

University of
Strathclyde
Glasgow

**“Know what you have and where to find it”:
Managing rare books collections to meet
information needs of multiple user groups**

Arthur B. Charney

Ph.D. Computer and Information Sciences

Supervisors: Professors Gobinda Chowdhury and Ian Ruthven

2026

Word Count: 72,709 words

Declaration of Authenticity and Author's Rights

This thesis is the result of the author's original research. It has been composed by the author and has not been previously submitted for examination which has led to the award of a degree.

The copyright of this thesis belongs to the author under the terms of the United Kingdom Copyright Acts as qualified by University of Strathclyde Regulation 3.50.

Due acknowledgement must always be made of the use of any material contained in, or derived from, this thesis.

Signed:

A handwritten signature in black ink, appearing to read 'A. Charney', with a large, stylized flourish at the end.

Arthur B. Charney

Date: 24 April 2026

Abstract

Managing a library's collections is a difficult enough process on its own; this difficulty increases significantly when that collection is one of rare or special books. Special collections librarians have to consider both the nature and needs of the items under their purview as well as the information-seeking behaviours and information needs of the internal users they serve. This leaves them with three questions this doctoral thesis aims to answer: 1) what key features uniquely define a "rare books and special collections" item and how they are captured in current metadata standards, 2) the potential user groups and their information-seeking behaviours (ISBs) and information needs when accessing these collections, and 3) collection management theories and how they lay the groundwork for combining the key features and metadata standards with the identified user needs in order to effectively meet those needs and promote access to these types of collections. Under the umbrella of interpretivist philosophy, this thesis uses a case study strategy paired with a user-centred design method to examine these research questions. Following both in vivo and literature-based coding of interviews conducted with members of two selected user groups from the Mitchell Library in Glasgow, personas representing both user groups were used to test two sample collections over multiple scenarios, with the conclusion that special

collections items require a much higher level of detail in their metadata to adequately meet the needs of both user groups than items published currently. Alongside this conclusion, this research also provides special collections librarians with a methodology they can follow to identify the needs of their particular users and tailor their bibliographic records to more appropriately meet the needs of their users.

Acknowledgements

My first thanks go to my stalwart supervisors, Professors Gobinda Chowdhury and Ian Ruthven. I had a slightly rough start to my PhD, but with their unending help, patience, and guidance, I managed to get this cart back on the track. Thank you both so much for being patient with me, being so helpful, and giving me so many opportunities to improve my skills both in my own research as well as in being a researcher overall.

Thanks also go to the entire Special Collections department staff at the Mitchell Library. Without you helping me so much with the collection, there would be no thesis.

Next, thanks go to my friends—both those who supported me in real life as well as those who cheered me on over the Internet. There are so many that I cannot name them all: Ellie, Charlotte, Kieran, Abbey, Robyn, Lorelei...you guys are some of the greatest friends I could ever have, and I'm not sure I can thank you enough.

Final thanks go to my family, but most specifically my younger sibling and my dad. Micah, I know me being so far away has been so unbelievably hard for you, but I still appreciate immensely your unconditional support and love through the whole process. Dad...I have no words to describe the gratitude I feel every day that I am

your son, and that I have such a loving and supportive father. I could get really sappy with this, but I'm sure you're probably already tearing up (you big mush!), so I'll save it for our daily video calls. I love you both so much.

Table of Contents

DECLARATION OF AUTHENTICITY AND AUTHOR'S RIGHTS	I
ABSTRACT	II
ACKNOWLEDGEMENTS	IV
TABLE OF CONTENTS	VI
TABLE OF FIGURES	XI
TABLE OF TABLES	XIII
GLOSSARY OF ABBREVIATIONS	XV
CHAPTER 1: INTRODUCTION	1
1.1 RESEARCH QUESTIONS	4
1.2 THE LIBRARY	6
1.3 THE COLLECTION	7
1.4 PREVIOUS WORK	10
1.4.1 WORK COMPLETED IN 2020	10
1.4.2 COMPARING CHARNEY (2020) TO THIS WORK	13
1.5 OUTCOMES AND GOALS	13
1.5.1 OUTCOMES	14
1.5.2 GOALS	15
1.6 So WHAT?	15
CHAPTER 2: LITERATURE REVIEW	19
2.1 WHAT MAKES A RARE BOOK?	20
2.1.1 RARE BOOKS, SPECIAL COLLECTIONS, AND ARCHIVES	20
2.1.2 WHAT SEPARATES RARE BOOKS FROM MODERN PUBLISHING?	28
2.2 THE ROLE OF METADATA AND CATALOGUING	32

2.2.1 WHAT IS IT?	33
2.2.2 DCRM(B)3	37
2.3 USER NEEDS ANALYSIS	44
2.3.1 WHAT IS IT?	44
2.4 COLLECTION MANAGEMENT POLICIES	71
2.4.1 WHAT IS COLLECTION MANAGEMENT?	71
2.4.2 STEPS OF CREATION	81
2.4.3 COMMON ISSUES PREVENTING CMP CREATION	88
2.4.4 THE IMPORTANCE OF COLLECTION MANAGEMENT	93
2.4.5 NEED FOR INTERDISCIPLINARY POLICIES	95
2.4.6 IN SHORT...	96
2.5 SUMMARY	96
CHAPTER 3: METHODOLOGY	98
<hr/>	
3.1 RESEARCH PROCESS DIAGRAM	100
3.2 RESEARCH PHILOSOPHY	101
3.2.1 WHY NOT POSITIVISM?	101
3.2.2 INTERPRETIVISM	103
3.3 RESEARCH STRATEGY AND METHOD	107
3.3.1 QUALITATIVE, QUANTITATIVE, OR BOTH?	107
3.3.2 CASE STUDY STRATEGY	109
3.3.3 USER-CENTRED DESIGN METHOD	112
3.4 THE QUESTIONS	116
3.4.1 QUESTION ONE'S METHODS, DATASET, AND ANALYSIS	116
3.4.2 QUESTION TWO'S METHODS, DATASET, AND ANALYSIS	126
3.4.3 QUESTION THREE'S METHODS, DATASET, AND ANALYSIS	150
3.5 ETHICAL CONSIDERATIONS	154
3.5.1 POSITIONALITY	155
3.5.2 POWER DYNAMICS	157
3.5.3 DATA SENSITIVITY	158
3.5.4 DATA HANDLING FOR THE COLLECTION ITSELF	158
3.5.5 QUALITATIVE RIGOUR FRAMEWORK	158
3.6 LIMITATIONS	161
3.6.1 EXISTING COLLECTION DOCUMENTATION	161
3.6.2 SAMPLE SIZE	164
3.6.3 COLLECTION ACCESS	165
3.6.4 ONLINE TRANSLATION SERVICES	166
3.7 SUMMARY	168

CHAPTER 4: DATA	169
4.1 DATASET ONE: TEST COLLECTION	170
4.1.1 CATEGORIES BASED OFF METADATA STANDARDS	170
4.1.2 CATEGORIES BASED OFF PREVIOUS WORK	172
4.1.3 DISTRIBUTION OF ITEMS ACROSS VARIOUS CATEGORIES	174
4.1.4 DCRM(B)3 AREAS VERSUS CATEGORIES	179
4.1.5 SAMPLE COLLECTION ENTERED INTO KOHA	182
4.2 DATASET TWO: INTERVIEW TRANSCRIPTS AND CODES	182
4.2.1 CODES COLLECTED THROUGH IN VIVO CODING	183
4.2.2 CODES COLLECTED FROM THE LITERATURE	188
4.2.3 IN VIVO CODES COMPARED TO LITERATURE CODES	192
4.2.4 PREVALENCE/NON-PREVALENCE OF CODES IN SAMPLE POPULATION(S)	193
4.3 DATASET THREE: PERSONAS/SCENARIOS	195
4.3.1 LIBRARIAN PERSONAS	195
4.3.2 PATRON PERSONAS	196
4.3.3 SCENARIOS	197
4.4 SUMMARY	199
CHAPTER 5: RESULTS	201
5.1 BEFORE PROCEEDING...	201
5.1.1 DISTINCTION BETWEEN CASE STUDY ONE AND CASE STUDY TWO	202
5.1.2 DISTINCTION BETWEEN SEARCH ENGINES USED BY EACH PERSONA TYPE	203
5.1.3 EVALUATING PERSONA SEARCH RESULTS	205
5.2 CASE STUDY ONE	208
5.2.1 SAMPLE COLLECTION	208
5.2.2 PERSONA TESTING RESULTS	212
5.3 CASE STUDY TWO	238
5.3.1 SAMPLE COLLECTION	238
5.3.2 PERSONA TESTING RESULTS	244
5.4 SUMMARY	275
CHAPTER 6: DISCUSSIONS	282
6.1 REMEMBERING THE OVERARCHING RESEARCH GOALS	282
6.2 QUESTION ONE DISCUSSION	283
6.2.1 DEFINING RARE BOOKS AND SPECIAL COLLECTIONS	284
6.2.2 DEFINING THE RARITY OF THIS COLLECTION	288

6.3 QUESTION TWO DISCUSSION	289
6.3.1 DIFFERENCES BETWEEN LITERATURE AND IN VIVO CODE PREVALENCE	290
6.3.2 COMMENTS ON THE LITERATURE	291
6.3.3 BENEFITS OF THE PROCESS	293
6.4 QUESTION THREE DISCUSSION	295
6.4.1 COMPARISON OF CASE STUDIES	295
6.5 LIMITATIONS, RECOMMENDATIONS, AND OTHER REFLECTIONS	308
6.5.1 QUESTION ONE LIMITATIONS AND RECOMMENDATIONS	309
6.5.2 QUESTION TWO LIMITATIONS AND RECOMMENDATIONS	313
6.5.3 QUESTION THREE LIMITATIONS AND RECOMMENDATIONS	315
6.5.4 TIME CONSUMPTION AND FINANCIAL COST	323
6.6 WHAT DOES THIS MEAN FOR THIS COLLECTION?	326
6.7 WHAT COULD THIS MEAN FOR SPECIAL COLLECTIONS?	327
6.8 SUMMARY	328
CHAPTER 7: CONCLUSION AND RECOMMENDATIONS	329
7.1 THE INITIAL PROBLEM	329
7.2 RESEARCH QUESTIONS	331
7.3 ESTABLISHED KNOWLEDGE	332
7.3.1 GAPS IN THE LITERATURE	333
7.4 CONTRIBUTIONS TO KNOWLEDGE	334
7.4.1 VERTICAL NATURE OF SPECIAL COLLECTIONS METADATA	334
7.4.2 INFORMATION-SEEKING BEHAVIOURS	335
7.4.3 CREATING A PROCEDURE FOR OTHERS TO FOLLOW	336
7.5 FUTURE RESEARCH OPPORTUNITIES	337
7.5.1 FOR SPECIAL COLLECTIONS	337
7.5.2 FOR INFORMATION SCIENCE	340
7.6 THE FINAL WORD	341
REFERENCES	344
APPENDIX A	351
APPENDIX B	352
APPENDIX C	353
APPENDIX D	356

APPENDIX E	357
APPENDIX F	358
APPENDIX G	359
APPENDIX H	360

Table of Figures

Chapter N ^o .	Figure N ^o .	Figure Title	Page N ^o .
1	1.1	Photograph of Sir William Stirling Maxwell (Annan, c.1870)	7
	1.2	Scans of six slips in the Stirling Maxwell slip index (Taylor, 2020a)	8
2	2.1	Metadata standards creation timeline	36
	2.2	Meho and Tibbo's model of information-seeking behaviour of academic social scientists (Meho and Tibbo, 2003: p.584)	52
	2.3	Rhee's model of the information-seeking behaviour of historians (Rhee, 2012)	53
	2.4	Types of policies (Clayton and Gorman, 2006: p.18)	75
	2.5	Collection life cycle diagram (Hibner and Kelly, 2013: p.2)	85
	2.6	Library information life cycle (Clayton and Gorman, 2006: p.188)	85
3	3.1	Screenshot of some spreadsheet entries demonstrating layout	119
	3.2	Functions of NVivo across the six stages of data analysis (Adu, 2019)	146
4	4.1	The three main styles of bookplates present within Dataset One	177
	4.2	An example of a Stirling Maxwell bookplate with three-character combination at the bottom of the bookplate	178
5	5.1	Advanced search using the OPAC	204
	5.2	Advanced search using the internal site	205
	5.3	An example of the front of a slip for an item selected for CS1's sample collection	209
	5.4	An example of the back of a slip for a separate item selected for CS1's sample collection	210
	5.5	MARC record display in OPAC for the record created from Figure 5.3	211
	5.6	Normal OPAC view of the bibliographic record created from Figure 5.3	212
	5.7	Normal internal view of the bibliographic record created from Figure 5.3	212
	5.8	Initial search parameters for patron personas in Scenario One	214
	5.9	Search field change used in Scenario Three	219
	5.10	Sole search result returned for Scenario Three	220
	5.11	Sole search result returned for second search in Scenario Four	222
	5.12	Search results returned for seventh search attempt	225
	5.13	Search result returned for seventh search attempt	229
	5.14	MARC record for sole result for seventh search attempt	230

	5.15	Normal internal view of sole search result for third search attempt	233
	5.16	MARC record display in OPAC for item described in Table 5.11	241
	5.17	Normal OPAC view of the bibliographic record created from Table 5.11	242
	5.18	Normal internal view of the bibliographic record created from Table 5.11	243
	5.19	Sole search result returned for second attempt in Scenario Two	248
	5.20	One result returned in third attempt for Scenario Two	249
	5.21	First result returned for final search in Scenario Three	252
	5.22	Second result returned for final search in Scenario Three	253
	5.23	One result for third search in Scenario Four	255
	5.24	Search results for first search in Scenario Two	260
	5.25	Example of a record with an “ex libris” inscription noted	262
	5.26	Example of a record with an ownership-indicating inscription	262
	5.27	An example of one record returned for fifth search attempt in Scenario Three	266
	5.28	Normal view of last result for fifth search in Scenario Three	267
	5.29	Item description containing information on conservation needs	270
6	6.1	Ex libris inscription present in the collection	286
	6.2	Truncated Stirling Maxwell family tree	286
	6.3	Internal MARC display for a record used in CS2	321

Table of Tables

Chapter N°.	Table N°.	Table Title	Page N°.
2	2.1	Differences between materials in libraries and archives (Hunter, 2020: p.5)	25
	2.2	Encoding DCRM(B)3 records in MARC (Committee, 2011)	42-44
	2.3	Archive users and their perspectives (Crockett, 2016: p.39)	48-49
	2.4	Comparing inclusion of social scientist ISBs between studies	55
	2.5	Listing information-seeking stages noted in two studies beside each other	56
3	3.1	Research process diagram breaking down approach based on question	100-101
4	4.1	Mapping of finding aid categories to DCRM(B)3 categories	173-174
	4.2	Breakdown of major languages used within texts in Dataset One	175
	4.3	Breakdown of minorly prevalent languages used within texts in Dataset One	175
	4.4	Full breakdown of languages present within texts in Dataset One	176
	4.5	Breakdown of estates present within Dataset One based on bookplate	177
	4.6	Breakdown of other special descriptors present in Dataset One	179
	4.7	Comparing DCRM(B)3 areas to MARC fields used for core-level records	180
	4.8	Comparing finding aid categories to MARC fields used for core-level records	180-181
	4.9	Quantified breakdown of in vivo codes in patron interview transcripts	186
	4.10	Quantified breakdown of in vivo codes in librarian interview transcripts	188
	4.11	Example application of literature-based codes to patron transcript excerpt	189-190
	4.12	Quantified breakdown of literature codes in patron interview transcripts	190
	4.13	Example application of literature-based codes to librarian transcript excerpt	191
	4.14	Quantified breakdown of literature codes in librarian interview transcripts	191-192
	4.15	Compiled breakdown of all codes present for sample population transcripts	193-194
5	5.1	Recall and precision of patron search results for Scenario One	216
	5.2	Recall and precision of patron search results for Scenario Two	217

	5.3	Recall and precision of patron search results for Scenario Three	221
	5.4	Recall and precision of patron search results for Scenario Four	223
	5.5	Recall and precision of librarian search results for Scenario One	226
	5.6	Recall and precision of librarian search results for Scenario Two	227-228
	5.7	Recall and precision of librarian search results for Scenario Three	231-232
	5.8	Recall and precision of librarian search results for Scenario Four	234
	5.9	Recall and precision of librarian search results for Scenario Five	236
	5.10	Recall and precision of librarian search results for Scenario Six	238
	5.11	Excel spreadsheet entry for item noted in Figure 5.3	239-240
	5.12	Recall and precision of patron search results for Scenario One	246-247
	5.13	Recall and precision of patron search results for Scenario Two	250
	5.14	Recall and precision of patron search results for Scenario Three	254
	5.15	Recall and precision of patron search results for Scenario Four	256
	5.16	Recall and precision of librarian search results for Scenario One	259
	5.17	Recall and precision of librarian search results for Scenario Two	264-265
	5.18	Recall and precision of librarian search results for Scenario Three	268
	5.19	Recall and precision of librarian search results for Scenario Four	269
	5.20	Recall and precision of librarian search results for Scenario Five	271
	5.21	Recall and precision of librarian search results for Scenario Six	274
	5.22	Recall and precision of all patron search results for CS1 and CS2	275-277
	5.23	Recall and precision of all librarian search results for CS1 and CS2	278-281
6	6.1	Comparing recall (%) and precision (%) per scenario per persona type for both CS1 and CS2	295-296

Glossary of Abbreviations

Abbreviation	Meaning
AACR	Anglo-American Cataloging Rules
AACR2	Anglo-American Cataloging Rules, 2 nd edition
BDRB	Bibliographic Description of Rare Books
BIBCO	Bibliographic Record Cooperative Program
CERL	Consortium of European Research Libraries
CIP	Cataloguing in publication
CM	Collection management
CMP	Collection management polic(y/ies)
CS1	Case Study One
CS2	Case Study Two
DCRB (BDRB2)	Descriptive Cataloging of Rare Books
DCRM	Descriptive cataloguing of rare materials
DCRM(B)	Descriptive Cataloging of Rare Materials (Books)
DCRM(B)3	Descriptive Cataloging of Rare Materials (Books), 3 rd edition
ESTC	English Short Title Catalogue
ICA	International Council of Archives
ISB	Information-seeking behaviour(s)
ISBN	International Standard Book Number
ISO/TC	International Organization for Standardization
KOS	Knowledge organisation system
LCCN	Library of Congress Control Number(s)
LCSH	Library of Congress Subject Heading(s)
MARC	Machine-readable cataloging
OPAC	Online public access catalogue
OSA	Open Society Archives
PCC	Program for Cooperative Cataloging
PIS	Participant Information Sheet
RBC	Rare Books Cage
RBMS	Rare Books and Manuscripts Section
RDA	Resource Description and Access
RQ	Research question
SAA	Society of American Archivists
s.l.	Sine loco (“without place”)
s.n.	Sine nomine (“without a name”)

Chapter 1: Introduction

Imagine a library receiving a donation of a pre-built collection—with thousands upon thousands of centuries-old books—and needing to integrate the new collection with their already-existing collections while simultaneously facilitating patron access to both new *and* old. This is an example of a serious problem facing librarians worldwide; one known as the broad and contentious problem of collection management (Clayton and Gorman, 2006; Hibner and Kelly, 2013).

Collection management has one main purpose behind it—to promote and therefore facilitate access to the collections and other items held under a library's or librarian's purview. A detailed look at the literature regarding the subgroup of steps needing to be taken under consideration will appear in the next chapter, but for brevity's sake, there are two main, broad categories that must be worked out for proper collection management to take place (Clayton and Gorman, 2006; Johnson, 2025):

1. Who—who are the patrons who will be needing access to this library/collection/item? What are their needs? What about any internal users (members of the library staff) and their needs?

2. What—what information is contained by the library/collection/item? What metadata can be found by examining the items?

It is through thorough consideration of these two questions that librarians are able to best manage the collections they oversee. As simply noted by Clayton and Gorman, “Libraries exist to serve the needs of their users...In the management of information resources, too, librarians must seek to serve the needs of their users” (Clayton and Gorman, 2006: p.4). Thus, collection development and management must include at its core a consideration of those needs—after all, how can a library best serve its users’ needs if it does not know what those needs are? This covers the first category, the *who* of collection management. As for the *what*, the importance of this is also noted by Clayton and Gorman:

Any institution would be justified in expecting its library to be able to state broadly (1) how selection decisions are made, (2) what is to be acquired, (3) what is to be preserved, and (4) what is to be relegated or withdrawn (2006: p.19).

To clearly answer this question then, it falls to reason that the library must understand what is held in their collections. This can be done through many forms from simple finding aids to complex cataloguing systems, but whatever form this understanding takes needs to be readily available to both library staff and any external users that may be perusing the collections.

The information contained within these various styles of catalogues is what is known as *metadata*. Metadata is colloquially defined as “data about data,” but the stricter definition is that metadata is “the structured, encoded data that describe characteristics of information bearing entities (i.e., *things*)” (Zeng, 2016: p.3). In essence, metadata is what is used to describe and identify an item—in the case of a library, for example, one type of metadata that could be used to describe a book would include the title, or the author’s name. Metadata, both in the creation and application of, is therefore a significant and important piece to the collection management process. Returning to the comment made by Clayton and Gorman from earlier, it can then naturally be extrapolated that the metadata style chosen by a library to catalogue its holdings must also be aimed at the service of the library’s users. This all then circles back to the overarching need for a library to identify its user groups and their particular information needs when they access the collections prior to making the choice of if they will follow a particular metadata style or create their own.

Take again the beginning example of a library receiving a donation of thousands of centuries-old books. Due to their age, these books are given specifically to the library’s special collections department for cataloguing and managing. This department must then identify its user groups and their specific information needs, select a method for capturing the appropriate metadata needed to satisfy those needs, and use that captured information to aid their patrons in accessing these

materials. This lengthy, time-consuming process thus requires the department librarians to answer a series of research questions as seen in the following section.

1.1 Research questions

To properly address the widely present problem of poorly documented special collections, this thesis examines the following research questions and sub-questions below:

1. What qualifies a collection as one that falls under the purview of “rare books and special collections” departments?
 - a) Are there any key features unique to these types of items or this type of collection that sets them apart from conventional, modern-day publications and collections?
 - b) How can these unique features be captured and recorded in the form of metadata?
2. What is the methodology by which librarians may identify who the typical users of rare books and special collections departments are, and what specific information needs they may require from these collections within these departments?
 - a) What features of rare books and special collections items are of particular interest to the identified users?

- b) Are there any changes to the metadata captured in question 1b required after identification of users and user information needs?
3. How can the unique features of rare books and special collections work with their associated metadata to meet the identified user information needs and improve access to rare books for these users?

These are the research questions for the study contained within this thesis. These three questions are closely linked, informing each other through the need to understand them. While the question regarding the *what* is the first research question, in truth it was developed following the formation of the second research question regarding the *who*. After all, to understand the definition of a “typical user” of special collections departments, one must understand what exactly belongs in a rare books and special collections department.

As for the phrase “typical user,” this phrase includes an immense amount of nuance. This nuance is explored in depth in the next chapter, but in short, defining the “typical user” as a historian of some sort is a simplification of the wide variety of those who could be identified as a historian—it includes users like academic researchers, family historians, and those who research history for television and/or film.

The author chose to examine these questions through the lens of a specific collection held by the special collections department at the Mitchell Library in

Glasgow, Scotland. This collection and the work performed on it serves as a case study example of the work and process that must be followed by rare books and special collections librarians both at this particular library and elsewhere.

1.2 The library

The Mitchell Library was founded in 1874 through a bequeathment by Stephen Mitchell “in a trust deed for the establishment and maintenance of a large Public Library” (*Stephen Mitchell (1789-1874) – The 150th anniversary of his death*, 2024). Following the establishment of this fund, the library first opened its doors in 1877—though it moved to its current location on North Street in 1911, where it has remained to this day (*Stephen Mitchell (1789-1874) – The 150th anniversary of his death*, 2024). The library began as a reference library, but added a lending service in 2005 (*The Mitchell Library*, 2026).

