covers at all on other tools like the latest books and top 100 books. As a result, these tools were not used as frequently as the bookshelves tool.

Categories were essential to every browsing experience, but the categories provided were extremely varied. The categories seem to have been most successful as implemented by the Book Depository. This library provided users with categories and sub-categories at every stage of their browsing session: on the homepage, on results pages, and on individual book pages. However, as categories become more narrow and specific, they ran the risk of being unrecognizable to users. As was discussed in the literature review, unfamiliarity with a needed keyword can be an absolute barrier in search. Fortunately, unfamiliar category labels are not so absolute a barrier in browsing. First, recognition is easier than recall, so users may recognize more category labels than they would be able to generate in the form of keywords. Second, unfamiliar category labels can sometimes borrow a degree of meaning from their context, or the categories with which they appear. Finally, unfamiliar words in a browsing system can still be explored, providing an opportunity to learn while browsing. Subcategories were among the most useful tools for users that were not expected at the beginning of the research.

An important takeaway from the research is the mutual relationship between book covers, overviews, categories and subcategories, and carousels. When thumbnails are grouped as they are on Netflix, these tools provide users with the ability to rapidly garner information about the content presented to them, and each individual movie thumbnail gathers additional meaning from



## Chapter 5: Discussion

the movies around it. One discovery that was made during the study that was not anticipated was users' understanding of the scope of a collection. This was observed with the open library where the number of books in each category is provided. Users were pleased to know how many books they could find in each category. Some made decisions on whether to pursue their current target based on the number of related books in the library's collection. Project Gutenberg provided a number of downloads statistic that several users mistook for the total number of books in that category. Even though these participants were momentarily misled, this leads me to believe they may value this information after repeated use. This scope falls in line with McKay's eighth guideline, overview, because it gives users a top down view of the library's collection or allow them to make inferences based on the amount of content in each category.

This research study has confirmed the utility of many of McKay's guidelines. Perhaps ironically, Netflix is a prime example of most of these guidelines in practice. Similarly, the browsing behaviors that were expected from users based on Bates' prior work were observed in every research session. Carevic's work with the utility of facets helped to focus extra attention on participants' use of subcategories and multi-step browsing. This research helped illuminate why digital libraries have similar architecture as streaming websites but do not yet provide as rich a browsing experience as Netflix. Netflix has used years of data analysis to ensure that the earliest content carousels on the homepage are the most relevant content. Netflix presents users with shows that have been recently released or shows that are trending in popularity. After these categories, Netflix begins supplying users with other pre-defined categories that, over time, are increasingly tailored to users' tastes. During the test sessions in this study, no previous selections had been made, yet Netflix still provided categories that are likely to appeal to a standard user (again, Netflix's years of data gathering provide good default suggestions). Netflix also mixes in popular subcategories. For example, a high-level category may simple be titled horror; however, other carousels may reflect specialized subcategories such as sci-fi horror or binge-worthy horror. Netflix has identified these subcategories through analysis of the movies that their users have grouped together, and can thus link particular movies together with confidence. And over time, as a new customer makes choices, Netflix will be increasingly successful at recommending the specialized subcategories that are appealing to that customer.

## Chapter 5: Discussion

Netflix also does a better job at making key information about each preferred movie available with a minimum of user effort. On each thumbnail, Netflix's users are provided with a trailer and up to three words to categorize the show's content. This extra information supplies users with more of the information that they use to make selections than digital libraries have yet figured out how to do. To some degree, this difference in treatment reflects the difference between a movie and a book. Viewers have always used movie previews as a way to make decisions about movies, whereas fewer readers would be willing to choose a book based on a 30 second audio clip. However, even before Netflix moved to automatic previews on mouseover, they used mouseovers as a cue for supplementing the movie thumbnail with fairly complete text information. Audio clips may not work like movie previews in the digital library environment, but the older text flyouts could be used to provide additional information for minimum user effort.