Today, it is among the largest European public libraries, holding more than a million items in its various collections. Amongst its collections are the “Glasgow City Archives, Special Collections, Leisure and Lifestyle, and General Services” (*The Mitchell Library*, 2026), but it also houses spaces for the community to use for various non-reading purposes.

1.3 The collection

The collection being examined in this project is known as the Stirling Maxwell Donation or Stirling Maxwell Collection interchangeably. This collection was built by Sir William Stirling Maxwell as part of his larger library sometime between 1843 and his death in 1878. Sir William Stirling Maxwell was an influential Glaswegian during the period this collection was built.



FIGURE 1.1: PHOTOGRAPH OF SIR WILLIAM STIRLING MAXWELL (ANNAN, c.1870)

Upon his death, his son John inherited the library, and split the full collection in two, sending the more unified selection of “emblem literature” contained in his father’s library to the University of Glasgow, and the rest of it in multiple batches to the Mitchell Library sometime during the 20th century. Viewed holistically, the collection donated to the Mitchell Library consists of more than 3,600 items of

literature spanning from the 15th century to the 19th century. Around the time of the various donations, a slip index was created—being a basic form of a card catalogue that only lists basic information about the books in the collection rather than any information about storage—but to date this is the only documentation uncovered regarding the collection.

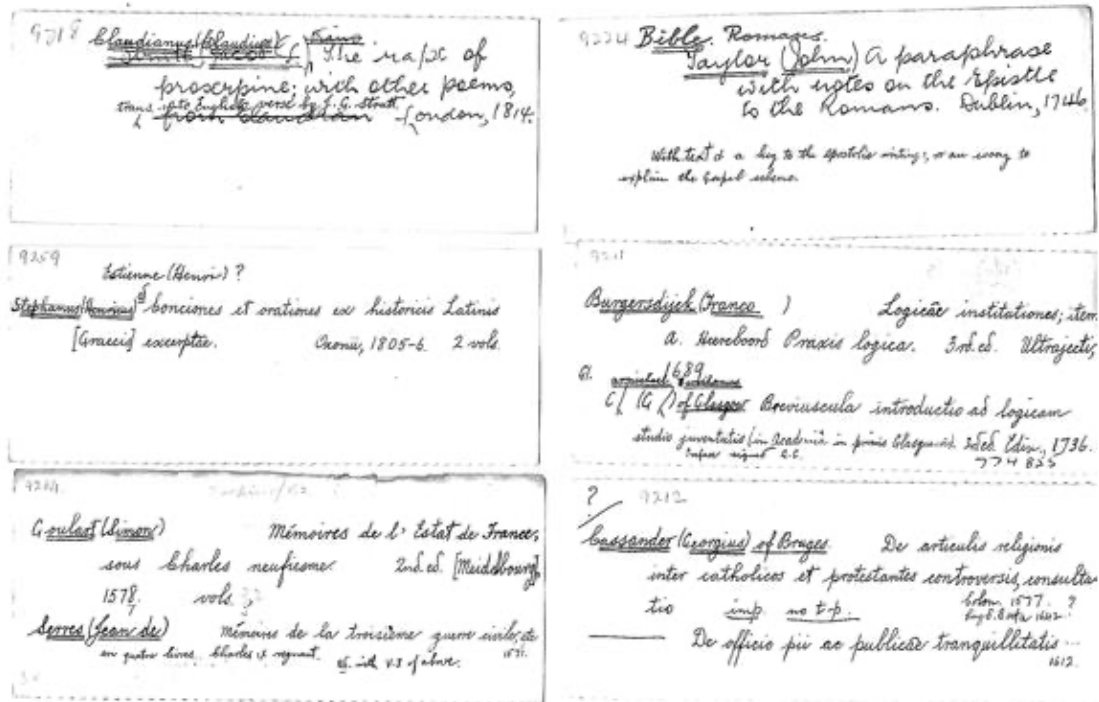


FIGURE 1.2: SCANS OF SIX SLIPS IN THE STIRLING MAXWELL SLIP INDEX (TAYLOR, 2020A)

In addition to it being the only confirmed overall documentation of the collection, it is unclear exactly who is responsible for this index. There is evidence of multiple potential creators present in various handwriting styles amongst the slips (seen above in **Figure 1.2**), as well as evidence that points to these creators possibly being either Mitchell librarians who created the index following the various

donation batches¹ or people employed by Stirling Maxwell's son prior to the donations². None of this evidence is conclusive, as to date there has been no documentation regarding these donation batches uncovered or known to exist. Once it was received by the Mitchell librarians, however, the collection was divided again into multiple other collections held by the Special Collections department (the Wing Collection, Early Foreign Printing, and Early Glasgow Printing to name a few), though very little, if any, documentation about the reasoning behind this decision or the paths the items took into and through storage as the years passed can be found or exists (Taylor, 2020b).

This lack of documentation regarding the donation and storage of the Stirling Maxwell collection is not an issue unique to this collection alone. In fact, it reflects issues that face librarians charged with managing rare books and special collections around the world. Special collections as a whole tend to not arrive at libraries with pre-recorded metadata, especially when these collections are donated from members of the family as was the case with the Stirling Maxwell collection. Oftentimes it is only the books themselves that are donated, and

¹ The evidence of this lies in the fact that multiple of the various handwriting styles present amongst the slip index can also be found in the department's location books, which are believed to only ever been accessible to department librarians throughout Mitchell Library history.

² This evidence can also be found in the department's location books, where there are scattered notes amongst the entries of Stirling Maxwell items that seem to vaguely reference the slip index being created prior to donation.

librarians must progress through the process outlined in this thesis to even be able to begin to understand what their library now holds.

The author's work on this collection began back in 2020 as part of the requirements needed to complete an MSc in Information and Library Sciences. A brief description of this work is contained in the following section.

1.4 Previous work

This postgraduate research project follows off of the author's previous work completing an MSc dissertation in summer 2020 (Charney, 2020). The overarching goal of the previous dissertation work was to create the skeleton of an interdisciplinary collection management policy (CMP) for a public library (the Mitchell Library) with a special collections department, again using the Stirling Maxwell collection as the research lens. While the full policy creation process is not the topic under review in *this* dissertation, there is a significant amount of work contained within this project that builds on work completed in the previous study, which calls for a brief explanation of the prior work.

1.4.1 Work completed in 2020

The work completed in 2020 was performed on the same collection as this thesis, though it was ultimately in pursuit of answering a different set of research questions.

1.4.1.1 An overview of Charney (2020)

To create a CMP based in best practices, a thorough understanding of the literature of CMP theory was developed. In the previous research, a gap in the literature for CMP theory appeared when attempting to examine policies for similar collections (meaning, collections held in libraries that span multiple fields of librarianship—the Mitchell Library is a public library, but has a robust special collections department, which is the sort of intersection not covered by current literature). Thus, to create a thorough CMP based in best practices, an examination of CMP theory in multiple fields had been undertaken.

In the previous work, three fields of CMP theory were chosen. The first two are self-explanatory: public library CMP and special collections CMP. Again, this is due to the fact the collection in question is held by the Special Collections department at the Mitchell Library, which is a public library. The third field chosen at the time was museum CMP, as there was the potential for the items for the collection in question needing to be viewed as exhibitable items (both for public access, but also in terms of preservation/conservation) (Charney, 2020). Further work with the Special Collections department alongside the master's dissertation had also revealed a desire by the department to align themselves much more closely with the museums side of the Mitchell Library's parent corporation, Glasgow Life, which deepened the need for this field of CMP theory to be examined as well.

A large portion of the previous work done with the collection included the beginning digitisation of the slip index via data entry into an Excel spreadsheet as requested by the department librarians. The formatting of this sheet was initially informed by the metadata structure of the slip index, but over time evolved to include other metadata categories the department librarians identified as important to them. It is this portion of the previous work that lay the foundation for the work contained in this thesis, specifically the foundation for the work completed in pursuit of answering research questions 1b and 2b.

1.4.1.2 Work from Charney (2020) contained herein

Information taken from the previous master's dissertation work is included in this doctoral thesis, though not necessarily verbatim. The primary information used herein lays in **Chapter 2, Section 2.4** discussing collection management policies. Some of the same sources are cited within this section as were used in the previous work.

The other information used as the foundation of the work completed for this project is the initial structure of the Excel spreadsheet being used as the finding aid created by the author for the Stirling Maxwell collection. A detailed description and explanation of the structure and its components is contained in **Chapter 3, Section 3.4.1.2**. At the time of submission of the master's dissertation, the information contained within the created finding aid was comprised solely of most of the information contained within the slip index—the slip index had been

alphabetized by author last name, and at the completion of Charney (2020), slips through the letter S had been entered into the digitised finding aid.

1.4.2 Comparing Charney (2020) to this work

The work described and completed in *this* project, rather than focusing on the comparatively broad topic of CMPs overall, narrowed the lens to creating the process librarians must undertake in order to best identify and understand the needs of the users they serve and how that interacts with metadata frameworks exemplified by the collection as a crucial *subset* of CMP theory. This process would then work to facilitate better access not only to the collection at the heart of this case study, but to other related special collections. The work contained in this research project is built on the foundation of the previous work as described in **Section 1.4.1.2** and forms a new method for answering the research questions noted in **Section 1.1**.

With the information about the Stirling Maxwell collection and the previous work completed on it laid out, the next topics to discuss revolve around the intended goals and outcomes of the work completed in this project.

1.5 Outcomes and goals

Now that the research questions have been identified, along with information about the collection being used as the lens those research questions will be

examined herein and the previous work completed in 2020, the outcomes and goals can be laid out.

1.5.1 Outcomes

The first part of the intended outcome of this research is the identification of the unique features that make a book part of rare books and special collections' departments and how those features may be captured through metadata in the pursuit of answering research questions 1, 1a, and 1b. The current idea of what makes a rare book a "rare book" is often vague and can be left up to organisations or even individuals to define the distinction on their own. It is through this work that the author hopes to work towards a more concrete definition that can be generalised both to other collections held at the Mitchell Library and to collections held at libraries elsewhere.

The second part of the intended outcome is the identification of unique user information needs and requirements of rare books and special collections items, as well as the identification of any necessary changes to the metadata captured in answering 1b. This will be laid out during the process of answering research questions 2, 2a, and 2b.

The work completed in answering research question 3 lays out the ultimate overarching outcome hoped for this research; in addition to the two previous potential outcomes, this work also aims to develop a methodology for indexing and

cataloguing rare books and special collections items. This methodology will then stand as a plan of action special collections departments worldwide can follow as they attempt to appropriately manage the collections under their purview.

1.5.2 Goals

Finally, the overarching goal of this research is to open the door for further research on the interactions examined herein. As mentioned earlier, there are numerous literature gaps around CMP theory overall, let alone in the niche of information needs interacting with metadata frameworks and standards. Collection management is an *essential* part of the process of running a library of any kind, and the literature does exist emphasizes the extreme need for comprehensive policies for librarians to follow. As user and content analysis are two pillars of CMP creation, this work aims to show how these analyses can be done in a feasible manner in the case of a collection with little to no pre-existing metadata on hand.

1.6 So what?

Why are these questions so important in the context of special collections? Why are this particular collection and its circumstances so unique in comparison to other collections held not only at the Mitchell Library, but at libraries across the globe, and thus the choice as the lens through which the research questions will

be examined? The answer is a simple one: it boils down to the uniqueness of the collection.

Terry Walker eloquently breaks this down when discussing the importance of special collections as a whole, stating:

Like most special collections, ‘the materials...are...uncommon, sometimes unique, often old, frequently fragile, or just plain expensive and irreplaceable.’ Also like most special collections, [Yakima Valley Libraries]’s special collection materials do not circulate. These materials are special precisely because they are uncommon. Regardless of their market value, they would be very difficult, if not impossible, to replace... (Walker, 2014: p.18).

The uniqueness of the items held in the Stirling Maxwell collection can be measured in such metrics as print runs and how many of those items printed may have survived the passage of time to be held in other collections throughout the world. As much of the collection was published in the 17th and 18th centuries and due to the hand-driven nature of the publication process at the time, books printed at the time were nowhere near as ubiquitous as they are today (Unwin, Unwin and Tucker, 2024). Is this what makes these items part of “rare books” or “special collections”? Is information like that what users of these departments are searching for, or is there some other information more important to them? What is

it about these items and these collections that requires so much work to be completed in the process of indexing and cataloguing them—or are librarians doing too *much* work? These are all questions that will be answered in the following pages of this thesis.

Again, Walker provides the most impactful reasoning as to why even just the physical presence of the items held in special collections are so important, and must be preserved:

If history is a story we tell ourselves about the past, these artifacts are the fabric of that past, linking elements that may allow researchers to build a factual and balanced understanding of not only what happened, but how it happened and perhaps even why. The opportunity for local researchers to have hands-on interaction with primary source documents in a local setting is not something that can be fully replicated on line. This is the reason these documents have been preserved and so carefully cared for—and this treasure trove is available for present and future generations to explore at the library (2014: p.21).

It is for this reason that the management of physical items in special collections is so important and must be examined as the world of librarianship moves ever further into the digital age. Preserving these tangible connections to the past in the physical format and properly managing them in a way to ensure their availability for

those patrons who may be searching for the information held within is tantamount. Important steps in the management of special collections are the identification of key features to capture with metadata as well as user groups and their information needs, followed by the process of indexing and cataloguing the items in the collection or collections being managed. This research will take the reader through this process and construct a methodology that can be followed by rare books and special collections librarians around the world.

Chapter 2: Literature

Review

Prior to proceeding with the data collection and analysis for this project, a thorough understanding needed to be gleaned to lay the necessary foundation for the data collection and analysis phases. The structure of this literature review follows the layout given by the research questions discussed in **Chapter 1**, starting with 1) what key features uniquely define a “rare books and special collections” item and how they are captured in current metadata standards, 2) the potential user groups and their information-seeking behaviours (ISBs) and information needs when accessing these departments, and 3) collection management theories and how they lay the groundwork for combining the key features and metadata standards discussed in the first section with the user needs examined in the second section in order to effectively meet those needs and promote access to these types of collections. This literature review aims to review the current body of research around the three broad questions defined in **Chapter 1**, with the intent to discover the gaps, if there are any, in the existing literature.

2.1 What makes a rare book?

In the recent past, “special collections was regarded as the dry and dusty backwater of a profession” (Ovenden, 2008: p.295). In many ways, perhaps, it still is, particularly by those outwith this particular profession. But what are special collections? Rare books? To truly understand what makes a book part of rare books and special collections, one must first define what the term “rare books and special collections” means—and indeed, if that is one term, or two terms merely masquerading as one.

2.1.1 Rare books, special collections, and archives

Rare books, special collections, and archives are terms often used interchangeably, both by members of the public as well as the custodians of the knowledge held within. As such, the definitions of each are nebulous and can confuse all but the most knowledgeable person. The wider consensus about these three terms is that they are separate terms, and while archives can be separated somewhat more readily, the remaining two have, in practice, become nigh inextricable from each other.

2.1.1.1 Defining special collections

Sidney Berger notes that special collections departments are typically:

...the repositior[ies] of items that are not necessarily monetarily valuable but are distinguishable by virtue of their being assemblages of materials of like

kinds or on like topics. In one sense, “special collections” are valuable simply because of the cohesiveness and extent of the items on a particular topic, in a particular genre, or by a particular author (2014: p.2).

Another defining feature of special collections for Berger (2014) is the need to protect the materials somehow from “mutilation,” whether that is by forces of nature or forces of man in some way or another. This means that a special collection is made up of materials that need either special protection or care, but also have some common theme or thread linking the items together in some way.

Alison Cullingford expands this definition by using the Association of College and Research Libraries’ definition of material items, meaning, “The entire range of textual, graphic, and artifactual primary source materials in both analog and digital formats, including printed books, ephemera, manuscripts, photographs, maps, artworks, audio-visual materials, archival materials, and other objects” (Association of College & Research Libraries, 2020, quoted in Cullingford, 2022: p.xvii). While Cullingford does agree with Berger about the necessity of special care and protection, as well as the common thread linking items with each other being defining features of a special collection, she adds a third defining feature that echoes the words of Walker in **Chapter 1**, stating that, “They are the cultural heritage of institutions, communities and society, have value as evidence, are vast sources of stories and inspiration, and have immense potential for research, learning and public engagement” (Cullingford, 2022: p.xviii).

2.1.1.2 Defining rare books and rare books departments

Defining a rare books department is harder. Since the idea of a rare books department was first conceived to today, the idea has shifted to more closely align with that of special collections. At that unnamed time, “certain volumes were segregated from the parent collection because of their value or vulnerability and put into a separate, more secure environment” (Berger, 2014: p.1). Not every item was a book, however:

But custodians would also have seen, for instance, maps of great rarity, manuscripts that were, by their very nature, unique and irreplaceable, and photographs that were equally unique. And they would have wanted to secure these items as well. It would be only natural to do so in the same place in which the rare books were kept (Berger, 2014: p.1).

Thus, the idea of a rare books department seems to be rooted in the idea of the holdings’ “monetary value and their vulnerability to mutilation” (Berger, 2014: p.2), even if the collection of holdings is lacking any form of common thread in regard to content, theme, or indeed anything outwith the financial value and the need to protect them. Summing it up, Berger states:

To call it a *rare book* collection would have been strictly inaccurate, but the designation does convey the idea of rarity and the need for protection as

well as the notion that this is where research can be done on manuscript and printed sources (2014: p.2).

This is distinguished from the idea of a “special collections department” in Berger’s view by virtue of the items lacking cohesiveness outside of these two qualities.

2.1.1.3 Defining archives

The term “archives” is one that has also become blurred as time has progressed and humanity has entered the digital and information age. According to Gregory Hunter, prior to the information age, archives were known as collections of some form of items—usually books at this stage, but could also include letters and manuscripts—that were stored for some length of time (2020). The advent of the information age brought with it digital repositories and forms of media, each with their own set of jargon, prompting a necessary shift in the idea of what an archive is. Unfortunately, this shift in definition was one more towards ambiguity and nebulousness.

The definition of the equally valid terms archive and archives Hunter thus decides to use is laid out as three possibilities:

1. *materials*: items created in the conduct of affairs and preserved because of their enduring value (also referred to as archival records)

2. *place*: the building or part of a building where archival materials are located
(also referred to as an archival repository)
3. *agency*: the division within an organization responsible for maintaining records of enduring value (also referred to as an archival agency, archival institution, or archival program) (2020: p.2)

He also mentions the two approaches used when speaking about the particular materials held by an archive, naming them as broad and narrow approaches. The overarching approach is, “to use the word *archives* broadly to refer to items of enduring value created or received by individuals, families, or organizations (public or private)” (Hunter, 2020: p.2). In contrast, the narrow approach creates a strict delineation between archives being “items created or received by organizations” (Hunter, 2020: p.2) and the use of the term *manuscripts* as an identifier for “materials created or received by individuals or families” (Hunter, 2020: p.2). The ambiguity created by this second, narrower approach can be most clearly demonstrated with the example of the Stirling Maxwell donation, where it could, in theory, fit *both* of these definitions at the same time: it could be considered an archive, as it was received by the organization known as the Mitchell Library, but could also fit the definition of manuscripts as materials received by an individual.

Hunter also defines archives by using the contrasting facets of what a library is and what an archive is, using the table below:

Category	Libraries	Archives
Nature	<ul style="list-style-type: none"> • Published • Discrete items • Independent significance • Available elsewhere 	<ul style="list-style-type: none"> • Unpublished³ • Groups of related items • Significance from relationship to other items • Unique
Creator	Many different individuals or organizations	Parent organization or institution
Method of creation	Separate, independent actions	Organic—normal course of business
Method of receipt	<ul style="list-style-type: none"> • Selected as single items • Decisions revocable 	<ul style="list-style-type: none"> • Appraised in aggregate • Decisions irrevocable (destruction is forever)
Arrangement	Predetermined subject classification	Provenance and original order (relation to structure and function)
Level of description	Individual items	Aggregate (record group or series)
Descriptive media	<ul style="list-style-type: none"> • Built into the published item (title page, table of contents, index) • Card catalog, online public access catalog (OPAC) 	<ul style="list-style-type: none"> • Must be prepared by the archivist • Guides and inventories, online systems
Access	<ul style="list-style-type: none"> • Open stacks • Circulating items 	<ul style="list-style-type: none"> • Closed stacks • Noncirculating items

TABLE 2.1: DIFFERENCES BETWEEN MATERIALS IN LIBRARIES AND ARCHIVES (HUNTER, 2020: P.5)

This distinction works best when viewing a library and an archive as two separate entities existing in two separate physical locations. Returning to the example of the Stirling Maxwell collection and the Mitchell Library however, again, ambiguity appears. One instance of this ambiguity can be seen in the “nature” category, where the nature of archives mentions that these are generally groups of items with some form of relation amongst themselves, and library items have

³ Hunter concedes this is not a hard-lined distinction, stating that, “archival materials are, by and large, unpublished” (2020: p.4), meaning that an item being published or not is not to be used as the sole method of distinguishing it as a library item or an archival item.

significance as individual items. What if an item held by an archive has significance as an individual item and does not necessarily rely upon it being related to other items in the archive? Or vice versa, the significance of an item held in a library is more apparent when taken as part of a group of other items? Finally, as evidenced by the requests from the Mitchell Library department librarians regarding the cataloguing of the Stirling Maxwell donation discussed in **Chapter 1**, the notion that the level of description remains in the aggregate seems to be an idea that the field is moving away from, albeit slowly.

While Hunter has a broad and somewhat meandering way of defining an archive, Crockett's definition is short and to the point. Her definition simply reads that an archive is defined as "records of one organization, family or individual, selected for permanent preservation because they provide key evidence of the entity's history" (Crockett, 2016: p.4).

2.1.1.4 Term interchangeability and usage herein

These three terms, as mentioned previously, have been and most likely will continue to be used interchangeably by all but the most pedantic. The authors examined above all agree on the fact that the items managed by these departments have a) value of some sort, whether monetarily or not, and b) require some sort of special care and preservation. Past that, the definitions become increasingly pedantic about the smallest of features. Berger even mentions specifically that "rare books" and "special collections" are used so

interchangeably now, both in function and name, that he himself chooses not to make the distinction unless absolutely necessary in his work (2014). The argument can also be made that the definition of “archives” given by Hunter (2020) could be included in this interchangeability, especially when a library *contains* an archives department.

The Mitchell Library, for example, has two separate departments named Special Collections and Glasgow City Archives that each serve different purposes. Going by the statements given on these departments’ websites, it is easy to see how the two can be conflated by the average person. The Glasgow City Archives states, “Here you can enjoy the wonderful documentary heritage of Glasgow and the west of Scotland over eight centuries” (*Glasgow City Archives*, n.d.), while the Special Collections department website answers the question as to what their holdings are with:

Special Collections collects, manages, and offers access to collections including family history, Glasgow local history and heritage, newspapers, maps, photographs, rare books, manuscripts, ephemera, and some artworks. We hold collections on a wide range of subjects which are unique and distinctive, and need to be treated with special care under supervision, such as rare and older books, and manuscripts (*Special Collections*, n.d.).

Returning to the definitions seen in the previous sections, it is clear how the distinction between these departments could be confusing to the average patron, until they interact with the items held by each. In practice however, it becomes clear that the City Archives function under the definition given by Crockett (2016), where it is specified that the records in an archive belong to one organization—in this case, the City of Glasgow.

Where this leaves the rest of this project is thus: Berger's (2014) argument that rare books and special collections departments are theoretically and practically interchangeable is undeniable, especially with the expansion by Cullingford (2022). While Hunter's (2020) definition of archives could fall into this same easy interchangeability, the specification given by Crockett (2016) provides enough of a distinction that means "archives" cannot be included here. Thus, for the remainder of this thesis, "rare books" and "special collections" are used equally with no distinction, and the term "archives" is not used.

2.1.2 What separates rare books from modern publishing?

Now that the basic ideas of what is meant when referring to rare books and special collections departments have been made clear, it is necessary to examine some features that set the items within such departments apart from modern, conventional publications contained elsewhere in the library. It is common knowledge that books are primarily a method of conveying information to their

readers, whether they fall under the purview of a general librarian or a special collections librarian. The idea of rarity and vulnerability being the driving factors for an item being placed in the holdings of special collections is a good start, but does not truly unveil exactly *why* these items may be considered rare or why they require more protection than an everyday book.