## Chapter 6: Conclusion

This qualitative, descriptive study was intended to understand what users experience while they are browsing through digital libraries. The methodology was well-suited for researching digital libraries during the pandemic. However, browsing is such a varied behavior that qualitative, remote observation alone could not address some aspects. If the pandemic had allowed it, this research would have benefited from additional qualitative and quantitative eye tracking data. Based on the results, the answer to the first research question is yes. Users were still able to browse through each digital library and make selections for each task. The research also demonstrated that users preferred certain tools over others while browsing each library. The tools that they gravitated to were primarily meaningful categories, subcategories, and carousels. However, these tools were implemented with differing levels of success on each website. Book covers, overviews, and meaningful categorization proved to be the most important factors in a successful browsing experience. The study also demonstrated that Netflix gives users access to all of these tools while requiring a minimum of user effort, whereas digital libraries were more likely to neglect one or more of these tools, to the detriment of the browsing experience. While users were navigating Netflix, they were far less hindered by the lack of a search bar than for every other digital library in the study. This paper has demonstrated some of the explicit ways that users traverse modern digital libraries. It has also shown that although digital libraries and streaming websites have different audiences and goals, their information architecture is close enough that McKay's browsing guidelines can apply to both contexts. Lastly, it has identified some steps digital libraries can take to provide an improved browsing experience, allowing browsing to become a more equal co-strategy with traditional search.

### Limitations

Due to the pandemic, several modifications to the research methodology for this study became necessary. To maintain a safe environment, user testing was held over Zoom. In addition, users were recruited through social media. This convenience sample is likely to have limited the potential diversity of users. For instance there were only three users over the age of forty that volunteered for the research.

## Chapter 6: Conclusion

More significantly, previous studies about browsing behavior often involved eye-tracking data. Prior to the pandemic, this study would have been held in the University of Baltimore's user research lab, with its eye tracking equipment. Eye tracking data would have provided additional insight into how participants interacted with page content, including order of consumption, duration of attention, etc.

It also became clear that additional data about browsing on streaming websites would have been valuable. Unfortunately, since participants were often somewhat familiar with Netflix, and because the site promotes their newest content at the forefront of the main page, the task occasionally finished too quickly. An alternative task, or even additional tasks, would have provided greater insight into the ways in which participants browse the content on this site. Some participants did not even need to scroll down on the homepage before completing their task. It would also have been helpful to assign tasks on a second media streaming site in order to see more than one approach.

I also needed to find a task that was sufficiently directive that success could be measured but that was also sufficiently open that it would require browsing and active selection. Asking participants to find two books about groups of people who changed the world fulfilled these needs. The nature of this task limited the choice of libraries. I needed to use digital libraries that had collections of history. This requirement precluded the choice of other digital libraries that were focused on fiction or other content. Other reasons for excluding particular digital libraries included having an abundance of broken tools, being an intermediary that directed users towards other digital libraries where they could find books, or being a photographic archive of texts rather than a repository of actual books. Some of these libraries provide interesting tools that merit study in future research.

For example, the World digital library has an interactive visual timeline feature that allows specification between media such as books, journals, manuscripts, maps, motion pictures, newspapers, prints/photographs, and sound recordings. It also allows users to specify between what two dates they can search. When clicking on a particular subject, like WWI, users can filter by sub-categories such as military, diplomacy/politics, culture, civil society, etc. This tool is presented with a map, timeline, filters, and an image of each item's cover, a title, and date.

## Chapter 6: Conclusion

The Florida Electronic Library allows users to browse books based on alphabetical order of titles, subject, grade level, e-books, and popular. This library offers also invites users to find books by self-identifying as a particular persona, under the navigational label “interests.” The options are: The avid Gardener, The Caregiver, The cultural enthusiast, the DIYer, The Environmentalist, The Financial Planner, The Health Care Professional, The History Buff, The Inventor, The Job Seeker, The Legal Eagle, The Librarian, The News Junkie, The Pop Culture Fan, The Researcher, The Self improver, The Small Business Owner, and The Student. While “interests” may not be an effective label for this tool, it would be interesting to explore the utility of this persona-based tool in the browsing experience.

Johns Hopkins Sheridan Libraries is a digital library for the Johns Hopkins School of medicine. It has subject guides to help students find resources. It is one of the few digital libraries that were found with a virtual bookshelf. This bookshelf sorts search results vertically with a graphical representation of books along with a title and date of publishing. Clicking on a book highlights the selected book within the shelf and brings in more metadata on the left side of the screen. Unfortunately, this library was rejected because there is no way to reach this virtual bookshelf, or indeed any library content without a keyword search. This factor, coupled with the specific nature of this library’s collection, disqualified it from being selected for the study.

Another notable digital library was the National Digital Library of India, because it has a well-developed system to browse via categories that it calls "subject levels" underneath a browse tab. Each subject level was a drop-down menu with different subcategories that users can select. Selecting a subject at one level changes the available subjects for lower levels. For example, if the first subject level is History and Geography, then the second level can drop down options for the history of Asia, Europe, or Africa. If a user selects Africa, then the third level options become Egypt, Algeria, or Sudan; whereas if users selected the history of Asia, the third level subjects provided will be Iran, Japan, or China.

### **Recommendations for future study**

If this study were to be conducted again in the future, I would make several changes. First, as I have already stated, I would incorporate eye tracking into the research sessions. Making this change would also require more control over the device that users used during testing. I recommended that users perform testing on desktops or laptops, but I allowed users to

## Chapter 6: Conclusion

participate using their chosen device, which introduced extra variation in their experiences. Some users were able to perform the tasks on smartphones. I would also administer a post-research survey to gauge how satisfied users were with their experience or get them to record what they considered difficult. I would also like to test a larger pool of users. Although users ended up being somewhat diverse, further demographic diversity would be desirable. For instance, the next study could compare users of different genders, ages, or education levels. I would also recommend further testing with other streaming websites. Netflix's interface may be too familiar for many users, resulting in shorter query times and less browsing behavior to observe. Finally, future studies should alternate the order of the tasks administered to each user. This alteration would help researchers understand how much of the difficulty observed during the first task were driven by the unfamiliarity of the website, or whether finding books that are not for recreational reading is simply more difficult regardless of one's familiarity with the interface at hand. These libraries do not represent every digital library that exists today. There are hundreds if not thousands more digital libraries that were not tested in this study. Even the dismissed libraries would be worth exploring despite not fitting this particular methodology. Future researchers could test user browsing habits that focus solely on virtual library interfaces like the bohemian bookshelf, the next on shelf system, or the system at the Johns Hopkins Sheridan Libraries. These graphical tools have yet to be implemented across multiple libraries with larger audiences. The importance of libraries is growing along with the size of their collections, and with that grows our need to understand happens while we browse through library interfaces.

## References

## References

- Bates, M. J. (1989). *The design of browsing and berrypicking techniques for the online search interface*. *Online Review*, 13(5), 407–424. <https://doi.org/10.1108/eb024320>
- Bates, M.J. (1990). *Where should the person stop and the information search interface start?* *Inf. Process. Manag.*, 26(5), 575–591.
- Bell, S. J. (2014). *Collections are for collisions: Design into the experience*. *American Libraries Magazine*, 46 -49.
- Carevic, Z., Lusky, M., van Hoek, W., & Mayr, P. (2018). *Investigating exploratory search activities based on the stratagem level in digital libraries*. *International Journal on Digital Libraries*, 19(2–3), 231–251. <https://doi.org/10.1007/s00799-017-0226-6>
- Carevic, Z., Schüller, S., Mayr, P., & Fuhr, N. (2018). *Contextualised Browsing in a Digital Library's Living Lab*. *Proceedings of the 18th ACM/IEEE on Joint Conference on Digital Libraries*, 89–98. <https://doi.org/10.1145/3197026.3197054>
- Ellis, D. (1989 ). *A behavioural approach to information retrieval system design*. *Journal of Documentation*, 45(3), 171-212.
- Elrod, Jefferson. (1976) *Is the Card catalogue's Unquestioned Sway In North America Ending?* *The Journal of Academic Librarianship*, 4-8.
- Flamenco Search. (n.d.). Retrieved from <https://flamenco.berkeley.edu/index.html>