Vulnerability and the need for protection of some sort is a factor that does not need much explaining to understand. When one thinks of a rare books and special collections department, the images conjured are generally of centuries-old books printed on fragile parchment and bound in some ancient binding that crumbles away at the simple brush of a fingertip. In this context, the need for protection is clear—this book requires constant preservation and conservation to ensure that it survives into posterity. Rare books and special collections departments do not only hold items like that, however. For instance, Berger points out that, “It is clear that if an institution has something it does not know what to do with and, for any reason, cannot dispose of, special collections is a good dumping ground” (2014: p.3), meaning that much younger, less physically fragile items are also found in special collections. Regardless of their age or physical condition however, all of the items held by these departments still exhibit some form of physical vulnerability and require some form of protection.

Rarity is another vague term that poses trouble for anyone looking to define it. Going by the first dictionary definition, rare means “seldom occurring or found”

(*Rare*, n.d.). This definition is exceedingly broad in the sense of rare books and special collections, especially when it comes to booksellers' listings of items that librarians may be pursuing for new acquisitions for their libraries. In most cases, the use of the terms "rare" or "scarce" are "like the word *very*, which has been so overused that it has lost its meaning" (Berger, 2014: p.11). Thus, the definition that more clearly marks the distinction in question is the second definition: "marked by unusual quality, merit, or appeal" (*Rare*, n.d.). A book with only six copies ever printed may be just as rare as a book that has been printed thousands of times but contains handwritten annotations from an influential local personality.

The true feature of rare books and special collections departments is that their holdings are not treated simply as conveyors of the textual information written by the author. Compared to modern publishing, which is typically printed on extremely similar paper with equally similar binding methods and formats, the variety in manufacturing methods present in most special collections is also treated as information contained by the books. For example, Berger notes a syllabus for a course on historical bibliography at the University College London that describes the learning outcome as, "students will have an overview of the history of the book and the materials that make it up (paper, vellum, bindings, illustrations, print, etc.)" (Welsh, n.d., quoted in Berger, 2014: p.19). Students such as the ones in this course learning about the history of the book most often will do so by examining examples of books throughout time—the exact type of

items held in special collections. There are multiple fields that examine these books as research data in their own right, such as academics tracing book history, printing history, and even something as specific as the history of paper manufacturing. Therefore, knowledge of these materials, their manufacture, and use throughout history is a necessity for special collections librarians to know.

Beyond the simple explanation of the need to know this information about their holdings, this knowledge is important to numerous other fields as well. Knowledge of the material components involved in special collections is needed to not only describe these items to patrons, but also in the sense of what treatment to use in the process of preservation, what environment is best to house them, what packaging may be needed for transportation, and so on (Berger, 2014). Preservation especially is a field that needs specific information about the material makeup of items in special collections, as Cullingford notes here:

To preserve materials in Special Collections, it is essential to understand their physical composition. A large research library is likely to hold clay tablets, papyri, bamboo, bark, bone, palm leaf, vellum, parchment, wood, textiles, paper from various eras, photographic material, sound recordings and other audiovisual media and, increasingly, digital formats. Even smaller Special Collections holding only printed book collections will have different kinds of paper, ink, glue and binding materials to consider (2022: p.32).

Cullingford's list of potential materials only scratches the surface of the massive variety of manufacturing methods that may be present in a special collections department. All of this information is valuable to numerous people for numerous reasons, and as custodians of these items—whatever form these items take—librarians must understand the material makeup of their holdings so as to appropriately manage these collections and facilitate access to the people who may be looking for this information.

The Stirling Maxwell collection falls under what Cullingford (2022) describes as a smaller special collections department, as it is made up solely of books. Even still, there is a large range of physical items used to make these books, each with their own requirements for handling, preservation, and storage. Recording material information about these books is more necessary for special collections because these items are not just textual information conveyors of the words their authors put to page, but provide research data purely by existing. This again returns to the two ersatz defining features: rarity and vulnerability. These items are marked by the uniqueness of their material makeups (rare), and those materials also require very specific and varied methods of preservation and protection (vulnerable).

2.2 The role of metadata and cataloguing

The importance of the extreme depth of knowledge required of librarians and cataloguers in rare books and special collections, once identified, naturally leads

to the question: how is this information recorded and communicated? The answer to that question is through the use of metadata.

2.2.1 What is it?

As Sidney Berger points out, “Libraries exist to link patrons with the information they need” (2014: p.9). Once a collection has been developed and the management has begun, a librarian must turn their attention to the process of recording the contents of that collection. This information is recorded through the use of metadata, or as Alison Cullingford defines it, “‘data about data’: information that makes it possible to find and use data” (Cullingford, 2022: p.113) and stored in some form of bibliographic records in the library’s catalogue.

More specifically, the definition given in **Chapter 1** by Zeng explains that metadata is “the structured, encoded data that describe characteristics of information bearing entities (i.e., *things*)” (Zeng, 2016: p.3). Metadata is generally stored digitally, through the use of either formal cataloguing systems or—as in the case of many special collections departments—more informal finding aids that exist outwith a cataloguing system.

2.2.1.1 Types of cataloguing

By far, the overarching style of cataloguing used in the field of special collections is what is known as *descriptive cataloguing*. This is the process of “creating a catalogue record by identifying and recording certain types of data from a book or

other object using a set of standardised rules” (Cullingford, 2022: p.115). In this form of cataloguing, the data that is taken is generally information such as the book’s title, name of the author, and publication statement, amongst others. Generally, in mainstream cataloguing, the records are comprised of metadata that is at what is called the edition level. In other words, this form of cataloguing “uses data common to every copy of that edition” (Cullingford, 2022: p.115).

Edition level cataloguing is not necessarily a sufficient level of detail when it comes to the creation and use of descriptive records for rare books and special collections, however. Records at the edition level do not allow for identification of a specific book from other books with the same title or edition. Books held in a special collections department are by definition unique in some form or another, whether that uniqueness comes in the form of binding, presence of handwritten annotations, or the like. Therefore, librarians must use a level that allows for the amount of description needed to distinguish one item from all others: copy-specific cataloguing (Cullingford, 2022).

Descriptive, copy-specific cataloguing can be invaluable in the field of special collections, especially when one is examining older materials like those in the Stirling Maxwell collection or even earlier. For example, “Biographical information about an author on a title page might be the only biographical information we have about an author, or the only remaining information after dust wrappers have been

discarded” (Attar, 2021: p.426), or including information about any illustrations (whether photograph or drawing) in a record—which Attar notes is “allowed but not mandated unless the information is part of a statement of responsibility” (2021: p.426)—can aid researchers interested in that information.

Another way detailed metadata can be useful is in the process of digitisation, where, for example, “To calculate the complexity of projects and time required, it helps to pinpoint the quantity of folded plates at the outset” (Attar, 2021: p.428). It is clear that this sort of depth in metadata used to catalogue special collections is important for any number of reasons, but so is the structure and framework used to describe these items.

2.2.1.2 Metadata standards

Crockett discusses the importance of using some form of standardised metadata framework, stating that “records creation is about ensuring that records are created consistently with a range of characteristics that make them usable and reliable” (2016: p.44). One of the first, and most widely used, forms of metadata standards is known as the Anglo-American Cataloguing Rules (AACR), developed in 1967 and updated with a second edition (AACR2) in 1978 (AACR, 2005). While AACR and AACR2 were incredibly helpful for the cataloguing field, they have their shortcomings. Because they were developed in the world of the card catalogue, the records are not able to be linked to other records in any meaningful way, especially when used in digital catalogues and other record storage

forms. Therefore, AACR2 was essentially subsumed by a new standard known as Resource Description and Access (RDA) in 2010 (Cullingford, 2022; Committee, 2010). RDA has been noted to have some difficulties regarding the level of description it allows for—some librarians feel that it “was not yet suited to effective copy-specific cataloguing” (Cullingford, 2022: p.177).

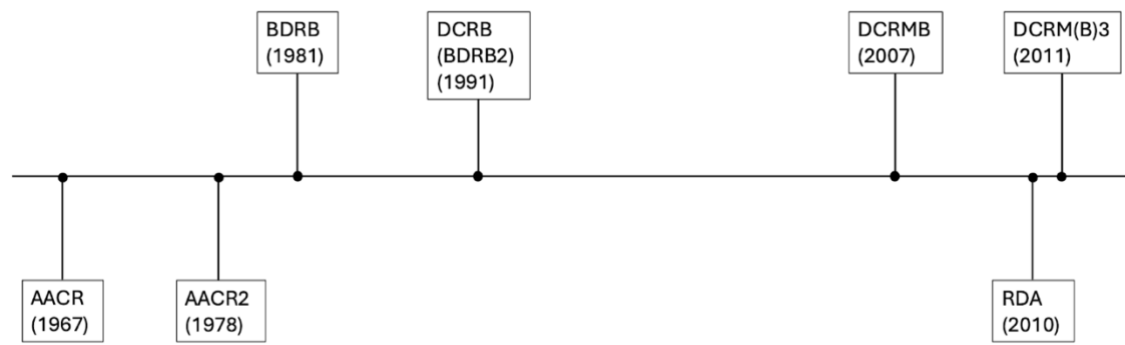


FIGURE 2.1: METADATA STANDARDS CREATION TIMELINE

The timeline for rare books and special collections cataloguing branches off from the publication of AACR2. Following the second revision of AACR in 1978, the Library of Congress in the United States published a set of standards titled Bibliographic Description of Rare Books (BDRB) in 1981 as a supplement to AACR2 dealing specifically with rare books cataloguing (Stalker and Dooley, 1992). Later that decade, “rare book catalogers had accumulated almost ten years’ experience with BDRB and increasingly were voicing a desire to see it updated” (Stalker and Dooley, 1992: p.8). The following revision led to the Descriptive Cataloging of Rare Books (DCRB) as the second edition of BDRB, published in 1991 (Library of Congress and Association of College and Research Libraries, 1991).

The next revision of DCRB became the first iteration of the standard currently generally used in the special collections field, known as the Descriptive Cataloging of Rare Materials (Books) (DCRM(B)), initially published in 2007 as a revision of DCRB (Committee, 2007). DCRM(B), along with its other rare books standards predecessors, was developed based off of the rules developed with AACR2, as its development slightly preceded the development of RDA—there is ongoing work on integrating important RDA principles around record-linking into DCRM as a whole, but this work was not completed in the publication of DCRM(B)3 (Committee, 2011). Other DCRM standards were also developed for other formats of special collections, but the DCRM(B)3—or the third edition of this standard—is the specific standard used in cataloguing rare books and special collections at the Mitchell Library.

2.2.2 DCRM(B)3

The DCRM(B) is defined as “provid[ing] guidelines and instructions for descriptive cataloging of rare books, that is, printed textual monographs receiving special treatment within a repository” (Committee, 2011: p.11).

2.2.2.1 Areas

There are six major areas discussed in the DCRM(B)3, each of which will be briefly defined and explained below, though most are self-explanatory. These areas are:

1. Title and Statement of Responsibility,

2. Edition,
3. Publication, Distribution, Etc.,
4. Physical Description,
5. Series, and
6. Note (Committee, 2011)

Title and Statement of Responsibility

This area deals with the title of the work and the statement of responsibility. The *title proper* is specifically defined as “the chief title of a publication, together with any title information preceding the chief title and any alternative title” (Committee, 2011: p.201). This area is also where further title information such as a parallel title, alternative title, and so on. The statement of responsibility, or the information on the title page regarding the person, persons, or entity responsible for the creation of the work is also part of this area (Committee, 2011). Most commonly, this is the author of the work, but it can also include translators, editors, compilers, etc.

Edition

This area is where information about the edition of the work is recorded. While in mainstream publishing this is usually denoted solely through the use of the specific word “edition,” this is not always the case when dealing with rare books. The DCRM(B)3 thus instructs cataloguers to treat works using the terms

“revision”, “issue”, and even “newly printed” as edition statements that must be recorded in this area (Committee, 2011).

Publication, Distribution, Etc.

This area is where cataloguers are instructed on rules regarding the person, persons, or entity responsible for the publication, distribution, and even manufacture of the item being catalogued. This also lists the publication date of the item in question (Committee, 2011).

Physical Description

This area records the extent, or the pagination, of the item, as well as records the presence of illustrations (of any sort), the size and format of the item, and if there are any additional materials accompanying the main body of the work (Committee, 2011).

Series

Any information regarding a series an item may belong to is recorded in this area. This is generally where volume numbering is recorded, if an item is part of a multi-volume publication. If there is a separate series title or statement of responsibility for the series as a whole, that information is recorded here rather than in Area One (Committee, 2011).

Note

This is by far the most expansive area within DCRM(B)3, as it essentially encompasses whatever information the cataloguer deems necessary to record that does not fall within any of the other areas. Specifically, 7A1.1 states that “Notes qualify and amplify the formal description, and are especially important for recording types of information not accounted for in other areas of the description. Notes can therefore deal with any aspect of the publication” (Committee, 2011: p.127).

Subject and genre/form headings

One part of the Note Area that requires expounding upon is the use of subject headings and genre/form headings. While these are not specifically mentioned in the main body of DCRM(B)3, their inclusion is implied in the definition of the Note Area. They *are* mentioned in the appendices detailing the application of DCRM(B)3 standards for machine-readable cataloguing (MARC) formats (Committee, 2011).

Subject headings deal with the content of the material being catalogued. There are many forms of vocabularies from which to draw subject headings, but Cullingford notes specifically that the Library of Congress Subject Headings (LCSH) list is “very widely used” (2022: p.129). This wide usage is exemplified in the records contained in the English Short Title Catalogue (ESTC), or, “the most comprehensive resource available for the output of the printing press in the English-speaking world before 1801” (*English Short Title Catalogue (ESTC)*, 2025), as noted by Garrett

(2007). The development of the LCSH began in the latter half of the 19th century as a compilation of several other subject lists, and after the first list was published at the beginning of the 20th century, work has been continuously performed to ensure the list is current and up to date (*Subject and Genre/Form Headings*, n.d.). Due to its wide use and the vast coverage it allows for, LCSH are then the most common forms of subject notation used in special collections libraries when subject notation is used.

Cullingford does note that while “some special collections use such schemes for shelfmarking, especially if collections contain books transferred from their main library collection” (2022: p.131), this is not necessarily a requirement—“closed access material does not need to be browsable” (2022: p.131). Additionally, especially for the oldest of materials stored in special collections, standard vocabulary lists like LCSH do not allow for texts where the subject terminology differs so vastly from the modern-day terminology used in LCSH (Cullingford, 2022). For libraries like the Mitchell Library however, with a main holding of mainstream items with the addition of a special collections department, the use of LCSH for both mainstream items and those held in special collections can provide for a standardised and streamlined process for the cataloguing department to follow.

Other sets of controlled vocabularies include those created by the Rare Books and Manuscripts Section (RBMS) to record information about Paper, Printing and

Publishing Evidence, Type Evidence, Genre, Binding, and Provenance Evidence

(Berger, 2014). The one of most use regarding the subject of a material is the Genre thesaurus. This thesaurus is not a required inclusion for item records in a catalogue, but its use does provide for a deeper description of the item in question (Committee, 2014).

Given the Mitchell Library’s position as a public library with a special collections department, the use of LCSH for subject matter notation has been chosen for use in the remainder of this project.

2.2.3.2 DCRM(B)3 encoding in MARC

In order for a DCRM(B)3 record to be encoded digitally in most forms, the translation of the DCRM(B)3 areas to MARC is necessary. A table laying out this translation for a core-level record is below:

MARC Encoding			DCRM(B)3 Area
Field	Subfield or Character Position (if applicable)	Instructions for Field/Subfield/ Character Position (if applicable)	
Leader	17—encoding level	Enter “4” to indicate a core-level record	
008—Control field	06—type of date/publication status	Code these as appropriate	
	07-10—date 1		
	11-14—date 2		
	15-17—place of publication, production or execution		
	22—target audience		
	23—form of item		

	28—government publication		
	33—literary form		
	34—biography		
	35-37—language		
	38—modified record code		
	39—cataloguing source		
010—Library of Congress control number (LCCN)		Must code if applicable to the item	
020—International standard book number (ISBN)		Must code if applicable to the item	
040—Cataloguing source	\$e	If Leader/17 is encoded with a “4”, this must be coded with “dcrmb”	
042—Authentication code		If the records are contributed by members of the PCC BIBCO program, this must be coded with “pcc”. All other records must leave this field blank	
050, 082, 086, 090, etc.—Call numbers		Not mandatory, but strongly encouraged to include at least one of these	
1XX—Main entry		Must code if applicable	
245—Title statement	\$a	Must code	Title and Statement of Responsibility
	Remaining subfields	Must code if applicable	
246—Varying form of title		Must code if applicable	
250—Edition statement		Must code if applicable	Edition
260—Publication, distribution, etc. (Imprint)	\$a—Place of publication, distribution, etc.	Must code if applicable. If data is unavailable, code with “[s.l.]”	Publication, Distribution, etc.
	\$b—Name of publisher, distributor, etc.	Must code if applicable. If data is unavailable, code with “[s.n.]”	
	\$c—Date of publication, distribution, etc.	Must code	
300—Physical description	\$a—Extent	Must code	Physical Description
	\$c—Dimensions	Must code	
	Remaining subfields	Must code if applicable	

4XX—Series statements		Must code if applicable	Series
500—Source of title proper (if other than chief source)		Must code if applicable	Note
501—With note		Must code if applicable	
502—Dissertation note		Must code if applicable	
505—Formatted contents note		Must code if applicable	
510—Citation/references note		Must code if applicable	
6XX—Subject headings		Must code if applicable	
655—Genre/form headings	\$a—Genre/form data or focus term	Mandatory if this field is being used	
	\$2—Source of term	Mandatory if this field is being used	
7XX—Added entries		Must code if applicable	
8XX—Series added entries		Must code if applicable	

TABLE 2.2: ENCODING DCRM(B)3 RECORDS IN MARC (COMMITTEE, 2011)

With this understanding of the information that needs to be recorded for these special collections holdings—and what defines a special collections item—and an overview of commonly-used metadata standards for rare books and special collections, it is now time to move on to the discussion of defining the users involved with these collections through a user needs analysis.

2.3 User needs analysis

2.3.1 What is it?

Knowing a library’s users is a key tenet of librarianship, especially as it comes to identifying what items a librarian will curate as part of their collections and how

those holdings are managed over time. Knowing their users will allow librarians to ensure that a given collection fits with the library's mission and/or vision statement, so it stands to reason that an analysis of the user groups they serve is also a necessary step. Evaluation, therefore, should not be limited solely to the collection itself (as discussed previously through defining what sets rare books and special collections apart), but should also extend to the current and prospective users of the collection in question to ensure a thorough and appropriate collection management. Understanding the way users may interact with the collection(s) under a department's purview will inform how the collection(s) must be managed to be most effective for those the collection(s) serves. This section discusses the choice of user groups analysed and a brief overview of current literature regarding these user groups' interaction with collections in a similar position as the Stirling Maxwell donation.

2.3.1.1 User group selection

The first step of conducting a user needs analysis is to define what a "user" is for the purposes of the analysis at hand.

Who are the potential user groups in question?

For this section, Alison Cullingford provides an excellent introductory breakdown of who can be considered a user of a special collections library.

The first thing to consider when defining a user group is their relationship to the library—are they *internal* or *external* users? Internal users are users who are members of staff at the library (both paid and volunteer), which comprise “A finite, known population, usually (not always) on-site” (Cullingford, 2022: p.188). External users are those who are part of the general public. Therefore, this population will most likely be significantly larger than the population of internal users (Cullingford, 2022).

As the finding aid being developed through this study will be used by both external and internal user groups at the Mitchell Library, two user groups must be selected for analysis in an attempt to encompass the variety of information needs and information-seeking behaviours involved in this process: one internal and one external. The internal group is self-explanatory based on the definition given by Cullingford—this user group will consist of the librarians in the special collections department. The external user group requires further discussion before selection.

External users, by definition, consist of a nearly infinite number of distinct user sub-groups, each with their own searching patterns and information needs that may need to be addressed in different manners. Cullingford provides a list of example user sub-group types, some of which are summarised below:

- Academic staff,
- Postgraduate students,

- Undergraduate students,
- Independent scholars,
- Family and personal historians,
- Local historians,
- Media or commercial researchers, and
- Tourists (Cullingford, 2022: p.188-9).

What must be considered for each external user sub-group in order to best provide services as a special collections librarian, as each sub-group will differ in large or small ways from each other, are plainly laid out as:

- Motivation for using Special Collections
- Experience of using such services
- Interest in research as a process versus need for an answer
- Timescale: how quickly they need an answer
- Relationship with the service, from a single encounter to lifelong engagement
- Emotional involvement with the collections
- Level of skills needed to access collections, including research techniques, digital skills, languages and palaeography (Cullingford, 2022: p.189).

These different user groups and some examples of their skills and expectations can also be seen in **Table 2.3** below, created by Crockett:

User/Researcher Group	Perspectives
Academics	<ul style="list-style-type: none"> • Usually highly educated, organized, focused • May be experienced archive users • Often expect archives to contain material relevant to their research and be organized as they want
Genealogists	<ul style="list-style-type: none"> • Could have any level of education, organizational skills and experience of archives • May be easily satisfied if lots of personal name indexes—or disappointed if not • New genealogists need coaching
Journalists, film and TV producers	<ul style="list-style-type: none"> • Confident and empowered • Expect records of subjects interested in to be there and organized in ways they imagine • Don't distinguish between archives, books and artefacts • May expect special treatment without offering remuneration for services received • Can be experienced researchers
Schoolchildren	<ul style="list-style-type: none"> • Should come with teacher or parent • Don't know (or need to know) complexities of archival description • Probably need resources based on school curriculum • Offers opportunity to introduce them positively to archives
Students	<ul style="list-style-type: none"> • Could be highly educated, organized, focused, depending on level and reason doing research • Can be inexperienced or more experienced • May be interested in course-related materials • May need lots of help if beginning research
Freedom of Information/Open Data/Big Data researchers	<ul style="list-style-type: none"> • Ordinary citizens • Very varied in terms of education and experience of using archives • Focused on particular information in archives/records not records in context • May need to think laterally to find what they are looking for
Personal history researchers (e.g. for adoption details)	<ul style="list-style-type: none"> • Probably inexperienced in archive use • Probably anxious or stressed • Interested in names, institutional records • Need sensitive and/or appropriately qualified support
Local historians	<ul style="list-style-type: none"> • Can be very knowledgeable and know archives better than archivists • May have own collections of archive materials

	<ul style="list-style-type: none"> • May expect archives to be organized like a local studies library
House History	<ul style="list-style-type: none"> • Very focused on own house • Probably not experienced in using archives • May be frustrated at lack of archival resources

TABLE 2.3: ARCHIVE USERS AND THEIR PERSPECTIVES (CROCKETT, 2016: P.39)

While it is easy to see that there can be any number of user groups that may interact with a special collections department like that at the Mitchell Library, examining *all* of these various user groups is outside the scope of this project due to the overly broad lens that would have to be used. Normally in a real-world scenario, “we don’t need a detailed understanding of the various sub-groups because, whilst we want to help them use the archives, we also want the archives to be an independent testimony to the history of the organization or individual” (Crockett, 2016: p.37). However, this project requires a much narrower focus due to the single collection being examined. Therefore, a single external user group was chosen for this project, comprised of historians who may be accessing the items in the department or in the collection itself for historical research, whether that is local history or general history. Henceforth, this user group is simply referred to interchangeably as “historians” or “patrons” for ease of communication. There will be no distinction in this project between historians acting in a professional manner and those merely conducting personal historical research, as this factor has no functional impact on their behaviours when accessing collections like the Stirling Maxwell donation within the scope of this project.

2.3.1.2 Information seeking behaviours

Now that the two user groups have been selected, the next step of this analysis is to examine the literature around these groups' ISBs. This part of the process will help inform the librarian as to how best to structure the overall management framework, including the choice of metadata design, in order to better serve the users at hand.