## References

- Frias-Martinez, E., & Chen, S. Y. (2005). *Evaluation of User Satisfaction with Digital Library Interfaces*. Proceedings of the 5<sup>th</sup> WSEAS Int. Conf. on SIMULATION, MODELING AND OPTIMIZATION, Corfu Greece, August 17 -19, pp. 172 -177.
- Gudinavičius, A., & Šuminas, A. (2018). *Choosing a book by its cover: Analysis of a reader's choice*. *Journal of Documentation*, 74(2), 430–446. <https://doi.org/10.1108/JD-09-2016-0111>
- Haider, S. , Foskett, . Douglas John , Francis, . Frank C. and Estabrook, . Leigh S. (2017, May 12). Library. Encyclopedia Britannica. Retrieved from <https://www.britannica.com/topic/library>
- Kleiner, E., Rädle, R., & Reiterer, H. (2013). *Blended shelf: reality-based presentation and exploration of library collections*. Paper presented at CHI 13, Paris, France.
- McKay, D., Buchanan, G., & Chang, S. (2018). *It ain't what you do, it's the way that you do it: Design guidelines to better support online browsing*. Proceedings of the Association for Information Science and Technology, 55(1), 347–356.  
<https://doi.org/10.1002/pr2.2018.14505501038>
- McKay, D., Chang, S., & Smith, W. (2017). *Manoeuvres in the Dark: Design Implications of the Physical Mechanics of Library Shelf Browsing*. Proceedings of the 2017 Conference on Conference Human Information Interaction and Retrieval, 47–56.  
<https://doi.org/10.1145/3020165.3020179>
- McKay, D., Chang, S., Smith, W., & Buchanan, G. (2019). *The Things We Talk About When We Talk About Browsing: An Empirical Typology of Library Browsing Behavior*. *Journal of the Association for Information Science and Technology*, 70(12), 1383–1394.  
<https://doi.org/10.1002/asi.24200>



## References

- McKay, D., & Cunningham, S. J. (2003). *Browsing a Digital Library: A New Approach for the New Zealand Digital Library*. In T. M. T. Sembok, H. B. Zaman, H. Chen, S. R. Urs, & S.-H. Myaeng (Eds.), *Digital Libraries: Technology and Management of Indigenous Knowledge for Global Access* (Vol. 2911, pp. 329–339). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-540-24594-0\\_34](https://doi.org/10.1007/978-3-540-24594-0_34)
- McKay, D., Shukla, P., Hunt, R., & Cunningham, S. J. (2004). *Enhanced browsing in digital libraries: Three new approaches to browsing in Greenstone*. *International Journal on Digital Libraries*, 4(4), 283–297. <https://doi.org/10.1007/s00799-004-0088-6>
- Nichols, D. M., Chan, C.-H., Bainbridge, D., McKay, D., Twidale, M. B., & Blandford, A. (n.d.). *METADATA TOOLS FOR INSTITUTIONAL REPOSITORIES*. 18.
- Rovelstad, M. V. (1976). *Open Shelves/Closed Shelves in Research Libraries*. *College & Research Libraries*, 37(5), 457–468. [https://doi.org/10.5860/crl\\_37\\_05\\_457](https://doi.org/10.5860/crl_37_05_457)
- Shevchenko, L. (2020). *Analysis of Library Website Users' Behavior to Optimize Virtual Information and Library Services*. *Journal of Information Science Theory and Practice*, 8(1), 45–55. <https://doi.org/10.1633/JISTAP.2020.8.1.4>
- Van den Haak, M. J., de Jong, M. D. T., & Schellens, P. J. (2004). *Employing think-aloud protocols and constructive interaction to test the usability of online library catalogues: A methodological comparison*. *Interacting with Computers*, 16(6), 1153–1170. <https://doi.org/10.1016/j.intcom.2004.07.007>

## References

- Waugh, S., McKay, D., & Makri, S. (2017). *'Too Much Serendipity': The Tension between Information Seeking and Encountering at the Library Shelves*. Paper presented at CHIIR '17 Oslo, Norway.
- Whitelaw, M. (2012). *Towards generous interfaces for archival collections*. *Comma*, 2012(2), 123–132. <https://doi.org/10.3828/comma.2012.2.13>
- Yeh, S.-T., & Liu, Y. (2011). *Integrated Faceted Browser and Direct Search to Enhance Information Retrieval in Text-Based Digital Libraries*. *International Journal of Human-Computer Interaction*, 27(4), 364–382. <https://doi.org/10.1080/10447318.2011.540492>

## Appendix A: User testing materials

**Appendix A: User testing materials****Research Call to Action**

Hello everyone, I am reaching out to ask participants to engage in a research study for my thesis. The subject will be about browsing within digital libraries. If you would like to help figure out why we are all so dependent on a search bar when browsing for books, then this is the opportunity for you! The study will be conducted completely online, so access to a desktop or laptop with a webcam, microphone, and high-speed internet would be best. First, I will ask you to fill out a survey with google forms, and then we will meet via zoom to test out some digital libraries! Please feel free to message me if you are interested in participating.

Thank You!

**Task Script**

“Thank you for participating in the study! Please click on the link I’ve placed in the chat to reach the first digital library. “

***Task one:***

Navigate to the main page.

You are a teaching assistant at a local start up university. A professor asks you to find two books for their world history class. They ask you to find one book each for two different **groups** of people who changed the world. The group can be countries, classes, warriors, organizations, leaders, nobility, workers etc. Once you have selected a book for each please let me know and I will record your discovery. You are not allowed to use the search bar at all. While you are browsing please say your train of thought out loud. If you would like me to repeat the instructions feel free to ask at any time. Please begin.

***Task two:***

Navigate to the main page.

## Appendix A: User testing materials

After finding the books you need for your professor, you decide to look for books to read in your own downtime. Look for two books that you would like to read on your own. You are not allowed to use the search bar at all. While you are browsing please say your train of thought out loud. If you would like me to repeat the instructions feel free to ask at any time. Please begin.

“Next we will navigate to a new digital library, and begin testing there. Please click on the link in the chat”

### **\*Repeat tasks for second DL\***

“Next we will travel to Netflix, and begin testing there. Please click on the link in the chat. I have included login information for the test account”

#### ***Task one:***

Please login with the email and password that I have provided.

Please look for two shows that you would be interested in that you have never seen before. You are not allowed to use the search bar at all. You will have 15 minutes to find them all. While you are browsing please say your train of thought out loud. If you would like me to repeat the instructions feel free to ask at any time. Please begin.

“Thank you for your participation in the research. Please contact me if you have any further questions, comments, or concerns about your participation. Have a great day!”

## Appendix B: User Book Selections

**Appendix B: User Book Selections**

Table 5.

*Selected Books for first digital library*

	DL 1 Task 1	DL1 Task 2
Participant 1	Why we're polarized, Churchill's ministry	One Piece Volume 1, Black Clover Volume 6
Participant 2	When They Call you a Terrorist: A BLM memoir, These Truths by Jill Lepore.	The Illusion of Life, AI Superpowers- China
Participant 3	Diary of a young girl , A short History of the Middle East	The Little Book of Mindfulness, The healing power of essential oils
Participant 4	A promised Land, Ratline	Enders Game, Solar Warden
Participant 5	Kaiser's Holocaust, Musashi	Grow Your own Vegetables in Pots & Containers, Post Office by Charles Bukowski
Participant 6	Famous Men of the Middle Ages, Martin Luther	Criminal Psychology: A Manuel for Judges, Space Station #1
Participant 7	Frederick Douglas, W.E.B. Dubois	The Republic by Plato, Don Quixote
Participant 8	Vie de M.Olier, Lives of the Queens of England from the Norman Conquest	Sphere by Michael Crichton, That Hideous Strength by C.S. Lewis
Participant 9	Romans at War, A brief History of the Vikings	Down to Earth by Bruno Latour, The Miracles of Mindfulness
Participant 10	The Radium Girls, Four Hundred Souls	The burning God, Bone Map
Participant 11	Legends of the Gods (Egypt), The woman and the right to vote	The practice and Science of Drawing, Lessons in Music Form
Participant 12	A brief history of the Samurai, The lost city of the Incas	Romans at War, The Bible Unearthed.
Participant 13	WW1: A history from Beginning to End, Why They Marched: Untold Stories of women	Death on the Nile, The Russian by James Patterson
Participant 14	Ancient Egypt, The Romanov Sisters	The warmth of other suns, The Color of Law
Participant 15	Black Against Empire, Hood Feminism	Athenais: The life of Louis XIV, A drinking Life
Participant 16	Rise of the Rocket Girls, My Own Words (RBG)	Why Im no longer talking to white people about race, Transgender history,

Table 6.