Historians

A brief overview of select studies of the ISBs of historians has revealed a set of ISBs that could be used to inform the patron-centred portion of the metadata framework design and evaluation process. These studies—which are discussed in detail below—build off of the foundation laid by David Ellis (1993). To fully appreciate the basis for these ISBs, they must be examined from their first instance and tracked through time as they shift, change, and in some ways remain the same.

Beginning with Ellis' grounded theory work in building “models of the information-seeking patterns of academic researchers” (Ellis, 1993: p.469), six basic information-seeking features were noted for the particular population of academic researchers in the social sciences:

- Starting,
- Chaining,

- Differentiating,
- Monitoring, and
- Extracting (Ellis, 1993).

Ellis compares this to the information seeking behaviours of English literature researchers, where the categories of browsing, differentiating, and extracting were replaced with the following three:

- Surveying,
- Selection and sifting, and
- Assembly and dissemination (Ellis, 1993).

In 2003, Meho and Tibbo built off of Ellis' previous work on ISBs in social science academic researchers and developed a newer model of the behaviours of social scientists. They found that the six social scientist behaviours previously identified remained present in their sample population, and added another four behaviours to Ellis' six:

- Accessing,
- Verifying,
- Networking, and
- Information managing (Meho and Tibbo, 2003).

An additional change this study brought about is in the creation of what they call "four interrelated stages" (Meho and Tibbo, 2003: p.570) of information seeking

that the ten behaviours fall into, naming them the searching, accessing, processing, and ending stages. They then developed the model shown in **Figure 2.2** below:

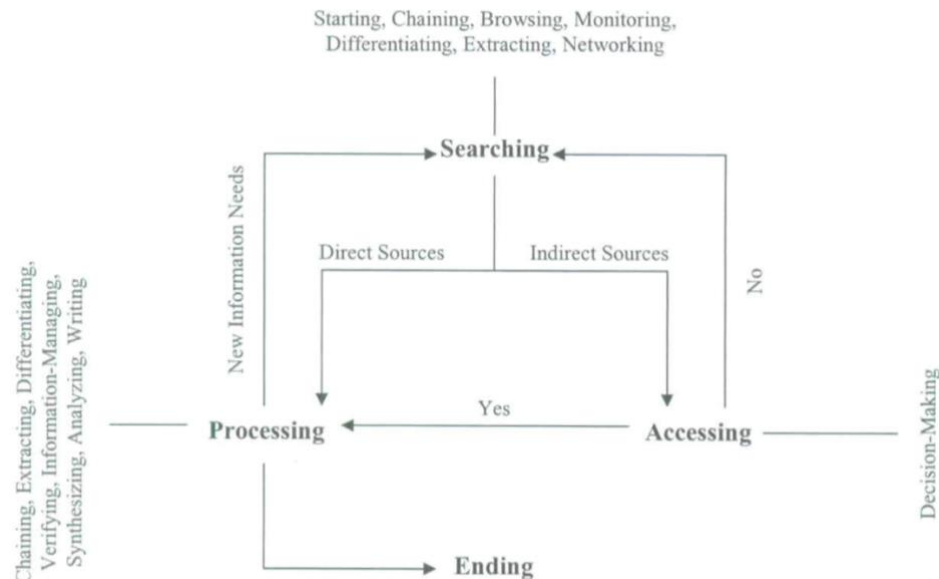


FIGURE 2.2: MEHO AND TIBBO’S MODEL OF INFORMATION-SEEKING BEHAVIOUR OF ACADEMIC SOCIAL SCIENTISTS (MEHO AND TIBBO, 2003: P.584)

In a study initially described in 2010 and fully published in 2012 comparing ISBs between historians and social scientists, H.L. Rhee revealed a list of behaviours that could be used to understand the information-seeking needs of historians. This list builds off of the ten generic features and four stages modelled by Meho and Tibbo in the above figure by adding an additional three behaviours:

- Starting,
- Extracting, and
- Verifying (Rhee, 2010; Rhee, 2012)

In the final study, Rhee developed the following model for the information-seeking behaviour of historians as an adaptation of the model shown in **Figure 2.3**:

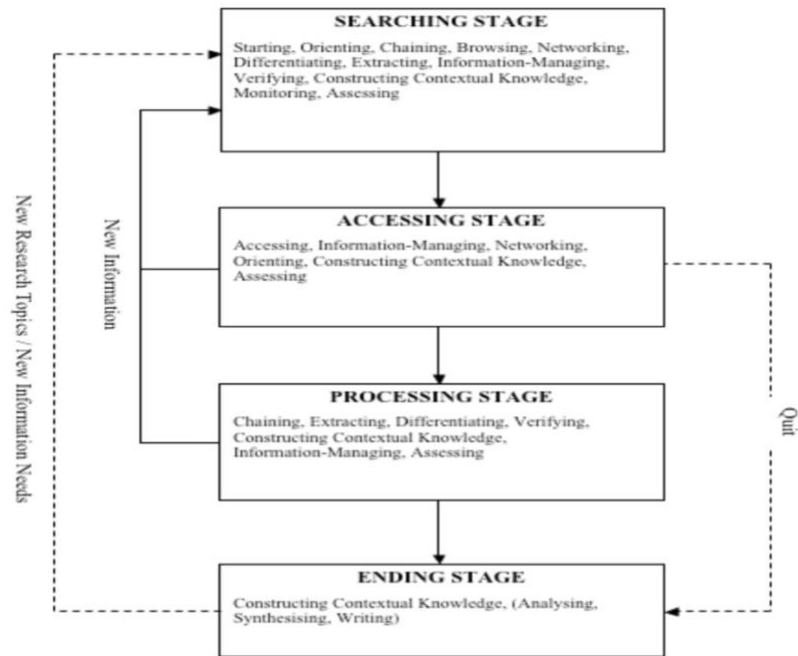


FIGURE 2.3: RHEE’S MODEL OF THE INFORMATION-SEEKING BEHAVIOUR OF HISTORIANS (RHEE, 2012)

Comparing the two models, Rhee points out that the four stages are the same as Meho and Tibbo’s model, but there are some significant differences as one moves from the first model to the second, notably:

- 1) the connections between interrelated stages; 2) features in each stage; 3) three new features that were included in [Rhee]’s model; and 4) analysing, synthesising, and writing activities considered to be information-using activities rather than information-seeking activities as in Meho and Tibbo’s model (Rhee, 2012).

Importantly, the specific definitions of each feature Rhee's model shares with Meho and Tibbo's do not change in the adapted model, only the way Rhee classifies their stages given the narrowing of the user type from academic social scientists to historians.

An earlier study by W.M. Duff and C.A. Johnson (2002) that used a more interdisciplinary method to examine the ISBs of historians specifically using archives, however, revealed *four* distinct features historians exhibited during their use of archives compared to the ten features identified by Meho and Tibbo and the three identified by Rhee. These four behaviours are:

(1) orienting oneself to, or becoming aware of the archives, the finding aids, the sources, or a collection; (2) seeking known material, including known items, known forms, or known collections; (3) building contextual knowledge; and (4) identifying relevant material (Duff and Johnson, 2002: p.480).

While Duff and Johnson call these behaviours "information-seeking activities" (2002: p.480), their meaning is in actuality closer to that akin to the term *stages* used initially by Meho and Tibbo, followed by Rhee. This means that all three studies identified the same four groupings of ISBs—herein, these groupings will be referred to as stages.

The development of the social scientist ISBs over time is visually displayed in **Table**

2.4 below:

ISB Name	Is it present in the study by...		
	Ellis (1993)	Meho and Tibbo (2003)	Rhee (2010, 2012)
Starting	Yes	Yes	Yes
Chaining	Yes	Yes	Yes
Browsing	Yes	Yes	Yes
Differentiating	Yes	Yes	Yes
Monitoring	Yes	Yes	Yes
Extracting	Yes	Yes	Yes
Accessing	No	Yes	Yes
Verifying	No	Yes	Yes
Networking	No	Yes	Yes
Information managing	No	Yes	Yes
Orienting	No	No	Yes
Constructing contextual knowledge	No	No	Yes
Assessing	No	No	Yes

TABLE 2.4: COMPARING INCLUSION OF SOCIAL SCIENTIST ISBs BETWEEN STUDIES

Important to note here is that while the ISBs identified by Ellis in the population of English literature researchers use different terminology than the social science ISBs, they are not included in **Table 2.4** because Ellis specifies that, “although the models differ in terminology and detail, when the features and characteristics of the models are compared they can be seen to be representing fundamentally the same activities” (1993: p.483). As such, the English literature terminology is not used moving forward.

Below is the side-by-side list of the four stages identified by Meho and Tibbo (also present in Rhee) and the four identified by Duff and Johnson:

Duff and Johnson (2002)	Meho and Tibbo (2003)
Orienting	Searching
Searching for known material	Accessing
Building contextual knowledge	Processing

Identifying relevant material	Ending
-------------------------------	--------

TABLE 2.5: LISTING INFORMATION-SEEKING STAGES NOTED IN TWO STUDIES BESIDE EACH OTHER

All of these studies make specific note that their four stages do not necessarily occur in any particular order—rather, they can not only happen at any point in the information seeker’s journey, but even overlap with other stages.

Defining behaviours

Now that the behaviours and stages have been identified, they must be defined. Rather than explain the definitions by the researcher responsible for identifying them, they are instead explained using the combination of definitions given by the various researchers per behaviour. This approach better allows for an understanding of their development over time. Important to note is that while the sheer number of behaviours may prove daunting to a librarian attempting to conduct a user group evaluation, these behaviours can in fact be quickly explained and understood.

Starting

Beginning in 1993, Ellis defined this behaviour as being comprised of “activities characteristic of the initial search for information” (Ellis, 1993: p.482). In his study of social scientists, he noted that some of these activities included the “identification of a key paper to commence the search” (Ellis, 1993: p.480), or “the employment of an online search to locate references” (Ellis, 1993: p.481). This definition is confirmed by Meho and Tibbo, who add that, “Although there were

some variations in starting methods among the participants, in general, when they initiate a new project, starting is usually carried out through a combination of searching and communication activities” (2003: p.579).

Rhee defines starting for historians in a similar way to Meho and Tibbo, saying, “Historians, like social scientists, begin the initial search for information related to new research topics and relevant sources by using both traditional tools (e.g., indexes and abstracts) and electronic tools (e.g., online finding aids and OPACs)” (Rhee, 2012). The reliance on traditional physical sources is greater in historians than in social scientists, however (Rhee, 2012). Part of this could be pure preference, but Rhee does note that “it is not clear whether historians do not frequently use electronic tools because...such tools are not available to them” (2012).

Chaining

The first definition of this behaviour comes from Ellis, who explains it as meaning that these researchers “follow[ed] chains of citations or other forms of referential connection between material” (Ellis, 1993: p.482). One of these other forms referred to by Ellis is clarified by Meho and Tibbo as “following references often obtained through reading and personal contacts” (2003: p.579). Rhee agrees with the definitions put forth previously, adding that along with using these citation chains as additional sources of information, the chains themselves can be used as a source of information in itself. These citation chains can be used “as markers of

intellectual history that chart changing research in a subject area and changing researchers' perspectives and views" (Rhee, 2012).

Duff and Johnson also call this behaviour *footnote tracing*, saying that it "involves locating key articles or books and following the references cited to other articles or books" (2002: p.475). Both Duff and Johnson as well as Rhee noted that this behaviour is among the most popular methods used by historians.

Browsing

Ellis explains browsing as the process of "semi-directed searching in an area of potential interest" (Ellis, 1993: p.482). In a physical library, this behaviour often displays as someone casually scanning over books stored in a particular area. On the other hand, it can manifest in an online search through heavy use of the subject search as the sole information filter. Meho and Tibbo agree with this description, adding that they were able to further specify two methods of browsing: "(1) the scanning of recently published issues of journals and tables of contents of relevant books; and (2) browsing the online catalog, indexes and abstracts, web resources, and references of materials found and/or read" (2003: p.580).

This is a behaviour Rhee notes as particularly prevalent in historians, stating, "Historians intentionally browse bookshelves and internet sources, trusting in serendipity" (2010: p.2). Serendipitous browsing—referring to the intentional use of browsing various forms of physical and digital search results with the hope of

finding relevant material—is not the only form of browsing used by historians, but there is evidence that it is one of the most used methods. In fact, Rhee notes that this supports an earlier finding “that most historians prefer high recall to high precision” (Orbach, 1991, cited in Rhee, 2012), or “the assumption that historians want to browse lots of search results in order to accidentally find unexpected but relevant sources that might be missed otherwise” (Rhee, 2012).

An important factor in the prevalence of serendipitous browsing that Rhee points out is the presence of poorly built or maintained finding aids and archive descriptions, resulting in problems locating one specific source. This results in a need to examine all materials held in a collection to more accurately judge what, if any, items may be relevant to their needs (Rhee, 2012).

Differentiating

Ellis defines differentiating as “using differences between sources as a filter on the nature and quality of the material examined” (1993: p.482). This could be as simple as distinguishing a book from a journal as a source, choosing between material published by two different publishers, etc. A more detailed definition is given by Meho and Tibbo as “when information sources are evaluated or judged according to their nature, quality, relative importance, and usefulness as a way of filtering the amount and nature of information obtained based on the participants’ own perceptions” (2003: p.581).

While historians use some of the same methods of differentiation that the social scientists studied by Ellis and by Meho and Tibbo, according to Rhee they also use an additional set of criteria. Historians make inclusion/exclusion decisions about materials using “evidential value and format of information sources and materials” (Rhee, 2012). Evidential value is important to historians because they “differentiate and select information sources and materials that they can use as evidence in historical argument...because the nature of their discipline is to pursue facts” (Rhee, 2012). In terms of format, historians differ again from social scientists in that they vastly prefer the original of the information source rather than a copy, because of “errors during reproduction and bad quality of reproduction results” (Rhee, 2012).

Monitoring

This behaviour is defined as the process of researchers “maintaining awareness of developments in a field through the monitoring of particular sources” (Ellis, 1993: p.482), usually through the use of subscription to regularly published journal issues or, as the use of online databases increases, a tactic such as a Google alert. Again, Meho and Tibbo agree with this definition and add on the further clarification that “both formal and informal information channels are used for keeping up-to-date” (2003: p.580).

In defining monitoring, Rhee refers to the similar need for historians to watch for new arrivals to the archives or collections they are using for their projects. The

informal methods, such as personal connections with archivists, seem to be the most useful for historians, again because of the general use of inadequate archive descriptions: “Collection-level archival description and finding aids seem to give historians problems in monitoring new arrivals in archives, particularly those inserted into existing collections” (Rhee, 2012). This is in comparison to social scientists, where the formal methods often prove of significant use for monitoring.

Extracting

Ellis explains that extracting is “systematically working through a particular source to locate material of interest” (1993: p.482). These sources can be direct or indirect, with direct sources being items like books or articles, and indirect sources being those like bibliographies or abstracts, among others (Meho and Tibbo, 2003). The only addition Rhee has to this definition is merely that “historians try to extract only relevant information and avoid the temptation to focus exclusively on research topics that, while fascinating, have little relevance to their work” (2012).

Accessing

Meho and Tibbo added this stage to the list previously identified by Ellis because, “For the information-seeking process to continue, however, researchers need to get hold of, or access, the materials or sources of information they identified and located” (2003: p.581). They noted the importance of this stage in informing the rest of the social scientists’ information-seeking process as a reason for including it as an ISB, explaining that “Because time, funding requirements, and

governmental restrictions for accessing some relevant materials are great barriers for many researchers, scholars sometimes resort to alternative sources or methods for obtaining information, often using secondary sources” (Meho and Tibbo, 2003: p.581).

Similarly, accessing is a term used by Rhee (2012) specifically in the physical sense (ability to access materials) and its impact on historians, whether positive (meaning continued research) or negative (meaning either an end of a project or a pre-emptive exclusionary decision). Historians tend to present this behaviour more often than social scientists, typically due to the singular nature of the sources they prefer to use compared to those used by social scientists. Remembering that historians vastly prefer to use the original (usually physical) form of a source—which Rhee points out is often “unique, dispersed, or classified” (2012), they end up dealing with far more access barriers. These barriers can have a much larger effect on a historical research project compared to a social science project: Duff and Johnson in particular discuss how the inability to access a source for whatever reason can lead historians to quit their project entirely (2002).

Verifying

Meho and Tibbo define this behaviour as being “characterized by activities associated with checking the accuracy of the information found” (2003: p.582). This behaviour appeared for Meho and Tibbo because their sample

population of social scientists was comprised of those researching stateless nations, which has always been a particularly sensitive subject. It was then important for these researchers to verify the information they gleaned through multiple methods—formal and informal, as in monitoring—to account for “‘bias,’ ‘disinformation,’ and lack of ‘reliability’ and ‘accuracy’ of many sources of information...especially among materials produced by ethnic and governmental organizations as well as those published on the Web” (Meho and Tibbo, 2003: p.582).

While Rhee agrees with the prior two studies on the definition of this term, there are some differences when it comes to the reasoning behind historians’ need to verify their sources. While Meho and Tibbo seem to link the need for verification to the sensitive nature of the research being conducted—meaning that researchers *not* investigating similarly sensitive topics in the social sciences may then not exhibit this behaviour—Rhee points out that the need for verification appears baked into the discipline of historical research. According to Rhee, “historians seem to learn verifying activities according to the conventions of history education” (2012). This behaviour then appears to be a permanent feature of the information-seeking behaviours of historians.

Networking

For the social scientists they worked with, Meho and Tibbo defined networking as being “characterized by activities associated with communicating, and

maintaining a close relationship, with a broad range of people” (2003: p.582). In many cases, these researchers built these networks not just for gleaning information, but for sharing information as well (Meho and Tibbo, 2003).

Networking involves the idea of historians building relationships both amongst other historians, as well as with the archivists in charge of the materials they are using or wish to use—Rhee notes that this behaviour differs notably in historians than social scientists in that, “Historians perform networking throughout the entire historical research process; social scientists perform networking only in the *searching* stage...” (2010: p.2). Additionally, historians rely heavily on connections with “special collection librarians, librarians with expertise in a specific field (e.g., art history), and especially archivists” (Rhee, 2012). The activities performed with the networks are the same as those formed in social science research, the difference lies mainly in the job descriptions of those being involved in historical networking.

Information managing

This behaviour is added by Meho and Tibbo again due to their participants “repeatedly talk[ing] about the need and importance of filing, archiving, and organizing the information they collect or use in facilitating their research” (2003: p.582). They noted that the information used by the population in their study was often not readily available or applicable, but that “it needs to be gathered, digested, organized, and stored for future use” (Meho and Tibbo, 2003: p.583).

Rhee defines this as the process by which historians store and process the information they glean from the sources they use. Again, Rhee notes that historians specifically “begin to file and/or organize collected materials in the *searching* stage and/or in the *accessing* stage even though they cannot finish systematically organizing all collected materials before the *processing* stage” (2010: p.2). Similarly to the reasoning behind why historians place such importance on the verification of sources, the process of managing information seems to be a baked-in feature of the entire historical discipline, as “basic research methodology textbooks for historians tend to detail note-taking and information-managing and highlight their importance” (Rhee, 2012). More is known about the information-managing process of physical notes, with Rhee stating that “historians often store these materials in their offices, often in some kind of card file to index collected materials, and organise them, primarily by topic” (2012), while research is still being conducted on this process when it is conducted digitally.

Orienting

Orienting is a behaviour that is present in historians but not in social scientists. Rhee defines this feature as being “characterized by activities associated with becoming informed about and acquainted with information repositories (e.g., archives and libraries), including their holdings and access tools” (2012). This is then considered a separate feature than starting, particularly

because this process is so vital to the field. While social scientists may tend to have a narrower field of resource databases they use for each research project, “historians tend to use multiple archives and libraries and various types of materials” (Rhee, 2012), requiring more work in orienting themselves. This behaviour—while separated from it—can occur alongside the *starting* feature, but due to its higher prevalence in the historians examined, it is made clear it is not a sub-feature of *starting*.

Duff and Johnson explain one method historians use for orienting themselves to collections is examining finding aids (either print or digital in form), “...to get a sense of the whole, or ‘geography’ (participant 4) of the collection” (Duff and Johnson, 2002: p.481). This form of orientation method, and orientation in general, is a behaviour used to “reduce uncertainty when visiting a new archive or starting to look at a new collection” (Duff and Johnson, 2002: p.481).

When viewing the research process of an individual historian as a whole, while orientation can be most commonly thought of as the beginning of the process, in truth, “Orientation to new archives, finding aids, sources, or collections are activities that occur throughout the research process” (Duff and Johnson, 2002: p.480). This links with Rhee’s (2012) explanation of the behaviour as a separate process from starting being partially due to the higher tendency to use multiple archives and finding aids, each with their own organisation method requiring learning.

Constructing contextual knowledge

Again, this is a process noted by Rhee as an inherent feature of the historical research field, noting that they “construct contextual knowledge to explain history” (2012). This information is used to help them in interpreting and understanding the resources they access, meaning that the process of gaining it involves examining sources outwith the particular source they are attempting to understand. In some cases, Rhee notes that this can even extend to the point of “many historians also visit historic places and investigate artifacts that are related to their research projects” (2012). Contextual knowledge is therefore not a feature identified in either of the social science studies mentioned previously, as this level of holistic examination is not needed to understand any single source as it is in historical research.

Assessing

This feature is one not explicitly noted by Meho and Tibbo, but Rhee points out that it appears to be implicitly referenced as “underlying a particularly important mode of decision-making” (2012). Neither term is included in the Meho and Tibbo model shown in **Figure 2.1** and listed as information-seeking behaviours of social scientists. By contrast, the definition chosen by Rhee—“evaluating the information found, such as for quantity and internal quality, and determining the significance or value of collected information sources and materials to the overall project” (2012)—is shown as an integral part of the historical research process. At every

stage of the research project, *assessing* takes place to determine whether the historian remains in the current stage (i.e., continues searching for more sources in the searching stage) or proceeds to the next (i.e., verifying sources in the processing stage). As such, this behaviour appears at every stage of the model shown in **Figure 2.2**.

Librarians

Compared to the range of literature discussed regarding the information-seeking behaviours of historians, there is remarkably little information about the ISBs of special collections librarians of the type relevant to this project. Literature exists for special collections librarians such as medical and law librarians, but these types of collections differ from the one being studied herein in both the subject matter as well as the access level required. Oftentimes medical and law libraries are limited to specific members of the group the libraries serve—this could mean medical students at a specific university or medical professionals at a particular hospital if discussing a medical library, or law students at the university hosting the law library—which brings with it a significantly different set of access expectations for librarians.

This difference can most clearly be seen by returning to Cullingford's discussion of what special collections librarians must consider when addressing the needs of various external user sub-groups:

- Motivation for using Special Collections
- Experience of using such services
- Interest in research as a process versus need for an answer
- Timescale: how quickly they need an answer
- Relationship with the service, from a single encounter to lifelong engagement
- Emotional involvement with the collections
- Level of skills needed to access collections, including research techniques, digital skills, languages and palaeography (Cullingford, 2022).

The answers to these questions can differ significantly when comparing users like medical and law students/professionals accessing medical or law libraries to users like those most likely to access the historic collections held at the Mitchell Library. The clearest difference between the two types when addressing the questions posed by Cullingford is seen in the discussion about “emotional involvement with the collections” (Cullingford, 2022: p. 189). For the first type of special collections (medical or law libraries), the emotional involvement on average tends to be low, while the opposite is true when dealing with the type of historical special collections like the Stirling Maxwell collection.

Crockett (2016) compounds the example of these differences in **Table 2.3**, where she lays out a list of nine separate sub-groups of archive users and the perspectives they may be coming from when accessing archive materials. While a

special collections department like the one holding the Stirling Maxwell collection at the Mitchell Library must provide access and aid to all nine of these user sub-types, medical or law libraries in the situations posited previously may only provide access and aid to three of these user sub-types:

- Academics
- Students
- Practicing professionals⁴

This smaller group of user sub-types brings with it a smaller list of access needs that must be considered by librarians managing collections. All of this information together leads to the conclusion that literature regarding the ISBs of librarians in charge of managing collections in highly specialized archives like medical or lab libraries cannot be used in lieu of the missing literature researching the ISBs of librarians in charge of managing less specialized archives such as the special collections at the Mitchell Library used in this case. The specific impact of this gap in literature will be discussed in further detail in the following chapters.

Following these sections defining what makes a rare book and the analysis of user ISBs and information needs for the two user groups specified for this project, it is time to turn to how these identifying features and their metadata can be combined

⁴ The term “practicing professionals” in this usage is replacing the user sub-type “local historians” in the original nine given by Crockett (2016) and is defined as professionals currently practicing their discipline in the field (i.e., surgeons, doctors, nurses, lawyers, and paralegals, to name a few).

for the purpose of best serving the identified user groups and their needs. This will be done through a discussion of collection management and collection management policies.

2.4 Collection management policies

2.4.1 What is collection management?

Once a library has a collection—whether they were the one to build it or merely acquired it—they cannot leave the collection as-is into perpetuity. If left alone, libraries run the very serious risk of becoming inaccessible buildings full of shelves overflowing with books in various stages of disrepair and relevance to the people they serve. Collections must be managed—and carefully so—to avoid a library spiralling out of control. This section discusses collection management and its definition with the aim of laying the basis for an understanding of collection management policy before diving into metadata creation and its role as a subset of collection management.

2.4.1.1 Development or management?

It is crucial that the terms “collection management” and “collection development” are defined as clearly as possible now to clarify the distinct (and often confusing) differences between the two as well as lay out the reasons why this paper and this project will be speaking more specifically about the former term rather than the latter.

Clayton and Gorman provide the most thorough analysis of the two terms and their distinctions, noting that while the two terms are sometimes treated as two distinct processes, it must be recognized that there is still a fair amount of overlap between them, and that even “In some of the most widely used literature in this field there seems to be considerable confusion about the distinction between the two concepts” (2006: p.17). While it seems easy for the complicated definitions of collection management and collection development to be distilled into simple terms such as Matassa’s “know what you have and where to find it” (Matassa, 2011), Clayton and Gorman argue that this level of simplification completely discards collection development in favour of the idea that collection management has not only replaced collection development as the “approved professional term” (2006: p.16), but also as the *only* process that takes place—a fact they say is untrue, stating, “in fact the two components continue to exist” (2006: p.17). Hibner and Kelly follow along with Clayton and Gorman’s understanding of the two terms, stating not only that the latter term only covers the beginning of the collection life cycle, but also, “The word ‘development’ implies growth, which points to the selection piece of the life cycle⁵, while the word ‘management’ implies control and attention” (Hibner and Kelly, 2013: p.2). While Matassa’s

⁵ The term *collection life cycle* is one given by Hibner and Kelly as one that describes a library as a living entity that is constantly changing and evolving. Their life cycle is comprised of eight stages: “selection, acquisitions, processing and cataloguing, shelving, checkout (use), re-shelving, repair/maintenance, and weed or replace” (Hibner and Kelly, 2013: p.1). The definitions for each of these stages may vary based on the particular field the library covers, but a couple will be explained further along in this paper as they may prove confusing the most often.

definition is not incorrect (and neither are the simple, common-sense definitions of the terms—collection development being about the building of a collection and collection management being the continued care of an already-built collection), the field of collection development is better understood as both a part of the process of collection management as well as a field capable of standing on its own. Specifically:

In our view collection management may be driven in part by collection development needs, but at the policy level a collection development policy clearly resides within the broader objectives of a collection management policy. Collection development, in other words, is a subset of collection management, and both have a clearly defined policy role (Clayton and Gorman, 2006: p.17).

Drawing on work done by Cogswell, Clayton and Gorman therefore choose to define the term *collection management policy* as “the systematic management of the planning, composition, funding, evaluation and use of library collections over extended periods of time, in order to meet specific institutional objectives” (Cogswell, 1987, quoted in Clayton and Gorman, 2006: p.17), adding that, “it is thus a global statement about a library’s collection, of which the collection development aspect is but a single component” (Clayton and Gorman, 2006: p.17). Following this, they then define a *collection development policy* as:

...a statement of general collection-building principles which delineates the purpose and content of a collection in terms relevant to both external and internal audiences. It would not normally include the procedures necessary for the implementation of these policies—which should, of course, be detailed in a separate, internal procedure manual. This is where the wider collection management policy fits in (Clayton and Gorman, 2006: p.17-18).

A prime depiction of this overlap can be seen in **Figure 2.4**, which depicts a Venn diagram made by Clayton and Gorman to explain the two types of policies. This figure aptly visualizes how collection management policies and collection development policies can not only stand on their own right, but also work together to create a much more in-depth, detailed *integrated* collection management policy.

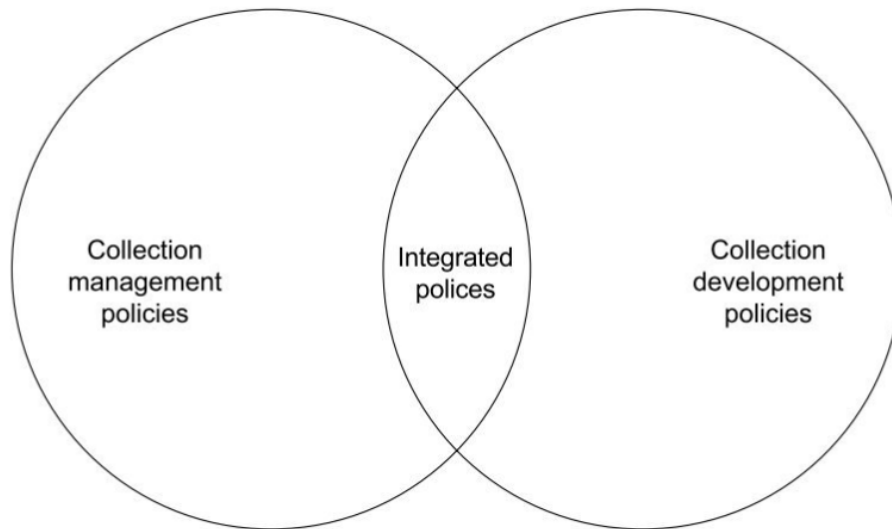


FIGURE 2.4: TYPES OF POLICIES (CLAYTON AND GORMAN, 2006: P.18)

This paper will thus be following this definition of the sometimes-nebulous distinction between collection management and collection development. Furthermore, while some authors use these terms interchangeably, the distinction provided by Hibner and Kelly (2013) will stand here, meaning only the term *collection management* will be used.

2.4.1.2 (Some) definitions

As noted previously, some terms noted in the collection life cycle by Hibner and Kelly (2013) are often used interchangeably—such as “deaccessioning” or “de-selecting”—but may have some subtle differences in their definitions that could prove confusing.

Two of these terms require defining for clarity, which are discussed below:

- Weeding/De-selecting*: Marcella Huggard provides an excellent, thorough discussion of this term, noting that the Society of American Archivists (SAA) defines this term as the “process of identifying and removing unwanted materials from a larger body of materials” (Huggard, 2019: p.2). Internationally, the definition is similar, though delves deeper by including the idea that “weeding, in [the International Council of Archives (ICA)]’s definitions, is clearly related to the removal of a small amount of materials from a larger grouping” (Huggard, 2019: p.2-3). This idea of weeding being on a more item-by-item basis is followed by Clayton and Gorman, where they note that “...books that are not used, books that no longer have readers and unwanted books that get in the way of readers are prime candidates for weeding” (2006: p.198), which seems to approach this process in the same way as Huggard, the SAA, and the ICA. Clayton and Gorman also add the reasoning behind *weeding* being used synonymously with *de-selecting*, stating: “The classic rule is that the criteria for weeding should be essentially those used in the first place for selection—in fact, weeding has often been referred to as ‘deselection’” (2006: p.198).
- Deaccessioning*: Returning to Marcella Huggard and the SAA, this term is defined as “the process by which an archives, museum, or library permanently removes accessioned materials from its holdings” (Pearce-Moses, 2005, quoted in Huggard, 2019: p.2). While weeding/de-selection is

defined as occurring at an item level, deaccessioning occurs at a higher level (often at the level of an entire collection). She notes that while the SAA definition of this term does make the distinction between item-level and collection-level, the International Organization for Standardization (ISO/TC) does not explicitly make this distinction—their definition is quoted as merely the “process of formally removing a document from a collection” (International Organization for Standardization ISO/TC, 2017, quoted in Huggard, 2019: p.2). The ISO/TC does make the distinction in their definition for weeding, where they explicitly state that weeding is used for “individual documents [being] weeded from a collection” (Huggard, 2019: p.2). The ICA, on the other hand, follows with the SAA in more explicitly defining the difference between weeding and deaccessioning as being at an item-level versus collection-level difference.

A final definition that must be given prior to moving forwards is one that will be discussed further along in this chapter, that definition being for *finding aids*. The clearest definition of finding aids comes from Margaret Crockett as, “documentation about the archives which helps to identify, locate and manage them” (Crockett, 2016: p.152). These records can take many forms, from handwritten lists at their simplest to fully completed catalogues at their most complex. Whatever form a finding aid takes, they “should act as maps for

researchers and users so that they can locate records of interest to them”

(Crockett, 2016: p.36). Crockett also points out that there are often two layers or types of finding aids that work together as a single finding aid:

One is a map or schema of how the records are related to each other, usually based on organizational functions or the separate activities of the family or individual. This helps us to preserve the evidence and knowledge of how the records arose from the creator’s activity. The second layer is a detailed description of the content, which may be indexed manually or, increasingly, part of a computerized system with a range of search functions (2016: p.37-8).

As mentioned previously, the finding aid being created through this project is in the form of an Excel spreadsheet, focused on describing the content of the collection itself rather than how it relates to other collections within the special collections department as a whole—making it a second level finding aid. While this somewhat limits the ability to link from this spreadsheet to other finding aids in the Mitchell Library’s holdings, its creation and use allow for *some* forms of linkage not present in the collection’s current state as well as the potential for an easier transfer in the future from the finding aid to the online catalogue.

2.4.1.3 Written or unwritten?

Having a written collection management policy is most often seen as a benefit—as Clayton and Gorman note, “contemporary requirements for accountability oblige institutions such as libraries to be able to justify decisions, especially concerning monetary expenditure” (2006: p.18-19). They also point out that by having a written policy, libraries then have a “contract” with their users that, “...is a useful way to summarize the whole exercise, as it demonstrates to users what they can expect of the library” (Clayton and Gorman, 2006: p.19).

Clayton and Gorman also bring up work done by Richard Gardner in 1981 on the subject, where he points out several advantages to having an explicitly written-out policy, not only to library staff, but users and administration as well. These advantages include, but are certainly not limited to, the following:

- “[forcing] staff to think through library goals and commit themselves to these, helps them to identify the long- and short-range needs of users, and to establish priorities for allocating funds
- “[helping] set standards for the selection and weeding of materials
- “[helping] assure continuity, especially in collections of any size, providing a pattern and framework to ease transition from one librarian to the next”

(Gardner, 1981, quoted in Clayton and Gorman, 2006: p.20).

There are some who may argue that libraries are able to function perfectly well without a written collection management policy, especially since the number of libraries without one far outnumbers that of libraries that have one. This may be a false equivalence, however, as, "...there is no doubt that they (and the collections of any major library without a written collection [management] policy) have been guided by a well-understood, unwritten policy that is part of the institutions 'oral tradition'" (Clayton and Gorman, 2006: p.20). What looks like a library without any form of a collection management policy may truly be one with an *unwritten* policy. Returning to Richard Gardner's observations, however, it can then be understood why having an unwritten policy may not be as beneficial to a library as a written policy. Peggy Johnson speaks to the "protection" offered by a written policy, stating that "the best defense against challenges to a library collection is advance preparation. This begins with a written collection development policy that can easily be shared" (2025: p.82). While this is specifically speaking of collection development rather than management, the premise holds true for both areas. Without a written plan to point to upon questioning or challenging, libraries may risk running afoul of entities holding them accountable, whether that is their users or stakeholders.

Now that the advantages of written policies have been laid out, the question then becomes how does a library create one? This will be discussed in the following section.

2.4.2 Steps of creation

Many libraries may not have collection management or development policies in place at the time of writing, and therefore face the oftentimes arduous process of creating one. In fact, “the literature strongly suggests that the percentage of libraries with a written collection development policy is in the minority” (Schleicher, 2010: p.252). Returning to the previous section, Clayton and Gorman also comment on the lack of explicit, written policy of either kind, stating:

Some libraries have a collection management policy (at various levels of comprehensiveness), others collection development policies, still others combine some of the features of both. Many, of course, have none. We know of no library that has integrated a fully developed collection management policy with a fully detailed collection development policy... (2006: p.18)

The first items that a librarian or team of librarians must understand before beginning fall into two overarching categories: what needs to be considered, and who needs to be involved.

2.4.2.1 What needs to be considered?

The clearest idea of what needs to be considered when creating a comprehensive collection management policy is provided by Johnson, where she states that in the creation process, some of what must be thought of includes the target users and

their needs, budgeting and funding, outreach and communications, resource sharing, and most importantly, the process followed for revising the CMP (2025). This definition, while a very broad overview of the practice of creating a collection management policy, is still detailed enough to stand on its own in this section.

When speaking to the needs of users, both current and potential, the true first decision to be made is whether the policy will be for a library as a whole or a particular department within the library. Both options are valid, and fall down to the management of the library—smaller libraries may tend to choose an all-encompassing policy for the entire library, while larger libraries may choose to specialize policies by department due to the sheer volume of items they hold. In the case of the Mitchell Library, the Special Collections department has been operating on the plan that the policy they need will be one that covers solely that department. Thus, the work completed in this project aims to provide an example for how other collections in the department may be integrated into a future CMP as seamlessly as possible.

The reason this decision must be made prior to the user needs assessment is that the scope of the policy effectively determines the user group being assessed—the larger the scope, the larger the user group(s). A policy addressing the management of a library's entire set of holdings serves a larger variety of user groups than one meant to manage either a single department's holdings or even a

single collection. For this project—which focuses on the narrower field of rare books and special collections departments—this means the range of user groups can be narrowed and simplified to their users and their needs, as discussed previously in this chapter.

Beyond the user needs analysis that must be conducted, librarians must consider multiple other factors such as laws that must be followed, security risks, and others noted by Johnson. These factors span multiple levels of a library, extending past the simple collection lifecycle described by Hibner and Kelly (2013).

Collection inventory and assessment

Alongside the user needs analysis is an analysis of the materials contained within the collection being managed, remembering Matassa’s simple statement of “Know what you have and where to find it” (2011). Understanding the full contents of the collection is necessary, as this is information that acts along with the results of the user needs analysis to inform an appropriately thorough CMP.

For collections that have a significant amount of documentation (i.e., collections recorded in an online catalogue that is kept up to date), this inventory process is simple. Most often these sorts of collections are formed of more mainstream books that arrive at a library with completed metadata already attached. In other cases, like those where items are particularly unique for one reason or another, or of a particularly unusual niche, the inventory process becomes more drawn out

and involves significantly more time and larger staffing requirements. There are benefits to the taking of a library's inventory outside of simply knowing what is *present*; a key benefit being that in the process of taking inventory, some of the information gleaned can then be used to inform the next steps that may need to be taken with the items.

Returning to the idea of the *collection life cycle* as used by both Hibner and Kelly (2013) as well as Clayton and Gorman (2006), this key benefit is clearly shown in their diagrams depicting this cycle:



FIGURE 2.5: COLLECTION LIFE CYCLE DIAGRAM (HIBNER AND KELLY, 2013: P.2)

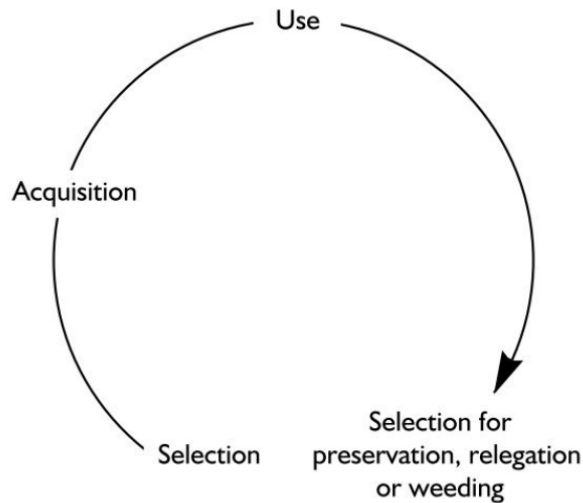


FIGURE 2.6: LIBRARY INFORMATION LIFE CYCLE (CLAYTON AND GORMAN, 2006: P.188)

The second of these diagrams—what Clayton and Gorman call the “library information life cycle” (2006: p.188) as opposed to the exact term used by Hibner and Kelly (2013)—is essentially a simplified form of the first, with some of the stages noted in **Figure 2.5** being folded into the four stages noted in **Figure 2.6**. Both sets of authors explain that these processes are *cyclical* and *perpetual*—collection *management* is a process that would ideally never end, if this were happening in a perfect world (Clayton and Gorman, 2006; Hibner and Kelly, 2013). It is also through these two diagrams that Matassa’s (2011) definition really takes shape; after all, none of these stages can be completed without knowing what is present and where it is in the library’s holdings.

This could be seen as placing the processes of taking inventory and creating a CMP into a chicken vs. egg situation, as one must take inventory to create a CMP while part of implementing a CMP is taking inventory, but in truth this dilemma is only

surface level. A CMP cannot be created without a librarian first taking a complete and thorough inventory of the entire collection, whether that means a single set of books or the library's entire holdings.

A major part of completing this detailed inventory involves the use of item records that contain the metadata discussed in **Section 2.2**. Metadata forms a key part of the cycles diagrammed by Hibner and Kelly as well as Clayton and Gorman. For example, **Figure 2.5** lays out a diagram of Hibner and Kelly's collection life cycle. One of these stages—what they call “processing and cataloging” (Hibner and Kelly, 2013: p.2)—is the stage where metadata most strongly comes into play. In truth, metadata plays a role in every stage of the collection life cycle.

Once the decision as to how the collection being managed will be defined and who the potential user groups of the collection will be is made, as well as the completion of a detailed inventory, the next consideration in CMP creation is the *who*.

2.4.2.2 Who needs to be involved?

There are two schools of thought as to who must be involved in the creation of a collection management policy.

The first is that the process of CMP creation can be completed by a single librarian. As noted by Hibner and Kelly, “Libraries of all types often assign specific selectors to specific collections. In this arrangement, one person makes the final

selections for their assigned collections...It is important to put a person in charge of each collection who can oversee its management and make executive decisions” (2013: p.4-5). This can be beneficial for libraries with thin staffing, as it does not require hiring a large team of people to oversee the library’s collections. It also gives these librarians “a sense of ownership of a collection and encourages deep understanding and familiarity with their assigned areas” (Hibner and Kelly, 2013: p.5), which can then benefit the users of their libraries by ensuring that the librarians have enough of an understanding of their assigned subject areas to answer most questions their patrons may have.

Some argue that having a single person making these decisions, even for one collection, is a net negative for the collection overall. Hibner and Kelly note that “When there is only one person making decisions for a collection, it can become stagnant; a reflection of one person’s philosophy, interests, and opinions” (2013: p.5), which will negate the benefits described previously. The detractors of the first school then argue that collection management policy creation is best completed by a team of people—not only librarians, but members of teams throughout the library and even patrons of the library as well—to combat the stagnation that can occur otherwise. By spreading the work of policy creation between multiple people, the amount of work undertaken by one person is lessened—which allows them to then be able to spend more time on other critical tasks at hand—and the invitation for competing viewpoints, opinions, and expertise is extended in a

manner not possible in the previous school of thought. Bird states this quite clearly: “You can’t be everywhere, and if you have a team in your library, be sure to share the burden...because many minds will see potential glitches, and think laterally when an insurmountable issue arises” (Bird, 2015: p.261). The message is clear in this school of thought: collection management policy should be undertaken by a team of people that includes members of teams from multiple departments of the library as well as from various groups of external stakeholders and library patrons in order to glean the clearest picture of the decisions that must be made.

While the ideal is thus to involve multiple members of staff in the process of CMP creation, this then illuminates one of the most common issues that stands in the way of every library attempting to complete said process.

2.4.3 Common issues preventing CMP creation

Now that the first two overarching considerations facing librarians beginning the process of creating CMPs have been thoroughly discussed, there are some issues that arise rather quickly. This section discusses some of the most common problems: staffing levels, funding restrictions, time consumption, and the emotional impact CMP-related decision-making can have on staff charged with creating and implementing them.

2.4.3.1 Staffing levels

One common reason many libraries, regardless of field, may not have a thorough written collection management policy in place can be linked to staffing levels. Clayton and Gorman note that, “It is unquestionably time consuming, taking both a great deal of senior staff time and, if adopted as the result of a genuinely consultative process involving library stakeholders, taking a considerable amount of elapsed time” (2006: p.20). While this is going on, the staff involved must divert their attention and time away from other, more immediate tasks. This may then force the library to make the decision to sacrifice funding meant for other functions in order to hire extra staff to assist with the creation of a written collection management policy or to continue with the ‘oral tradition’ version of an unwritten policy while acknowledging the risk of, and impact of, collective amnesia.

Staffing levels in themselves are an increasing issue in libraries regardless of sector, with a report from 2009 stating that “many libraries nationally are considering cutting hours and freezing positions” (Oder, 2009: p.32), especially as they attempt to adjust to the U.S. recession that began in 2008. In some smaller libraries, Oder notes that “often just one person works in the library at a time” (2009: p.33). It is clear, then, how staffing levels are just one of multiple, tightly intertwined factors that may come into play in preventing the creation of a thorough written policy. In considering staffing issues, the additional problems of

funding and time arise as well. In approaching either of those problems first, the other two quickly appear. Libraries must then consider the three factors as a shifting balance that must be adjusted in a way that suits them—what suits one library will not suit another.

2.4.3.2 Funding restrictions

Funding restrictions are a familiar problem for librarians in any field, let alone in special collections. In 2009, *Library Journal* conducted a survey of libraries and their budget expectations to compare to a similar survey conducted the year before. This survey found that:

Some 70% of all libraries project an increase in total operating budgets, but only 53% are looking forward to an increase in materials funding. Indeed, in some size categories, including the smallest and largest libraries, materials budgets are expected to decline (Oder, 2009: p.32).

In 2008, Ovenden writes that “the aspect of special collections librarianship that has changed most dramatically in the past thirty years or so is the emphasis on funding” (2008: p.300). While previously, these departments relied on the funding they received from their parent organisations (in Ovenden’s case, the university where his special collections department was located), as this funding stream dwindled, the need for fundraising increased.

Years later, this was still a problem: “Discussions about flat or declining funding have become the ‘new normal’ in public libraries...Since the latest recession started, more than half of the public libraries in the United States (approximately 57 percent) have experienced decreased or flat funding” (Stenstrom and Haycock, 2015). An article published in 2024 by the BBC notes that, “While a study by The Chartered Institute of Public Finance and Accountancy (CIPFA) found library visits increased by 71% between 2021-22 and 2022-23, its chief executive Rob Whiteman said funding still lagged behind the ‘rising demand for services’” (Lynch, Tomas and Hattenstone, 2024). The link between this “lag” in funding and the ability of a library to provide services is further evidenced by the fact that “The poorest areas were around four times more likely to lose a local venue than the richest when permanent closures were mapped to the government's Indices of Multiple Deprivation (IMD) - a system which ranks areas according to income, living conditions and other poverty measures” (Lynch, Tomas and Hattenstone, 2024). Hibner and Kelly note that “each part of the collection life cycle has an implication for budgeting” (2013: p.127), and that is for libraries with already developed CMPs—the increase in demand on funding involved in creating a CMP can prove what seems to be an insurmountable barrier for particularly cash-strapped libraries. Funding is inextricably linked with staffing and time, but it is a vital consideration that must be allowed for when considering a project such as the creation of a CMP.

2.4.3.3 Time consumption

The creation of a thorough and functional CMP does not happen overnight. For libraries in desperate straits regarding funding and staffing levels where they may only be able to spare a single employee to take on this process, this is not a short project. And in cases like those Oder (2009) noted where only a single employee works in a library at a time, this will end up extending the process even further as this person attempts to juggle even the most important tasks to keep the doors open.