*Selected books for second digital library*

	DL2 Task 1	DL 2 Task 2
Participant 1	Hiroshima by John Hershey, History of United States Naval Operations in WW2	The Invisible Man, Death Note
Participant 2	Concerning Christian Liberty, Myths of Babylonia and Assyria	The Critique of Pure Reason, Dialogues concerning natural religion

## Appendix B: User Book Selections

Participant 3	Sula by Tony Morrison, Americanah	The Invisible Man, On Cooking
Participant 4	Frederick Douglass, The Life of Harriet Beecher Stowe	Mechanical properties of Wood, Wood Carving
Participant 5	History Will Absolve Me by Fidel Castro, All Men are Brothers by Ghandi	The Master and the Margharita, The little Paris Bookshop
Participant 6	Gettysburg, The Cold War	Age of God-Kings, Sense and Sensibility
Participant 7	Diary of a young Girl, Letters from a Stoic	1984, Man's Search for Meaning
Participant 8	The Fairy Land of Science, Aircraft and Submarines	Aircraft and Submarines, The Mastery of the Air
Participant 9	Roman Stoicism, Dante and the Early Astronomers	Dante and the early astronomers, The origin of the mound builders
Participant 10	English Costume from prehistoric times to the end of the eighteenth century, Decision in Normandy	Becoming attached, Enders game
Participant 11	Black Against Empire, The Rise and fall of ancient Egypt	Color and Light, Stick Control- Snare drummer
Participant 12	A journey to the western lands of Scotland, Life in the woods.	The Works, The Bible as History
Participant 13	Classical Art: from Greece to Rome, The connected Discourses of the Buddha	The Sign of Four, The power of Habit
Participant 14	Commentaries on the laws of England, The origin of Species	Betty Crocker's Healthy New Choices, Alice in Wonderland
Participant 15	Knights of Art: Stories of the Italian painters, The Lives of the Poets of Great Britain and Ireland	Mov Rule in New Orleans, The death of the Scharnhorst and other poems
Participant 16	Alice in Wonderland, Speeches by Fidel Castro	The Giver by Louis Lowry, OCD Love Story

## Appendix C: Researched libraries

**Appendix C: Researched libraries**

Institution name	URL	Gallery View	Search Filters	Full text retrieval
Library of Congress	<a href="https://www.loc.gov/">https://www.loc.gov/</a>	Yes	Physical location in library, place of publication, Material types, Language	Yes
Langsdale Library	<a href="http://library.ubalt.edu/">http://library.ubalt.edu/</a>	No	Held by library, Format, Content Type, Publication Year, Author/Creator, Subject, Database, Language, Audience (juvenile, not juvenile), publisher, geography, collection, content provider.	Yes
Dekalb county public Library	<a href="https://dekalblibrary.org/">https://dekalblibrary.org/</a>	Yes	anywhere, subject, title, author, series	No
The British Library	<a href="https://www.bl.uk/">https://www.bl.uk/</a>	No	Access Options, Material type, Author/Contributor, Subject, Creation Date, Language, Publisher, Form/Genre, Journal Title, Additional Features	Yes
Library and Archives Canada	<a href="https://www.bac-lac.gc.ca/eng/Pages/home.aspx">https://www.bac-lac.gc.ca/eng/Pages/home.aspx</a>	No	Relevance, date (new or old), lifetime views, recent views. Word semantics (all these words, exact phrase, any of these words) database, search in (archives, genealogy, images, library), available online, user contributions	No
World Digital Library	<a href="https://www.wdl.org/en/">https://www.wdl.org/en/</a>	Yes	language, exact phrase match	No
National Library of Australia	<a href="https://www.nla.gov.au/">https://www.nla.gov.au/</a>	Yes	title, subject, author, subject, publisher, series, ISBN, Occupation. and you can use Boolean operators to refine your search. Format (Book, video, music, etc.), language, date published.	Yes
Trove	<a href="https://trove.nla.gov.au/">https://trove.nla.gov.au/</a>	No	newspapers, magazines, newsletters, images, research, books, diaries, letters, people, organizations, websites, lists	Yes
National Library of New Zealand	<a href="https://natlib.govt.nz/">https://natlib.govt.nz/</a>	Yes	Type (books, images, etc.), high resolution images, availability, date, collection, usage options (modify, share, use commercially, Unknown, all rights reserved), Creator, Subject	No
Project Gutenberg	<a href="https://www.gutenberg.org/">https://www.gutenberg.org/</a>	No	The search page presents you with options to further sort alphabetically, by quantity, and release date on the left. and the actual options on the right.	Yes