2.4.3.4 Emotional impact of decision making

The emotional impact of CMP-related decision-making can be felt across every field of librarianship, especially when it comes to weeding and deselecting items. Jankowski, Schultz, and Soito concur, stating, “stress and aversion associated with making withdrawal decisions have been documented not only in libraries but across other collection-based professions” (2018: p.115). This stress is trouble enough, but can become debilitating in fields like special collections and archives due to the emotional weight attached to the items being held. An example of the weight special collections/archives librarians may feel on their shoulders is eloquently stated by Terry Walker:

We hold a treasure trove of local history and artifacts of cultural heritage that donors have entrusted to the library, expecting that

their donations would be preserved and used to understand who we are and where we came from (2014: p.18).

Possibly making weeding or deselection mistakes is troubling enough for librarians in public libraries and should not be ignored, but there is added urgency when faced with making the same decisions about archival items with emotional weight and history that must be addressed as well. This stress is most likely also present in museums, but there are avenues to alleviate said stress potentially available there that are not present for special collections librarians or archivists.

2.4.4 The importance of collection management

As discussed throughout this literature review, collection management is an incredibly important process that must be conducted by any collection-based organization around the world. Freda Matassa says it best: “collections management can be summed up as: know what you have and where to find it” (Matassa, 2011). In order to prevent a library from becoming nothing more than a building overrun with books, even the most basic collection management policy should be in place.

Having a written policy is especially important for librarians, as it ensures that the librarians can point to an easily shareable document outlining the collection management process. This is especially useful when libraries conduct the heavily politically and emotionally charged practices of weeding/de-selection and

deaccessioning, where items are removed from collections and/or from the library entirely. Earlier, it was noted how stressful these particular practices can be for librarians, partially due to the fact that they are so emotionally charged and can upset patrons and stakeholders alike. By providing librarians with a written policy as a defence, this stress can be lessened significantly, or even removed entirely. And once librarians are more comfortable conducting these practices more often, the better and more honed their collections become, thus ensuring better service to the users they serve.

2.4.4.1 Impact on public interaction

Specifically when it comes to weeding and deaccessioning, there can be the worry of upsetting patrons in particular, which has already been shown to cause stress that ends with librarians abandoning these programs. Clayton and Gorman note that “the ultimate aims of weeding must be to increase accessibility, to improve efficiency and to reduce costs” (2006: p.198)—three aims that can only be accomplished by regular weeding and deaccessioning. When they talk about accessibility, they specifically mean “[facilitating] reader access to useful items” (Clayton and Gorman, 2006: p.210), which in turn serves the patrons better than before.

2.4.5 Need for interdisciplinary policies

As noted previously, there are three main problems preventing the creation of comprehensive collection management policies in single-discipline libraries that must be balanced. These factors are staffing, funding, and time. These three are difficult enough to balance in a single-discipline library, but as soon as the shift from single- to multidisciplinary library is made, a fourth factor is added which is the literature gap identified previously.

The gap in literature not only comprises the fourth factor possibly preventing policy creation, but also intensifies the three pre-existing problems. This is due to the fact that the gap itself must be addressed before any work can be completed on the policy—this requires research, which itself takes time and spare staff to conduct, which in turn increases the need for funding to allow for the other factors to be met. Therefore, by addressing one type of gap (between public libraries, special collections, and museums) and creating an interdisciplinary collection management policy, this project provides a blueprint for other interdisciplinary libraries to follow, effectively streamlining the process to reduce the pressure of the now-four factors. In turn, there may be less pressure *preventing* this process from occurring in other libraries in the future.

2.4.6 In short...

As stated above, using a user-centred review—while not the primary review that must occur during the creation of a collection management policy—is necessary to use as a subordinate review to the materials-centred approach in order for the policy to be as thorough as possible. Rhee seems to agree with Kelly on the need for a user-centred approach, stating, “Archivists should consider historians’ information-seeking and information-managing activities when designing archival description and finding aids” (Rhee, 2010: p.2)—here, the archival description and finding aids mentioned can be extrapolated to then mean collection management policies. By understanding the methods and behaviours used by the most common users of archives and special collections (whether these faculties are free-standing or part of a larger library), archivists and librarians can then design collection management policies that benefit the historians using their services.

2.5 Summary

This chapter contains an overview of the three areas of literature identified in **Chapter 1**: 1) identifying features of “rare books” and “special collections” and how those features are currently being captured by metadata, 2) the potential user groups and their ISBs, and 3) making the metadata containing information about these items’ key features work to best support the identified user groups and their information needs through collection management policies. The difficulty in

pinning down an exact definition of rare books and special collections departments and the items they hold was discussed, prior to a discussion about the exact features that set these items apart from mainstream, general collections, culminating in laying out the commonly used metadata standards used for recording these features. Following this, the beginnings of a user needs analysis was presented, with the identification and justification of the two primary user groups being examined herein. Finally, the importance of collection management as a pillar of library management, and, in turn, how the use of the metadata collected in the collection inventory can be structured and used to best support the identified user groups' information needs was clearly demonstrated through a discussion of prevailing theories.

This literature review has made clear that appropriate and effective cataloguing of rare books and special collections that meets the unique information needs of their users requires a highly specialised and deep knowledgeability on behalf of the cataloguer. Additional gaps in the literature were primarily identified in the user needs analysis, particularly around the information-seeking behaviours of special collections librarians. As such, this literature review sets the basis for and will continue to inform the following doctoral project.

Chapter 3: Methodology

There are three major intended outcomes that are examined and answered throughout this research in pursuit of answering the identified research questions in **Chapter 1**. In brief, these outcomes can be simplified to (1) the identification of the unique features denoting a book as part of rare books and special collections' departments and the associated metadata capturing those features, (2) the identification of unique user information needs and requirements when accessing items in these departments, and (3) the examination of how these unique features of rare books and special collections items work with their associated metadata to meet the various identified user information needs to improve access to these items. The process by which RQ3 is answered argues that this includes the development of a methodology for indexing and cataloguing rare books and special collections items. The aim herein is to unveil these outcomes through the lens of a case study examination of a specific collection and the creation of a digital finding aid for said collection in the form of an Excel spreadsheet meant for use by multiple user groups–this spreadsheet would then contain the metadata required by the needs of the user groups examined in this project.

The process of this finding aid creation requires the data-gathering of three main areas in order to understand these areas: (1) the collection and its representation

via hands-on work with the collection, (2) user information-seeking behaviours and needs via conducting interviews of user group members, and (3) how the collection representation data (i.e., metadata) and user needs can be matched to develop a finding aid with the intention of facilitating access to the collection via the creation of user personas (discussed in **Section 3.4.2.6**) and persona testing (discussed in **Section 3.4.3.2**).

Metadata standards are very useful for librarians of any field, let alone the field of special collections, but this usefulness is less clear in cases where cataloguing software (such as Koha—which is discussed further along in this chapter—among others) is not the primary method of information seeking for multiple user groups. Where finding aids outside of online catalogues or cataloguing software serve as the container of information about the items in a collection and will be accessed by multiple user groups, it needs to be evaluated whether current metadata standards need to be adjusted or changed to maintain their usefulness in cataloguing these sorts of situations. It is for this purpose that this study was conducted using the Stirling Maxwell collection as a test collection being accessed by two test user groups consisting of historians and special collections librarians at the Mitchell Library.

The purpose of this chapter is to discuss the methods employed within this study and their relation to the research aims and questions posed in **Chapter 1**. This includes an overview and discussion of the chosen research philosophy as well as

the research method, prior to a detailed explanation of how exactly the data was collected and analysed, broken down by the research question being examined in each phase. A summary of the ethical considerations of the chosen data collection method is included, before a final discussion of the limitations of all the above rounds out this chapter.

3.1 Research process diagram

Table 3.1 below diagrams a breakdown of the research process used here to answer the three questions and related sub-questions posed in **Chapter 1**. For the purposes of this diagram, only the larger three research questions will be listed, and are treated as having subsumed their respective sub-questions.

<p>QUESTION ONE:</p> <p>What qualifies a collection as one that falls under the purview of “rare books and special collections” departments?</p>	<p>QUESTION TWO:</p> <p>What is the methodology by which librarians may identify who the typical users of rare books and special collections departments are, and what specific information needs they may require from these collections within these departments?</p>	<p>QUESTION THREE:</p> <p>How can the unique features of rare books and special collections work with their associated metadata to meet the identified user information needs and improve access to rare books for these users?</p>
<p>INTERPRETIVIST PHILOSOPHY</p>		
<p>CASE STUDY STRATEGY</p>		
<p>USER-CENTRED DESIGN METHOD</p>		
<p>QUALITATIVE</p>	<p>MIXED-METHOD</p>	<p>MIXED-METHOD</p>

TRANSLATION ⁶	CODING	PERSONA TESTING
DATASET ONE: <i>Test Collection</i>	DATASET TWO: <i>Interview Transcripts and Codes</i> DATASET THREE: <i>Personas/Scenarios</i>	DATASET FOUR: <i>Synthesized Persona Testing Analysis</i> ⁷

TABLE 3.1: RESEARCH PROCESS DIAGRAM BREAKING DOWN APPROACH BASED ON QUESTION

3.2 Research philosophy

This project chose to follow an interpretivist philosophical approach rather than a positivist approach, for reasons discussed below.

3.2.1 Why not positivism?

Positivism is a philosophical approach that could be seen as an appropriate approach for this project, but that idea merely scratches the surface of what this project aims to examine and understand. By definition, positivism “argues for the use of natural science methods to study social reality and beyond” (Clark, 2021: p.23) using the five main tenets, as follows:

1. “Only phenomena, and therefore knowledge confirmed by the senses, can genuinely be considered as knowledge (a principle known as phenomenism).

⁶ Translation in this case refers to the act of applying the combination of rules given by the DCRM(B)3, AACR2, and the special collections staff at the Mitchell Library.

⁷ Dataset Four is not a new dataset, per se, but an amalgamation of the results of Datasets One, Two, and Three.

2. “The purpose of theory is to generate hypotheses that can be tested so that they enable explanations of laws—patterns and regularities—to be assessed (the principle of deductivism).
3. “Knowledge is reached by gathering together facts that provide the basis for laws (the principle of individualism).
4. “Science must (and presumably can) be conducted in a way that is ‘value-free’: in other words, that is objective.
5. “There is a clear distinction between scientific statements and normative statements—judgements about what is ‘good’ or ‘bad’—and a true scientist should only make the former. This last principle is implied by the first one because we cannot establish the truth—or otherwise—of normative statements by using the senses” (Clark, 2021: p.23).

The reason where this project does not truly fall under the approach of positivism is seen specifically in the first and fifth principles. These principles clarify that social research must be done in a way that aims to examine an objective *truth* underlying the world that is unaffected by any sort of social conditions. The premise of this research rests on the understanding that an individual’s information-seeking behaviours are in fact affected by their occupation or reason for research, which is a social condition affecting the underlying truth being sought herein. Additionally, the clear effect of social conditioning—which is not knowledge that can be gained purely from the senses—means that the statements made herein are then only

what Clark (2021) would define as *normative* statements. Normative statements, Clark (2021) states, have no place in a positivist philosophical approach, as they are not statements that discuss an *objective* underlying truth. It is because of these reasons that positivism is disqualified from being an appropriate philosophical approach of any use in this research.

3.2.2 Interpretivism

Clark gives a very basic definition of interpretivism as:

It is based on the view that there are fundamental differences between people and the objects of the natural sciences. Therefore, social scientists need distinct research methods that respect the differences between the natural world and the human one. These methods require the researcher to grasp the subjective experience of social action, what these experiences mean in practice, how those experiences and meanings are understood by others, and why they are interpreted in such ways (2021: p.25).

Where positivism focuses on an objective truth underlying the world, interpretivism denies the idea of this objective truth. Instead, it posits that reality is a *subjective* thing affected by a subject's environment, upbringing, and the like. As such, interpretivists argue that because there is no *objective* truth to the world of social sciences, methods normally used to examine the natural sciences not only cannot be used, but *should not* be used (Clark, 2021). The reasons previously given for

positivism being an inappropriate philosophical approach for this research are then the exact reasons why interpretivism *is* the appropriate philosophical approach here. The entire premise of the research relies on reality being a subjective thing—the information-seeking behaviours—affected in this case by the subjects’ occupations or reasons for using the special collections department.

As an interpretivist approach aims to understand not only the particulars of the subject being examined but the subject as a whole, this makes the approach much more suitable for a case study like this project. Thomas agrees, stating that, “Interpretative inquiry is a form of inquiry that employs a particular approach to answering questions—an approach that assumes an in-depth understanding and deep immersion in the environment of the subject” (2021: p.159). For Thomas, the similar natures of interpretative inquiry and case study structure make them, “not only natural bedfellows but also obvious marriage partners. They were made for each other: it is love and marriage—and they go together like a horse and carriage” (2021: p.159). As the aim within this project is to examine the interaction between user information-seeking behaviours and metadata standards in a case study setting, where the behaviours in question are affected by the individual’s environment, a holistic approach like that of interpretivism is more appropriate than a reductionist approach such as positivism. Holism (specifically, Gestalt psychology), as it is an integral characteristic of the interpretivist approach, must

then be explained to further clarify the choice of philosophical approach used herein.

3.2.2.1 Holism

Holism in its simplest form is defined as a philosophy that views a thing in its totality rather than a collection of parts (Thomas, 2021). The beginnings of holism are found in Aristotle's teachings, where he asserts that, "we progress only by using our practical reasoning, craft knowledge or tacit knowing...It cannot be communicated by reducing it to general principles" (Thomas, 2021: p.51). For Aristotle, because the influence this "practical reasoning, craft knowledge or tacit knowing" has on an individual's worldview—and thus their motivations going forward—cannot be undone, it *must* be considered when attempting to understand the human experience.

This assertion became the driving force in an argument known as the "paradigm wars" (Oakley 1999, quoted in Thomas, 2021: p.51) still ongoing even now. Aristotle's opponents in his time (Plato and Socrates) are "winning," with their idea that "universal truths developed from generalisation" (Thomas, 2021: p.51), but there is still an argument to be made for Aristotle's way of thinking. Flyvbjerg and Sampson argue that this view dominating the world of science as a whole is entirely incorrect (2001), and that the adherence to Plato's positivist, reductionist thinking fails:

...to distinguish between different kinds of inquiry for different purposes and it leads us, *in extremis*, to the absurd position that it is inappropriate to argue or learn from particular examples for fear that this might be thought ‘anecdotal’ and, therefore, unscientific (Thomas, 2021: p.51).

A positivist approach to the type of work done in this thesis would then—as the entire process of answering Question Two is comprised of an examination of anecdotal examples—deem the entire thing unscientific and impossible. To Plato and other positivists, the ultimate reliance of the entire work on the information gleaned through answering Question Two would then not be considered as true knowledge. By contrast, Flyvbjerg and Sampson (2001) contend that the purpose of this research indeed requires the use of what would commonly be denoted as anecdotal evidence. This approach treats the thing being examined as a complete thing, rather than reducing it to a sum of disparate parts.

3.2.2.2 Gestalt psychology

Gestalt psychology in particular demonstrates the prevailing purpose of holism, as it comes down to believing that to understand something—in this case, the “something” to be understood is not just the metadata standards or user information behaviours and needs by themselves, but how the interaction of the two can inform the creation of a finding aid (defined in **Chapter 2, Section 2.4.1.2**)—it must be viewed in its totality, and that, “The key idea to emerge here was that the mind works not by perceiving things in isolation, separately, but,

rather, as wholes—as integrated units” (Thomas, 2021: p.53). Essentially, reductionism and its component philosophies may end up working to *prevent* a true understanding of a thing because they refuse to look at them as a whole picture. Other psychologists—notably Max Wertheimer, Wolfgang Köhler, and Kurt Lewin—concur with the premise of Gestalt psychology, explaining that, “The mind works not in the convenient building blocks of psychological science—variables such as gender, age, reaction speed, time, class—but *wholes*” (Thomas, 2021: p.53). Because this research is meant to examine a particular behaviour of the human mind and its interaction with current metadata standards, a philosophy meant to understand a thing in the most human way is the only appropriate research philosophy to choose.

3.3 Research strategy and method

3.3.1 Qualitative, quantitative, or both?

The overarching research method used in this study is the mixed method. While the qualitative method fits best for *most* of the analytical methods used within due to the processes required to understand the three areas mentioned earlier, there are instances where a quantitative approach needs to be used as well. Quantifiable data alone, provided by methods such as ranking surveys or simple questionnaires—while helpful in providing a basic surface-level understanding of user needs—would not give near the depth of understanding

required to examine the interaction of these needs and current metadata standards. Qualitative data, on the other hand, opens research method design options that *do* allow for the depth of understanding necessitated by this project, such as interviews, focus groups, or persona-based testing (the latter of which was chosen for this project as discussed further along in this chapter). Data such as that gleaned through the recording and translation process in the first stage and the content analysis of the interview transcripts examined in the second stage are not necessarily quantifiable, meaning that the use of purely quantitative research methods applied to the entire project is prevented. Instead, the qualitative method was applied to all the research within, while the brief use of a quantitative method was applied to the applicable data—the quantification of the codes produced in the examination of Question Two and the method of evaluation used to judge the two metadata frameworks examined in answering Question Three. This makes the overarching research method a mixed method involving both qualitative and quantitative methods being used within.

In approaching the research questions under these methods, a multi-modal approach was necessitated here, utilizing a user-centred design method implemented in a case study strategy to accomplish this research. An exact step-by-step explanation of the process followed in this project follows in the rest of this chapter, while this section will focus on a discussion of the theory behind the chosen research methods and why exactly these approaches were chosen.

3.3.2 Case study strategy

The strategy used in the pursuit of this research was that of a case study. A case study is “a focus and the focus is on one thing, looked at in depth and from many angles” (Thomas, 2021: p.9). Here, the focus of this research is the interaction between historians and special collections librarians with the Stirling Maxwell collection currently held at the Mitchell Library in Glasgow, Scotland. More specifically, Thomas defines a case study as, “...analyses of persons, events, decisions, periods, projects, policies, institutions or other phenomena which are studied holistically by one or more methods to illuminate and explicate some analytical theme” (2021: p.12). Yin, on the other hand, defines a case study, “A social science research method, generally used to investigate a contemporary phenomenon in depth and in its real-world context” (2018: p.286). The two definitions are similar, in that they both speak to examining some “concrete entity” (Yin, 2018: p.286), or a “case,” in order to try and explain some phenomenon in a real-world setting, but they differ in one specific area. Thomas argues that a case study is not necessarily a method of researching, but a strategy, as emphasised by Simons here:

Case study is an in-depth exploration from multiple perspectives of the complexity and uniqueness of a particular project, policy, institution, programme or system in a ‘real life’ context. It is research-based, inclusive of different methods and is evidence-led. The primary purpose is to

generate in-depth understanding of a specific topic (as in a thesis), programme, policy, institution or system to generate knowledge and/or inform policy development, professional practice and civil or community action (Simons, 2009, quoted in Thomas, 2021: p.10).

By virtue of being a single project being examined using a user-centred design method, this research meets both of Thomas' aspects to be considered a case study. Going even further by using Simons' definition, this research meets these criteria as well, as it is:

1. On a specific topic—the Stirling Maxwell collection and the creation of its digitised finding aid;
2. Using multiple perspectives—examining the materials held in the collection as well as the needs of multiple user groups who would be accessing the eventual finding aid;
3. Inclusive of different methods—see the discussion on materials-centred and user-centred design methods below;
4. Evidence-led—the data collected in each phase informs the direction the next phase takes; and,
5. Aimed at generating knowledge and informing professional practice—the ultimate goals of this research are to expand the field of knowledge around special collections management and user needs, as well as potentially

creating a framework for special collections librarians to follow for other collections in their care.

In Yin's broad definition of a case study research method, two criteria that he points out are that:

The more that your questions seek to *explain* some contemporary circumstance (e.g., 'how' or 'why' some social phenomenon works), the more that case study research will be relevant. Case studies also are relevant the more that your questions require an extensive and 'in-depth' description of some social phenomenon (2018: p.4).

This research fits these two criteria as well, despite case study being used as a strategy rather than method, as it:

1. Examines *how* two user groups' information-seeking behaviours interact with metadata in rare books and special collections; and,
2. Requires an in-depth examination and description of the above phenomenon.

An important note to make about conducting case studies is that, due to their defining individual focus, the findings cannot be generalised. The goal here is not to widely generalise the exact results of this research to all libraries, as the significantly restricted case sample actively prevents that—users of similar collections at other libraries may have slightly different needs, which may end up

informing a slightly different metadata structure for their own use. Instead, the end results will be used to provide a framework for future research with larger samples (both in population size and geographic size).

Now that it has been made clear that this research fits the criteria to correctly use the case study strategy, the method by which this case is examined must be defined and laid out as well.

3.3.3 User-centred design method

There are two major design methods that must be defined prior to the explanation of which is used in this research: the materials-centred approach and the user-centred approach. Both are named in a somewhat self-explanatory manner, but the definitions are given here for reference. A materials-centred design method means librarians “must assay what quality of materials are needed and what ranges and depths of subject treatment can satisfy a broad cross-section of users” (Kelly, 2015). In other words, the central pillar of this design method is the materials themselves, and the user group is simply left at the general idea of a person rather than any form of specific type(s) of user(s) of the collection at hand—what kind of materials are being held and may need to be added to the holdings? It is easy to define the user-centred approach as one where the needs of specific users are considered the central pillar—what are the library users’ needs?

For any collection management method to function at the highest potential, it must be structured in a way that enables its users to navigate the collection with ease—therefore, while there is an argument for using a materials-centred evaluation approach in place of a user-centred evaluation approach in collection development, both approaches must work *together* to create the best possible collection management structure. Calling the united design method created by using both approaches the “materials-and-user-centred design method” can quickly become both confusing and tedious, so instead the design method named here—user-centred—is used to denote which of the two methods is leaned upon the heaviest in the overall project, rather than to the complete exclusion of the other method.

As noted by Kelly, “Use and user-oriented considerations are analogous to a roadmap for travel showing places of interest and how to get there but without a well-made road constructed out of a quality collection the journey could not be made” (2015), supporting the argument for not solely relying on one approach over the other when developing a collection management policy. Therefore, while the focus of this project begins with a materials-centred evaluation approach during the analysis of Question One, the user-centred approach becomes the focus and driving approach in the analysis of Questions Two and Three.

Clayton and Gorman provide the most succinct support for the involvement of the user-centred approach, stating:

Libraries exist to serve the needs of their users. This service philosophy has always been a basic tenet of our profession. In the management of information resources, too, librarians must seek to serve the needs of their users. The difficulty is that, while some needs of current users are explicit, those of would-be or of future users may not be – and even current users may not even know of the existence of all that they need to have, or have access to. Hence, serving the needs of users is not a passive but an active role. (2006: p.4)

Their argument that library staff must understand the needs and behaviours of their users to appropriately manage their collections is clear by the use of a basic tenet of the librarian profession. With this understanding, librarians may most effectively provide access to their patrons in the ways they both need and desire the most. For an example that fittingly suits the special collections field under examination in this project, Clayton and Gorman state, “...it is too easy to assume that, because users have access to the latest technology, they will use it. Some will, and a library needs to be able to support that use. Others will rely on traditional library materials, so these too will continue to be needed” (2006: p.6)— in this case study, the traditional library materials can be understood to be the books held in the Rare Books Cage (RBC).

While both the materials-centred and user-centred approach are used in this project, the emphasis of the user-centred approach herein means that this project

then becomes a user-centred design project. This is backed up by Allen and Emerald, who note that the contrasting of materials-centred (though they use the term *data-centred*) and user-centred designs is “an overly simplistic dichotomy” (Allen and Emerald, 1996: p.1). They then go on to point out:

Information-system designers can be influenced by the nature of the data contained in their information devices, by the technology used to store and retrieve that data, or by the information needs of users. In reality, the design process always takes all of these factors into account, and the difference between the user-centered, the data-centered, or the technology-centered approaches to design is the emphasis given to one influence over the others (1996: p.1).