## Appendix C: Researched libraries

University of Texas libraries	<a href="https://web-prod.lib.utexas.edu/">https://web-prod.lib.utexas.edu/</a>	No	Available online, peer reviewed, open-access, physical copy, Resource type, creation date, which library holds it, location in the library, subject, author/Creator, language, new records.	No
Georgia State University Library	<a href="https://library.gsu.edu/">https://library.gsu.edu/</a>	No	Full Text, Catalog only, Peer Reviewed, publication date, source types (Academic journals, conference materials, Books, e-books, electronic resources) subject, publisher, publication, language, content provider	Yes
Online Archive of California	<a href="http://www.oac.cdlib.org/">http://www.oac.cdlib.org/</a>	No	Allows you to limit what institution to search to as well as what date the item was from.	Yes
California Digital Library	<a href="https://cdlib.org/">https://cdlib.org/</a>	No	Collections Program, Contact CDL, Discovery & Delivery Program, Information Services, Licensed Resources, News and Media, News and Media: CDLINFO, UC Curation Center, User Experience Design Services	No
Florida Electronic Library	<a href="https://www.flelibrary.com/">https://www.flelibrary.com/</a>	Yes	A-Z, subjects, Grade level, e-books, Popular, Homework & Research, DIY & small business, Body & Mind, Magazine & Newspaper, Art & Culture, History, Florida	Yes
Harvard Digital Collections	<a href="https://library.harvard.edu/digital-collections">https://library.harvard.edu/digital-collections</a>	Yes	Date range, Type, Language, Repository	Yes
Morehouse school of Medicine Library	<a href="https://www.msm.edu/Library/DigitalLibrary.php">https://www.msm.edu/Library/DigitalLibrary.php</a>	No	Relevance, date, url, word ordering, word proximity, database frequency, document frequency, position in text, depth in site, date bias.	No
University of Maryland Digital Libraries	<a href="https://digital.lib.umd.edu/">https://digital.lib.umd.edu/</a>	No	Item type (audio/video, book, essay, image), collection, and year range, relevance.	No
Johns Hopkins Sheridan Libraries	<a href="https://www.library.jhu.edu/">https://www.library.jhu.edu/</a>	No	Format, library location, publication year, author, organization, language, subject, region, era, series, musical instruments.	Yes
MIT Libraries	<a href="https://libraries.mit.edu/">https://libraries.mit.edu/</a>	No	Relevance, title, year, Boolean operators, publication type, full text, MIT Barton catalog, Print books at MIT, Publication type, Subject, Language, Geography, Publication, Content provider.	No



## Appendix C: Researched libraries

Audible	<a href="https://www.audible.com/">Audible.com</a>	Yes	Keyword, title, author, Narrator or Host, Publisher, Category, Audible Packages, New Releases, Duration, Language, Format, Abridgement, Whispersync.	Yes
Tulane University digital Library	<a href="https://digitallibrary.tulane.edu/">https://digitallibrary.tulane.edu/</a>	Yes	Resource type, language, publication date, Boolean operators,	Yes
Howard digital Library	<a href="https://dh.howard.edu/">https://dh.howard.edu/</a>	No	abstract, subject, author, institution type, document type, publications	Yes
MLK memorial library	<a href="https://www.dclibrary.org/">https://www.dclibrary.org/</a>	Yes	New titles, e-Reader Format, Format, Material Type, Audience, author, Shelf Location, Library, Pub Date, Genre, Subject, Language	No
National Digital Library of India	<a href="https://ndl.iitkgp.ac.in/">https://ndl.iitkgp.ac.in/</a>	No	Access restriction, author, subject category, educational degree, difficulty level, educational level, file format, language, source, content type	Yes
Bartelby	<a href="https://www.bartleby.com/">https://www.bartleby.com/</a>	No	edition, publisher, subject	Yes
Which Book	<a href="https://www.whichbook.net/mood-emotion/">https://www.whichbook.net/mood-emotion/</a>	Yes	Mood and Emotion, Character and Plot, World Map, Bestsellers	Yes
Google Books	<a href="https://books.google.com/">https://books.google.com/</a>	No	Safe search, results per page, spoken Answers, search Activity (uses things you've already searched for on google), Region Settings	No
Internet Archive	<a href="https://archive.org/">https://archive.org/</a>	Yes	Title, Author, ISBN, Subject, Person, Publisher. Also, full text search. Metadata, text contents, news captions, radio transcripts, and websites. Boolean operations.	Yes
Open Library	<a href="https://openlibrary.org/">https://openlibrary.org/</a>	Yes	author, subjects, people, places, times, publisher, language	No
National Library of Medicine	<a href="https://www.nlm.nih.gov/">https://www.nlm.nih.gov/</a>	No	Health Information, Programs & Services, Exhibitions & Collections	No
Digital public library of America	<a href="https://dp.la/">https://dp.la/</a>	Yes	Type, subject, date, location, language, contributing institution, partner	Yes
Hathi Trust Digital Library	<a href="https://www.hathitrust.org/">https://www.hathitrust.org/</a>	No	Provides advanced catalog search (title, author, ISBN, etc.) and advanced full text search (Boolean operators).	Yes
ACM Digital Library	<a href="https://dl.acm.org/">https://dl.acm.org/</a>	No	Boolean operators, publishing date, title, author, people, publications, conferences, reproducibility badges, videos, software, dataset, names, institutions, authors, editors, advisors,	yes

## Appendix C: Researched libraries

			reviews.	
Barnes and Nobles	<a href="http://bn.com">bn.com</a>	Yes	Product type, price, age, best sellers, newest to oldest, title, books, e-books, Textbooks, Teens & YA, Kids, Toys, Games & Collectibles, Stationary & Gifts, Movies & TV, Music	Yes
Amazon	<a href="http://amazon.com">amazon.com</a>	Yes	amazon prime, department, average customer review, book series, New releases, Book format, Author, Book language, amazon global store, international shipping, condition, availability, award winners, bargain books, best of month, best of year, BLM books to read right now, celebrity picks, children's books, deals in books, prime book box, top 20 lists in books, 100 books to read in a lifetime, Amazon books on Facebook, Amazon books on Twitter, Amazon Book stores, amazon first reads, book club picks, from page to screen, start a new series	Yes
Book Depository	<a href="https://www.bookdepository.com/">https://www.bookdepository.com/</a>	Yes	Age Range Keyword, title, author, publisher, ISBN, language, Top Authors, Bestselling series, Art & photography, Biography, Children's books, crafts & hobbies, Crime & Thriller, Fiction, Food & Drink, Graphic Novels, History & Archaeology, Mind Body & Spirit, Sci0fi, Fantasy, Horror, Audio Books, Business Finance & Law, Computing, Dictionaries & Languages, Entertainment, Health, Home & Garden, Humor, Medical, Natural History, Personal Development, Poetry & Drama, reference, Religion, Romance, Science, & Geography, Society & Social sciences, Sport, Stationery, Teaching Resources, Technology, Teen & Young Adult, Transport , Travel.	Yes