Thus, following this reasoning, this project can be called as using a user-centred design method despite the influence of the materials-centred method, as the vast majority of the emphasis within is placed on the influence of the user-centred design. The chosen angle works extremely well with the case study strategy, as the combined use of the materials-centred and user-centred design methods (defined again as being under the user-centred design title following Allen and Emerald’s logic) provide a deeply-informed picture of the collection and its users’ needs, and the case study strategy is defined in part by its ability to drill down to what is called a “polyhedron of intelligibility” (Foucault 1981, quoted in Thomas, 2021: p.5)—also known as the same sort of well-rounded, deeply-informed picture of the

subject. This easy cohesiveness and ability to develop a clear image provided by using user-centred design and case study strategy is why the two were chosen for this research project.

3.4 The questions

This section will discuss the methods, collections, and analyses used in the project, organised by each of the three questions.

3.4.1 Question One's methods, dataset, and analysis

Question One asks, "What qualifies a collection as one that falls under the purview of 'rare books and special collections' departments?" To answer this question, data must first be obtained about the test collection through an in-depth examination of the materials held within. This examination must be conducted to appropriately identify features of the sort mentioned in RQ1(a), and is thus detailed in the following sections.

3.4.1.1 Method

A qualitative approach was necessary to address this question, through the process of translation. *Translation* in this context refers to the application of rules given by AACR2 and DCRM(B)3 (AACR, 2005; Committee, 2011), as well as the requirements noted by the Mitchell Library staff. As this process requires no

quantification, it therefore eliminates the need for a quantitative or mixed-method approach.

3.4.1.2 Previous collection work

Concurrently to, and somewhat preceding, the process noted to this point, work with the collection itself was continued from the point reached in Charney (2020). This work consisted of several steps. The first step was entering all information contained within the slip index—which was defined in **Chapter 1**—as it was written into an Excel spreadsheet. The headings chosen for each column in the spreadsheet are as follows, in order:

- ID Number⁸
- Found in Location Book? (y/n)
- Location Book Number
- Location Noted in Location Book
- Duplicate? (maybe or blank)⁹
- Current Location
- Authorised Author/Editor/Translator
- Transcribed Author/Editor/Translator

⁸ This column is empty currently but was included to provide the opportunity for future renumbering, especially given that not every item was positively identified with a Location Book Number.

⁹ Some of the slips contained in the slip index were thought to be duplicates, but it was unclear if this was the case or if they were simply denoting multiple copies of the same item. This column was used to denote the possibility by including a function whereby entering “maybe” would highlight the entire row in orange.

- See:
- Transcribed Title
- Translated Title
- Number of Volumes
- Transcribed Publication Statement
- Place of Publication
- Translated Place of Publication
- Transcribed Publication Date
- Publication Date
- Language
- Estate Donated From¹⁰
- Description (Extent)
- Notes on Item
- Notes on Slip
- Notes in Location Book
- Notes in S.T.C. Catalogue (1964)
- Conservation Needed? (Y/N/M)
- Found in RBC? (Y/N)

¹⁰ There are three main bookplates seen on these items if there is a bookplate present. This denotes which bookplate is present on the item.

- Upper Left Number¹¹

An example of some entries in this spreadsheet is shown below in **Figure 3.1:**

The screenshot shows a spreadsheet with multiple columns. The 'Conservation Needed?' column is highlighted in red for 'yes' and green for 'no'. Other columns are highlighted in green or yellow.

FIGURE 3.1: SCREENSHOT OF SOME SPREADSHEET ENTRIES DEMONSTRATING LAYOUT

This example is here merely to show the layout of the spreadsheet, rather than show the exact contents of the various cells. As seen in this example, all columns with a yes/no/maybe option were conditionally formatted to colour-code them in green, red, and yellow, respectively. The only change to this pattern was made in the “Conservation Needed?” column, where a “yes” would highlight the cell in red, and a “no” would highlight the cell in green, to clearly denote the need for conservation of the item.

Following this, an attempt was made to locate the specified items within the Mitchell Library’s location books. Entry into these books gives items a 6-digit location book number, some of which were noted on the slips, and most often notes a physical location for the corresponding item. These locations were entered into the spreadsheet, along with any notes written for each item in the location books. As noted in **Figure 3.1**, not every item was found within these

¹¹ The purpose of this number is unclear, but as it is present on the vast majority of slips in the index, it is included herein.

books. When an item was located in these books, there was some confusion on if it was noted in the location books as part of the Stirling Maxwell collection or not. When an item was positively identified as part of the collection, a “yes” was recorded in the “Found in Location Book?” column. If an item was identified by a title, author, and/or publication date match¹² but was not noted as being part of the Stirling Maxwell collection, a “maybe” was recorded in the column instead. There were some problems around the location books and the documentation of the Stirling Maxwell collection, including the directionality of previously transferred information from the slip index into the location books, which will be discussed alongside other limitations and problems later in this chapter.

Once the location book analysis had been completed, a brief examination of the library’s S.T.C. Catalogue from 1964 was conducted in the same manner as the location book analysis. Again, the vast majority of the items recorded in the slip index were not identified in this catalogue, but the information given for the few items that were was recorded in the specified column.

It was at this point that work with the physical items themselves began. The librarians pulled crates of approximately 20 books at a time for analysis. While up to this point the slip index had served as the most accurate record of the items (e.g., if the publication place for an item in the location book does not match that

¹² A combination of two of these items was considered enough to earn a “maybe”, but precedence was given to items that matched all three categories between the location books and the slip index.

listed on the slip for the same item, the publication place noted on the slip was treated as the “correct” one of the two, while the other was only noted in the corresponding location book notes column), the physical items superseded that accuracy from here. A pure transcription of the title, author, imprint, and any volume notation replaced the information given by the slip index. From this point on, the information in the “Description (Extent)” column was recorded according to the current metadata standards in the field as noted by the RBMS Bibliographic Standards Committee (2011). Other columns containing information gleaned from the items were pure transcriptions, including case sensitivity in the interest of gathering as much information as possible from the items themselves prior to them being returned to the RBC.

Following the collection of this information via pure transcription, the AACR2 capitalization rules (AACR, 2005) were applied to the items as specified by the DCRM(B)3 (Committee, 2011). Once the capitalization rules were applied to the information contained in these columns, the next step was to translate the names listed as publication places for items in foreign languages. Here, rather than using Google Translate like was used for the titles, resources such as the RBMS Latin Place Names, CERL, and Britannica were used.

The creation of Dataset One proceeded from this point.

3.4.1.3 Dataset One: Test Collection

Selection of test collection contents

A test sample of 509 items was chosen of the first 509 items to be retrieved by the librarians, rather than attempt to use the entire 3600+ item collection for analysis due to its immense size. A discussion as to the particular limitation of the sampling process here is found in **Section 3.6** later in this chapter. At this point, the pure transcription of the titles, author statements, and imprints for these items was corrected, again using the capitalization rules specified by the AACR2, Appendix A (AACR, 2005) During the process of this transliteration, any titles in foreign languages (most commonly Greek and Latin, though there are some Italian and French titles in this test collection) were translated via Google Translate. This translation was necessary to follow the AACR2 capitalization rules for non-English languages, as the researcher is not fluent in these four languages and was unable to perform the translations without this assistance.

The Koha records

In order to implement a smoother testing process using a knowledge organisation system (KOS) that could be used by a library at any location, a further sample of approximately 50 items were chosen from the test collection spreadsheet for entry to Koha Library Software, which is, “the world’s first free and open source library system” (*Koha Library Software*, 2025)

As the intent of this research is not tied to a particular KOS and the search of the smaller sample is used more as a demonstration than anything else, Koha was the obvious choice for three main reasons:

1. As Koha is an open source and open access system, it functions as an easy comprehensive KOS for the purpose of demonstration.
2. There are some examples of other special collections and archives using Koha to organise their collections, most notably being that of the Vera and Donald Blinken Open Society Archives (OSA) Archivum located in Budapest, Hungary. The Blinken OSA focuses mainly on areas very different from the content of the Stirling Maxwell collection ('Blinken OSA Archivum', n.d.), but this example was followed due to the similarities of being an archive using Koha to catalogue their collections.
3. Finally, returning to the open source and open access of Koha, this system is highly flexible and customizable, both features that make it a suitable tool for demonstrating how metadata can be standardised for situations such as this.

Once the items were selected from the larger test collection, two versions of each record were created—the first being a record based off the information gleaned from the metadata contained in the slips of the slip index, and the second being a record based off of the information contained in the spreadsheet after examination

of the physical items themselves. Details about the selection process for the sample collection and their entry into the Koha system are in the following section.

Selection of sample collection contents

For the purposes of demonstration, a smaller sample collection of items from the records contained for the test collection items was selected for entry into the Koha system. These items were selected using the following criteria:

- More than one third of the items selected needed to be illustrated items. These illustrations include illustrated title pages, mathematical figures, and maps rather than being solely limited to illustrated pictures published within the main body of the item.
- Items selected needed to have some form of information regarding the entity or entities responsible in some way for publishing the item, whether that be in the form of compilation, editing, printing, publishing, or selling.

Both criteria were chosen to account for the fact that the users of rare books collections often need to use the items in said collections not just as carriers for information, but also as valuable information in themselves. Rare books and special collections librarians also use this information to aid in the management and access to these items by staff and patrons alike. Information such as the presence of illustrations is a necessary inclusion in the bibliographic record in the aim of this management, as “those involved in the rare book world usually need to

talk about, describe, and know the terminology of the items in their care” (Berger, 2014: p.130). Cullingford agrees with Berger on this count, stating, “Manuscripts and early printed books require knowledge of their manufacture and meaning to unlock them, along with an understanding of the ways in which they can help scholars bring their period to life” (Cullingford, 2022: p.23). A librarian that knows about the manufacturing processes of items held in a collection armed with records that reflect that knowledge is therefore better poised to assist any patrons that may need that type of information to further their own information needs.

Outside of these two requirements for the items selected for the sample collection, best efforts were made in attempting to select items that had a variety of additional descriptors attached in the spreadsheet (meaning terms like “printed annotations” or “index,” etc.). Items selected that were published in more than one volume meant that all volumes present in Dataset One were then selected for entry into the sample collection. This sample collection was then duplicated and entered into the Koha system: the first version consisted of entries based off the pre-existing slip index and the information contained therein, while the second version consisted of entries based on the updated detailed spreadsheet built over the time spent working hands-on with the books. One important note to make here is that in the slip index, items that were published with multiple volumes present in the overall test collection were combined on one slip, leading to a slight numerical discrepancy between the two versions—despite the fact that both sample

collections contain the exact same items. The slip index Koha records made up a collection of only 34 items, while the spreadsheet Koha records were a collection of 51 items. Examples of the records created are found in **Chapter 5**.

While entering the metadata into the relevant MARC fields for both versions, the procedure for creating core-level records as laid out by the DCRM(B)3 (Committee, 2011: p.168-172) was followed, with some exceptions. Subject headings, genre/form headings, added entries, and series added entries were not included, as these fields are noted as “mandatory if applicable” (Committee, 2011: p.172) and required a depth of work just outside the scope of this project. A thorough discussion of this decision and its effect on this thesis can be found in the discussion chapter discussing the results of all parts of this research.

3.4.2 Question Two’s methods, dataset, and analysis

Now that data has been collected regarding the collection itself, the focus must shift to understanding the previously defined internal and external users of the Special Collections department. This phase of the project aimed to understand the information-seeking behaviours of the users in question and create a set of librarian personas and a set of historian personas to be used in the analysis in the third phase of the research.

3.4.2.1 Method

The mixed method was chosen to be used in pursuit of answering this question, as it required both qualitative and quantitative data analysis. While the qualitative method was used for a larger portion of the data collection and analysis for this phase, the brief use of the quantitative method in the analysis of Dataset Two makes the approach here a mixture rather than purely qualitative.

3.4.2.2 Literature review

To collect the data for this research, a scoping literature review was conducted of three areas: 1) any key features that uniquely denote an item as part of a “rare books and special collections” department as opposed to mainstream library departments and how these features are recorded via current metadata standards, , 2) the selected user groups (historians and department librarians) and their information-seeking behaviours (ISBs) and information needs, and 3) relevant theories behind collection management and how they build the foundation for combining the features and metadata standards researched in the first section with the ISBs and information needs identified in the second section. A detailed analysis of this literature is found in the previous chapter.

Simply put, a scoping review “sets the scene for a future research agenda...[it] documents what is already known, and then, using a critical analysis of the gaps in knowledge, it helps to refine the questions, concepts and theories to point the way to future research” (Jesson, 2011: p.76). It is not as exhaustive or strictly

prescriptive as a systematic review, and was chosen because the literature review conducted here is solely to gain a summary overview of available knowledge in the three chosen areas with the intent to direct specific research questions being examined herein, rather than the breadth and thoroughness given by a systematic review. Another definition that backs this reasoning is given by Arksey and O'Malley, stating that these kinds of reviews may “aim to map *rapidly* the key concepts underpinning a research area and the main sources and types of evidence available” (Mays, Roberts, and Popay, 2001: p.194, quoted in Arksey and O'Malley, 2005: p.21).

The three areas identified by the research questions given in **Chapter 1** and discussed above each required their own literature review, but the steps for each remained roughly the same. The methodological framework chosen will be discussed in the next subsection, followed after by sections containing the details for each of the three areas of the overall literature review.

Scoping review methodological framework

The process undertaken for these scoping literature reviews is the five step, or five stage, framework below:

1. Identifying the research question
2. Identifying relevant studies
3. Study selection

4. Data charting
5. Compiling and summarising the results (Arksey and O'Malley, 2005; Mak and Thomas, 2022; Peters *et al.*, 2020)

The one notable exception created by this framework's application to this study appears in the fourth stage, "data charting." The bulk of available sources on this form of literature review, and thus the framework, are based on research done in the healthcare and medical research fields—this means that the primary definitions for this stage describe heavily quantitative methods of data charting. This form of data charting, while possible, is unhelpful in this project due to the heavily qualitative nature of the information being sought herein. Therefore, for the purposes of this project, the stage will be referred to as "note taking" instead of "data charting" to more clearly represent the documentation method employed at this point in the review. Additionally, rather than referring to each item selected for inclusion in this research as a "study," these items are instead referred to as "sources."

Research question identification

The process of identifying the research questions being examined with this literature review broadly built off of the overarching research questions identified in

Chapter 1. The three main questions are repeated below:

1. What qualifies a collection as one that falls under the purview of “rare books and special collections” departments?
2. What is the methodology by which librarians may identify who the typical users of rare books and special collections departments are, and what specific information needs they may require from these collections within these departments?
3. How can the unique features of rare books and special collections work with their associated metadata to meet the identified user information needs and improve access to rare books for these users?

Further breakdown of how each research question informed the portion of the literature review search methodology is contained in their respective sections.

Identifying relevant sources

For all three portions of the literature review, the search and identification process was the same, save for what is mentioned in their respective sections below.

Two main databases were selected at the outset: SUPrimo—or the University of Strathclyde Andersonian Library’s online catalogue search engine—and Google Scholar. These databases were selected due to their vast coverages, ease of access, and high propensity for fully digitised texts. Once search terms and phrases were selected for each portion and entered into the search engines, the initial round of relevance decisions were made based on the chosen inclusion and

exclusion criteria. Specific inclusion/exclusion criteria are described in the sections discussing each portion of the literature review, but the main, overarching criteria (and their justifications) applied to all sources are as follows:

- Priority was given to sources published after 2000, but this was not a *strict* inclusion/exclusion criterion, especially as research progressed and information about the development of the metadata standards involved in this project was uncovered. The purpose of this decision was due to the ever-evolving nature of metadata standards along with the constant growth of the field of knowledge regarding ISBs and information needs.
- Sources in languages other than English were excluded. This is due to the researcher's very limited understanding of other languages, combined with the relative inaccuracy of online translation services such as Google Translate when attempting to understand more than the overall gist of a piece of text. This relative inaccuracy is discussed in further detail in **Section 3.6.4** later in this chapter.
- Similarly, sources discussing any of the three chosen areas that discussed studies conducted outwith of North America or Europe were also excluded. The reasoning for this choice is similar to the previous criterion, along with the fact that the case study described in this thesis is based in the United Kingdom, which limits the amount of applicable literature that appropriately represents the situation at hand.

All potential sources were judged based on these criteria as well as the portion-specific criteria discussed throughout the section titled “Portion-specific methodology,” and included or excluded for further study in the next stage of the methodological framework.

Source selection

Once potential sources passed through the initial stage of evaluation, the third stage of this framework involved the deeper examination of relevance to the research question being sought through a reading of the source titles and abstracts. Alongside this process, the referencing software EndNote was used to store referencing information for all sources at this stage, along with abstracts where those were available.

Note taking

This stage in the framework was the extraction of information from each source in the form of shorthand notes taken during thorough readings of the full texts, including selected quotations. These notes were stored in two places: digitally through the use of the note-taking software known as Microsoft OneNote, and physically in the form of handwritten notes in multiple notebooks.

Throughout this process, chaining was employed in many cases to obtain yet more sources not found during the use of the two search engines named above. This meant that if further information was deemed necessary on any topic mentioned in

a specific source, the citations used by that source were then extracted and the available sources were then cycled through the same relevance evaluation process.

Compilation and summarisation

Once the note taking stage was completed and no other sources were deemed relevant or necessary for inclusion in the literature review, the next stage commenced. This stage involved compiling the various notes taken on each source in three groups of sources—one for each portion of the literature review. Following this grouping step, brief, informal summaries of each group of notes were created to ensure that the choice to end the searching process for further sources was indeed correct.

Portion-specific methodology

With the basis for the methodological framework of the literature review laid out, this section discusses the steps specific to each of the three portions of the literature review. These steps occurred in tandem with the steps discussed throughout the previous section, and are discussed herein to ensure future replicability.

Key features and metadata standards

This portion of the literature review focused on the key features of rare books and special collections along with the associated metadata standards. As such, the following additional inclusion/exclusion criteria were identified:

- Sources discussing what makes a rare book a rare book were included in order to thoroughly answer the research question for this portion.
- Sources discussing features of historic special collections were included in order to thoroughly answer the research question for this portion using information relating to the case study collection being examined herein.
- Sources discussing features of special collections that are not composed of historic materials were excluded, as these types of special collections do not relate to the case study collection examined herein, and therefore are not relevant to this study.

As a result, the search terms chosen for entry into the specified databases were “features of rare books” and “features of historic special collections”. Additionally, further information that fit the inclusion criteria for this portion was also found in the note taking stage of examining sources identified as relevant in the third portion (the process of which is detailed in “Collection management theories”).

Alongside this process, a second form of chaining was also employed where information mentioned within the text of relevant sources (noted in the note taking stage) about the various metadata standards—particularly if standards were mentioned by name. The search phrases for this information were “metadata standards” and “metadata standards special collections”. The only additional inclusion/exclusion criterion added to those mentioned above was as follows:

- Sources not related to the DCRM(B)3 (Committee, 2011) or the preceding standards were excluded. This choice was made due to the DCRM(B)3 having been identified as the metadata standard currently in use at the Mitchell Library, rather than any other named standard.

User ISBs and information needs

For this portion of the literature review, the following additional inclusion/exclusion criteria were added to those specified in “Scoping review methodological framework”:

- Sources discussing the ISBs and needs of historians were included, as this was the primary external user group specified for this research.
- Sources discussing the ISBs and needs of academic or other researchers in fields related to history were not completely excluded, as they may provide some basis for those discussing historians, but were viewed as a lower priority.

- Sources discussing the ISBs and needs of special collections librarians that work with historic collections were included, as this was the primary internal user group specified for this research.
- Sources discussing the ISBs and needs of special collections librarians that work with materials in special collections fields other than historic collections (i.e., medical and law libraries to name two) were excluded, as the populations these libraries serve are not primarily made up of the chosen external user group, and thus may require different librarian ISBs.

The search terms and phrases used to find sources in the specified databases were “information behavior of historians”, “information behaviour of historians”, “information behavior of special collections librarians”, and “information behaviour of special collections librarians”. All other steps conducted for this portion of the literature review are those discussed in Section “Scoping review methodological framework”.

Collection management theories

The methodology for this portion of the literature review differed most significantly from the prior two sections, as the basis for this portion was based on prior research conducted in Charney (2020). Rather than the first step being creating inclusion/exclusion criteria in addition to those explained in “Scoping review methodological framework,” the sources about CMP theory used in the previous work were re-examined.

In this re-examination, the following inclusion/exclusion criteria were developed:

- Sources primarily discussing collection management *policy* were marked for potential exclusion. As this portion of the literature review in this thesis discussed the theory around collection management as a field of practice, specifics about the creation of such policies were not necessary. Rather than outright exclusion however, it is important to note that sources discussing the policies are also likely to include some relevant discussion on the theory of the field as a whole—thus preventing complete exclusion of these sources.
- Sources that had since been updated with newer editions were excluded in favour of the newer editions, to ensure that all information taken was up to date.

Following this step, the steps previously laid out in “Scoping review methodological framework” were conducted in the proper order, including the process of chaining.

In addition to these sources being re-examined, other sources needed to be obtained to inform the review specifically on the area of metadata standards alongside other information about collection management not previously captured. In order to obtain the necessary information, the search terms “special collections management”, “special collections collection management”, and

“rare books collection management” were used on the selected databases. Sources were selected through this process based on the previously discussed inclusion/exclusion criteria.

Following the information obtained in the literature review (seen in **Chapter 2**), the project proceeded with the next stage by conducting user interviews.

3.4.2.3 User interviews

The aim of conducting user interviews was to gain an understanding of the second area mentioned at the beginning of this chapter: user information seeking behaviours and needs. The information gleaned through these interviews was then analysed with the intent of creating a set of personas with which to test the chosen sample collection as discussed further along in this chapter. As the literature examined in the previous section provided only a broad understanding of the ISBs of the two user groups chosen in this project, which may have created personas that do not completely represent the populations interacting with the collection in this case study if used on its own, the decision was made to additionally interview members of both populations. By combining the information gleaned from these interviews with the information given by the literature review, the personas created more closely represented the actual populations who would be interacting with the Stirling Maxwell collection. This would in turn provide recommendations that would be of greater use for the staff at the Mitchell Library, as well as provide a stronger framework for other libraries in similar positions to follow in future cases.

Interview development

There are three major types of interviews in social research: the structured, semi-structured, and unstructured, each with their own advantages and disadvantages depending on the type of research being conducted. The structured interview, where “the questions and procedures are standardized so as to minimize differences between them” (Clark, 2021: p.191), is most commonly used in quantitative research, with its rigidity and uniformity lending itself well to the number-heavy nature of quantitative work. On the other hand, the unstructured interview, where “the researcher uses little more than a brief set of prompts to deal with a certain range of topics and the interviewee is relatively unconstrained in terms of the discussion” (Clark, 2021: p.426), is much more suited for qualitative research. Finally, there is the semi-structured interview, where there is a list of questions the interviewer must ask, but there is significant freedom to the order in which the questions can be asked as well as in the interviewer’s ability to dig deeper into a participant’s responses with questions not on the list–this is also more suited for qualitative research (Clark, 2021).

The decision was made to conduct these interviews in a semi-structured manner rather than either a fully structured or fully unstructured manner, meaning that the interviews were conducted via conversations with the participants based on a select group of questions, with the flexibility to change the order, wording, and number of questions asked, instead of strictly adhering to a pre-determined list of

questions in a specific order and with specific wording. While all three methods have their merits, the semi-structured method was determined to have a higher likelihood of obtaining the desired information over either the structured or unstructured method, particularly because the external user group being interviewed may not have a prior understanding of what is being sought after.

Once the type of interview method had been decided, attention then turned to the list of questions that would be asked throughout the interviews. The core theme of the interview was built with the twin primary purposes of understanding the ISBs each participant exhibited in their own searches of the Mitchell Library's special collections department along with information around the various scenarios which led to these participants searching for information. Once these purposes had been identified, a list of questions was constructed. This list can be seen in **Appendix B**.

Participant advertisement and selection

Since the author had been working with the special collections librarians on this collection for multiple years, the two librarians closest involved with the collection were approached in person for their willingness to be interviewed about their ISBs. Flyers advertising the need for historians were posted on every floor of the Mitchell Library, inviting participants to contact the author via email. A copy of this flyer is included in **Appendix A**.

These flyers relied on the participants self-identifying with the chosen inclusion/exclusion criteria, so these criteria were kept brief, broad, and clearly communicable. Notably, these criteria were:

- The participants must be 18 years of age or older,
- The participants must have used the resources within the special collections department at the Mitchell Library at least twice, and
- The participants must be able to partake in a 30 minute to 1 hour interview conducted over Zoom.

These criteria were chosen to solicit a wide array of patrons of this department, as this wide array would provide for varied and colourful personas that, in turn, would prove to be much more useful for the end recommendations. The term “historian” is used to speak about these patrons in the broadest sense of the word; here, it means a person who is interested in history for any reason, whether professional, personal, educational, or otherwise. This broad definition was chosen for the same reason as the criteria—to provide a broad data base for the persona creation.

Very few participants who fit all these criteria reached out to participate, which led to the need to expand the criteria to allow for a wider population from which to draw. The only change to the criteria was in the location of special collections resources the participants accessed: this was changed from the specific department at the Mitchell Library to special collections departments overall (i.e.,

the interview participants were no longer limited by location). These advertisements resulted in a final total of three patron participants volunteering their time for this study. Two of the special collections librarians most familiar with the Stirling Maxwell collection were approached for interviews; both agreed to participate.

Once contacted by a potential participant and before scheduling the interview, a combination Participant Information Sheet (PIS) and Consent Form was given for the participant to sign and return to the researcher. Interviews were only scheduled once signed copies were returned to the author. A copy of this combination PIS and Consent Form is attached in **Appendix C**¹³. It is at this point that the participants were given pseudonyms to ensure anonymization.

3.4.2.4 Dataset Two: Interview Transcripts and Codes

The interviews were conducted over the free video-calling service Zoom. This allowed for the interviews to be recorded and automatically transcribed, with the transcriptions serving as Collection Two. The interviews were conducted in a dynamic, conversational format with the overall aim of gleaning information about individual scenarios (both hypothetical and actual) laying out the search process(es) the interviewee follows.

¹³ The PIS in **Appendix C** notes that there may be need for a possible future contact with the participant “sometime in 2024 to evaluate the new finding aid.” While all participants agreed to this, this step was eventually removed from the research plan once the personas took shape.

The recording and transcript files were labelled with the participants' pseudonyms and held on a secure, password-protected external USB drive. Corrections were made to the transcripts to fix where the software had misinterpreted what either speaker had said, then the files were uploaded into NVivo for qualitative coding analysis.

The remainder of this dataset is defined as the set of codes created through the NVivo analysis performed on the interview transcripts that is discussed below, both the codes gleaned directly from the data and those noted upon comparison with the list of codes in literature.

3.4.2.5 Coding

Philip Adu defines coding by stating, "Qualitative coding, which is a sub-category of qualitative analysis, is a systematic, subjective and transparent process of reducing data to meaningful and credible concepts which adequately represent the data and address the research problem, purpose or question(s)" (2019).

The raw data (interview transcripts, in this case) was analysed for relevant data, which was then labelled with a term, or coded. These codes can be defined by the data itself or externally defined, though the former is generally the method used in a grounded theory study. The approach of using internally defined codes follows the idea that a researcher conducting a grounded theory study does not allow outside factors to affect the analysis of the data in order to let the data speak for

itself. Therefore, the first round of coding performed on each transcript was inductive, or more specifically, in vivo coding. In vivo coding creates the terms used for codes “from the actual language found in the qualitative data record” (Saldaña, 2021: p.137), or works from the data itself—using the words spoken by the interviewees—to create the codes. This coding strategy worked in conjunction with interpretation-focused coding, which is a step above that of description-based coding. While in description-based coding the participants’ own words are often used in the creation of the codes to prevent the researcher from performing any of their own interpretation (Adu, 2019), the pre-creation of the codes in this case requires *some* interpretation to be performed. Adu does note that it is possible for the two strategies to be used together, stating, “After selecting an empirical indicator, you could first find out what it is (with the description-focused coding) and then determine what it means (with the interpretation-focused coding)” (2019), which ended up being the case in this research project.

Once the transcripts were first coded in this manner, the codes were then compared to the list of 13 behaviours identified by Hea Lim Rhee as an expansion on the features defined by Meho and Tibbo as well as Ellis (Ellis, 1993; Meho and Tibbo, 2003; Rhee, 2012):

- Starting
- Chaining
- Browsing

- Differentiating
- Extracting
- Verifying
- Monitoring
- Accessing
- Networking
- Information-managing
- Orienting
- Constructing contextual knowledge
- Assessing

These behaviours are explained in more detail in the literature review chapter of this thesis. They are listed here to clarify the codes used when analysing the interview transcripts and identifying relevant data. For the purposes of this project, the behaviours noted for both historians and librarians were used to create the containers for the codes used in the NVivo analysis, as the literature review contained in **Chapter 2** revealed a significant gap in the literature around the ISBs of special collections librarians. The comparison of the in vivo codes to these behaviours was to further identify if these behaviours are identified in the participant interviews. This meant that two passes were needed at a minimum for each transcript—one to identify the descriptive/in vivo codes, and the second to compare those codes to the pre-created interpretative codes.

NVivo was chosen as the analysis software due to the relative ease of use in coding raw transcripts and the output display options. These can be seen in **Figure 3.2** below:

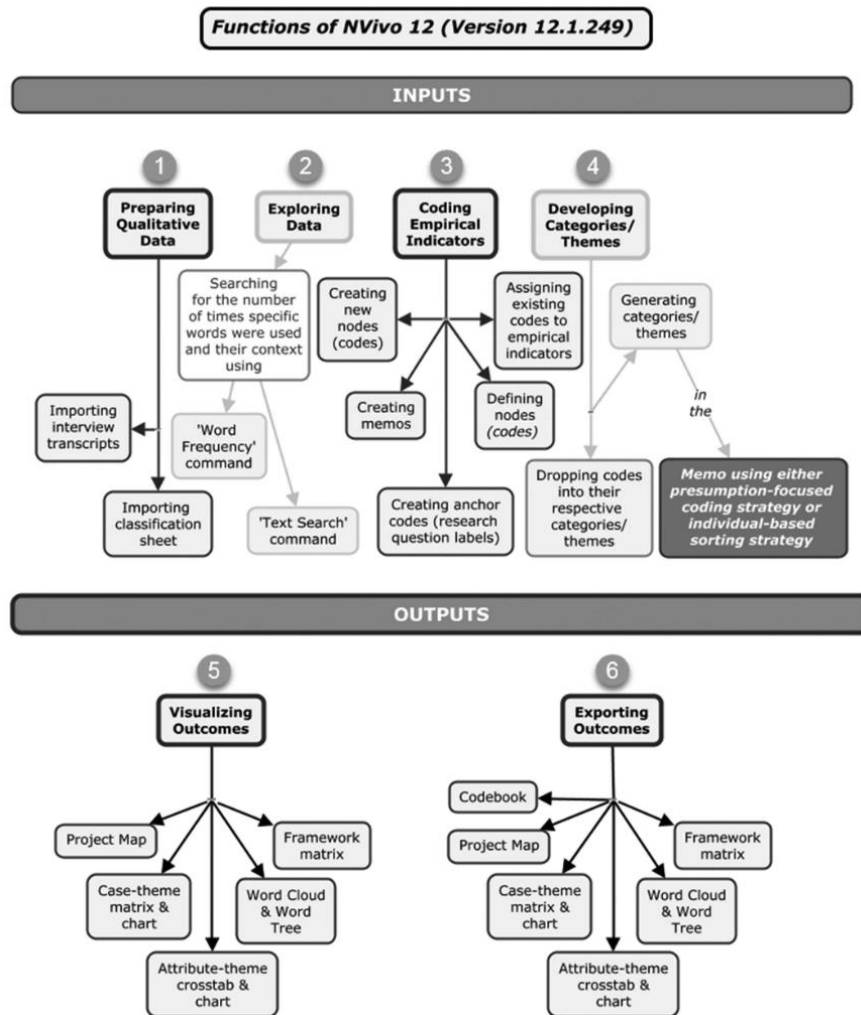


FIGURE 3.2: FUNCTIONS OF NVIVO ACROSS THE SIX STAGES OF DATA ANALYSIS (ADU, 2019)

Once all the transcripts had been coded, a sum total of the instances of each of the identified behaviours was taken to see if there was a prevalence amongst any

of them over the rest. This information, along with the information gleaned in the aforementioned literature review, was then used to create a set of personas.

3.4.2.6 Dataset Three: Personas/Scenarios

This dataset is comprised of the personas created through the quantification of codes within Dataset Two as well as information gathered in the literature review in

Chapter 2. This dataset was then analysed in accordance with the following section on Question Three to create Dataset Four.

Creation of personas

Henka and Zimmermann define personas as “a description of a hypothetical user. This includes a name, a description of the daily routine and of devices and technologies that the hypothetical user uses to access a specific software product” (2014). While in their case personas were used to develop accessibility evaluation guidelines for Web applications, the overall premise of this form of testing and analysis still stands for this research. Again, the goal of this research was to examine the information-seeking behaviour of multiple user groups and how those behaviours interact with current metadata standards. A set of personas that represents average users from both user groups was therefore the best way to adequately test the suitability of the current metadata standards in the various hypothetical situations a typical user may end up facing.

Laubheimer discusses three general types of personas that can be used in product accessibility testing, defining them via the basis of the data used to create them and stating:

1. Proto personas, meant to quickly align the team's existing assumptions about who their users are, but not based on (new) research
2. Qualitative personas, based on small-sample qualitative research, such as interviews, usability tests, or field studies
3. Statistical personas, where initial qualitative research informs a survey instrument that is used to gather a large sample size, and the personas emerge from statistical analysis (2020).

The first type, proto personas, can be ruled out without further explanation based on the quick definition given by Laubheimer; proto personas are based on existing knowledge of a user group's needs and behaviours, rather than any sort of new research. As the research conducted in this project *does* include new research past the use of existing knowledge found in the literature, it is clear that proto personas are excluded as a type of persona to be considered here.

Similarly, statistical personas are also not suited for this research. The creation of a set of statistical personas requires the use of large sample groups, as well as statistical analysis, both of which are not possible in this research project. Firstly, as discussed in **Section 3.6** of this chapter with regards to the sample size of

interviewees for each user group, conducting interviews of large groups was not indicated based on each of the population sizes. Alongside this counterindication, Laubheimer notes that this method requires either access to someone with or personal expertise in statistical analysis (2020). Statistical personas are “like cracking a walnut using a hydraulic press—yes, you can be certain that the walnut’s shell will be thoroughly cracked, but it’s massive overkill in most situations and can leave a mess if not done carefully” (Laubheimer, 2020).

The exclusion of proto and statistical personas then leaves the option of qualitative personas. This technique involves interviewing a much smaller sample size than statistical personas and analysing the responses for common behavioural patterns over many (if not all) of the interviewee responses. The involvement of user interviews provides an accurate representation of user information behaviours, while being significantly less time-consuming than the process of creating statistical personas. The downsides to this method are relatively minor, being mostly concerns of the researcher missing some unique behaviours or outlooks of users due to the small sample size being interviewed and the inability to precisely determine “the proportion of [the] user base that each persona represents” (Laubheimer, 2020). Despite these downsides, the qualitative persona technique is the best fit for this research, and the process for creating these types of personas was followed in this project. The template used to create personas for

each user group (one for members of the public and one for the department librarians) is included in **Appendix D**.

Creation of scenarios

After the creation of the set of personas, a total of six scenarios were also created: four to be used solely by the patron personas and an additional two for the librarian personas to encounter as well as the initial four. These scenarios were also based on the coded interview transcripts and literature review to faithfully represent the kinds of instances in which the created personas would be searching for information within the Stirling Maxwell collection.

3.4.3 Question Three's methods, dataset, and analysis

This phase of research builds on the datasets obtained in the previous two stages, using the information collected about the Stirling Maxwell collection and its potential users (represented here by personas) to test the two metadata frameworks exemplified through Dataset One and its sub-collections stored in Koha for what works best for both user groups.

3.4.3.1 Method

Once again, as both qualitative and quantitative methods were used to analyse this question, this means that this qualifies as the mixed method. The data here was analysed qualitatively through notes taken by each persona about the results

of each search attempt, and quantitatively using recall and precision as an objective method of evaluation.

3.4.3.2 Persona testing

Following the creation of the personas and scenarios detailed in Dataset Three, persona testing was conducted on the two sets of Koha records created from Dataset One.

What this means was that each set of records was searched through by each persona in each scenario, with any comments generated by the searching process noted down. These notes included:

- Complaints generated when searches failed in one or both records sets,
- Points where searches in one set were successful but failed in the other,
- Points where searches in one or both records sets were technically successful, but required additional specific adjustments to be considered fully successful, and
- General comments.

In addition to these qualitative notes, the number of results per search query in each scenario was also recorded to perform quantitative evaluation via recall and precision. The final analysis gleaned from this dual evaluation then forms Dataset Four, the final dataset.

3.4.3.3 Evaluation using precision and recall

The results of each search performed by the personas in each scenario were evaluated in terms of *recall*, or, “the percentage of the relevant documents in a collection that were found by the retrieval system in response to a query” (Buckland, 2017: p.155). This information is calculated using the formula:

$$\text{Recall \%} = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Number of Relevant Documents Present in Collection}} \times 100$$

Alongside the use of recall to evaluate the search results of each scenario, *precision* was used as a second method of adjudication. While recall is a simple measure of the number of relevant documents returned for any given query, *precision* “is a measure of purity...a technical term for the proportion of the documents in a retrieved set that is relevant to the query” (Buckland, 2017: p.155). Using the two concepts together thus provides a researcher with a holistic understanding of the strength of a search and retrieval system, with the ultimate ideal goal being a system that retrieves what Buckland calls, “all of the relevant items (perfect recall) and only relevant items (perfect precision)” (2017: p.156). This value is calculated using the following formula:

$$\text{Precision \%} = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Number of Documents Retrieved by a Search}} \times 100$$

A discussion of what constitutes a “relevant document” for each scenario is contained in **Chapter 5** containing the results of the persona testing process.

As for *why* recall and precision were chosen as an evaluation method over other forms of measurement methods in information retrieval, the reasoning lies squarely in that they best fit what is actually being measured in this research. The Koha search engine is not what is being evaluated here—the ability of the more detailed metadata to provide better matches to persona searches compared to the basic metadata of the slip index is. Therefore, measures illustrating the number of results returned and the number of *relevant* results returned—in other words, recall and precision—are the most appropriate evaluation methods for use herein.

3.4.3.4 Dataset Four: Synthesized Persona Testing Analysis

This dataset is made up of the final analysis of the notes and comments generated by the persona testing, alongside the calculated recall and precision values between the two metadata frameworks for each persona type, seen herein as **Chapters 5 (Results)** and **6 (Discussions)**. This analysis, along with the spreadsheet finding aid, was then informally presented to library staff to help inform the creation of future finding aids by the Special Collections department.

This presentation took place in January 2026 in the Special Collections Department of the Mitchell Library. The contents of this dataset were presented to one of the department librarians via a short PowerPoint presentation and delivery of the two chapters that make up Dataset Four. As work continues on cataloguing the Stirling Maxwell collection, the researcher and librarian came to an agreement to progress

in the same manner, with a re-evaluation planned for after all items in the collection have been accounted for (whether or not they are present in the RBC).

3.5 Ethical considerations

As this project required contact with people, it was necessary to seek ethics approval from the department ethics committee. This required the creation of a Participant Information Sheet and a Consent Form. Both forms were created using the template available from the University of Strathclyde Department of Computer and Information Sciences Ethics Committee website.

An ethics application was submitted to the Department of Computer and Information Sciences Ethics Committee and approved prior to soliciting participants and conducting interviews. No demographic or identifying information was collected, as it was deemed unnecessary for the creation of the personas needed for this research. Additionally, this research did not involve any discussion of sensitive material. The application was formally approved by the Department Ethics Committee on 8 February 2024, and it was only after this point that the flyers soliciting participants (**Appendix A**) were posted in the Mitchell Library.

The expansion of the population approached for the public user group interviews did not require additional ethical approval, as the manner of approach was not significantly changed and there were no additions of vulnerable groups (e.g., minors).

The following sections discuss some specific ethical concerns regarding this research and how they were addressed.

3.5.1 Positionality

Positionality is a topic that comes from feminist research that came into being in the 1970s and 1980s. It means “acknowledging that experience captured by research is subjective, power-imbued, and relational” (Clark, 2021: p.132).

Because this research is using an interpretivist philosophical approach, where reality is held as a subjective phenomenon, it is necessary to recognise the positionality of the researcher within this project.

Due to the researcher having worked with this collection for approximately five years prior to the writing of this thesis as both the researcher and practitioner/cataloguer, this defines the researcher as an insider researcher. An insider researcher is a role that occurs “when a researcher has cultural, linguistic, ethnic, religious, and/or national continuity with the group studied. In a similar way, professionals who decide to carry out research in their workplace could be defined as research insiders” (Clark, 2021: p.133). This role is most strongly present in the research in pursuit of answering RQ1 and in the interviewing and coding process for the librarians, but it is also present in the interviewing and coding process for the patrons, if less so. After all, the insider/outsider debate is

not a binary one, but one presented on a spectrum from one to the other (Clark, 2021).

This insider position comes with both benefits and drawbacks. In terms of benefits, there was much less of a trust threshold to overcome when asking for librarians to participate in the interview process, and the librarians were more comfortable speaking with the researcher in the interviews. Clark points out that one drawback of the insider position, in that “when researchers are insiders, both participants and researchers might assume certain shared understandings without questioning them sufficiently” (2021: p.133). These shared understandings could better be explored if the researcher sat closer to the “outsider” position on the spectrum.

An additional drawback that comes with the insider position is the inherent impossibility for the researcher to maintain complete objectivity. In many ways—such as the research conducted here—the subjectivity is a necessity. Returning to Walker’s comments about the importance of special collections, because the items in these collections hold such emotional weight, emotions are front and centre in decision-making, no matter how hard one may try to obtain complete objectivity.

3.5.2 Power dynamics

There were two sets of power dynamics at play in this research: one between the researcher and librarians, and another between the researcher and the patrons interviewed.

As discussed above, the complicated power dynamic between the researcher and librarians combines that of supervisor-supervisee in both directions, peer-peer, and the traditional researcher-participant dynamic. The work conducted on the Stirling Maxwell collection is conducted by the researcher under the supervision of the librarians, while the discussions around decision-making for the collection operate as if the researcher is the supervisor and the librarians are supervisees.

During the interviewing process, the trust built between the researcher and librarians changed the dynamic from being a purely researcher-participant relationship to that of a peer-peer relationship. This allowed the interview to progress more conversationally, and could have allowed for a deeper understanding of the transcripts prior to the coding process.

On the other hand, the power dynamic between the researcher and patrons fell much more along the traditional researcher-participant line, as neither party had met or worked with each other prior to the interviews. This comparative lack of trust could have resulted in some answers being held back or restrained, which could skew the data. This possibility is why the literature-based codes were used.

3.5.3 Data sensitivity

With such a small sample size from both populations, data sensitivity and appropriate handling can be a concern. For this project, past the signature required for the Consent Form, no identifying information regarding the subjects' age, race, gender, sex, or any other identifiers. Where any such information was freely given by the participants, any record of it was summarily destroyed to maintain anonymity and comply with the stated data management plan in the PIS and Consent Form.

3.5.4 Data handling for the collection itself

When it comes to the Excel spreadsheet finding aid and how it will be handled in the future, the department librarians have agreed to its use and dissemination upon its completion. Other than the electronically-created metadata attached to the file at its inception, there is no authorship information related, as it is a digital file meant to be edited (i.e., location changes, condition changes, etc) far beyond its initial completion under the researcher's hands.

3.5.5 Qualitative rigour framework

Clark (2021) describes criteria for evaluating qualitative research given by two researchers—Lincoln and Guba—as consisting of two main criteria, each with their own sub-criteria: trustworthiness and authenticity.

3.5.5.1 Trustworthiness

The first, trustworthiness, has the following sub-criteria (Lincoln and Guba, 1985 and Guba and Lincoln, 1994, cited in Clark, 2021):

1. Credibility, or the establishment of the data's integrity given interpretivism's view that there are "multiple accounts of social reality" (Clark, 2021: p.364);
2. Transferability, or the ability to take the results from this research and apply it to other situations;
3. Dependability, or the process by which the "merit of research" (Clark, 2021: p.366) is established; and,
4. Confirmability, or the recognition that "the researcher [shows] that they have acted in good faith" (Clark, 2021: p.366).

The credibility of this research was established through triangulation of the data, or "the use of more than one...source of data to study social phenomena" (Clark, 2021: p.363). In this study, interviews were used alongside literature to study the information-seeking behaviours of both populations under examination. The rich depth of research in both the literature review and interviews is shown in the depth of the personas created herein, lending this research its level of transferability.

This chapter acts as a record of the methods used in this research, meaning it meets the sub-criterion for dependability. Finally, the examination and addressing of the researcher's positionality in this chapter also gives this research

confirmability, rounding out the four sub-criteria given by Lincoln and Guba and Guba and Lincoln for research to be trustworthy.

3.5.5.2 Authenticity

Authenticity, the other main criterion for assessing research rigour and validity, has five sub-criteria of its own (Lincoln and Guba, 1985 and Guba and Lincoln, 1994, cited in Clark, 2021):

1. Fairness, or, “Does the research fairly represent different viewpoints among members of the social setting?” (Clark, 2021: p.366);
2. Ontological authenticity, or, “Does the research help members to arrive at a better understanding of their social environment?” (Clark, 2021: p.366);
3. Educative authenticity, or, “Does the research help members appreciate the perspectives of other members of their social setting?” (Clark, 2021: p.366);
4. Catalytic authenticity, or, “Has the research prompted members to engage in action to change their circumstances?” (Clark, 2021: p.366); and,
5. Tactical authenticity, or, “Has the research empowered members to take the steps necessary for engaging in action?” (Clark, 2021: p.366).

The answer to all these questions after performing this research and analysis is a clear and resounding yes. This research is fair as it has represented the varied viewpoints of the participants through both interviews and review of the literature, given the limitations already addressed regarding sample size. The remaining four

authenticity sub-criteria are also answered in the affirmative as seen throughout **Chapter 6**. Therefore, this research meets the authenticity criteria posed by Lincoln and Guba in 1985, and Guba and Lincoln in 1994 (Lincoln and Guba, 1985 and Guba and Lincoln, 1994, cited in Clark, 2021).

3.6 Limitations

There are several limitations to the research performed in this project that must be discussed prior to continuing through this thesis.

3.6.1 Existing collection documentation

3.6.1.1 Missing documentation

Returning to the earlier comments regarding the location books, documentation around how items were organized within the location books could not be located. The location books were not completed chronologically, some collections are missing names, and collections that do have names are often split amongst widely varying location books and numbers. This led to significant confusion about the location of the items listed in the Stirling Maxwell slip index—there are questions around where these items have been moved throughout the years since their donation, or if they are even still in the library's possession. Another piece of missing documentation relating to use of the location books is that which describes the meaning of the "Class Letter" designation. This is a letter given to most (but not all) items documented in the

books, sometimes accompanied by a number in superscript. The *only* information that has been found to date about what this information means comes from an interview of a former Mitchell librarian named Joe Fisher. He stated that, "...the Library had its own classification system. A letter of the alphabet and a broad classification. A, I think, was theology; H was fiction. By my time it was used for practically nothing" (MacDougall, 2017). The Mitchell Library opened in 1877; Fisher began working there in 1947 (MacDougall, 2017). Somewhere in those 70 years, the class letter system had been decided and defined, with the definition eventually falling away. Past the instance of this quotation, nothing is known about this designation in the present time.

3.6.1.2 Illegible documentation

Alongside the chaotic organization of items recorded in the location books came the age-old issue regarding anything handwritten—some of the staff writing in these books had near-illegible handwriting. For some items, enough information was recorded that enough could be deciphered to check against the spreadsheet of the slip index, but other entries were truly unreadable. The same can be said of the handwritten slip index itself. Best guesses could be made, but there is no guarantee that what could be read is an accurate representation of the information about the item.

3.6.1.3 Unclear documentation origin

The final problem presented in the comparison between location books and slip index falls in the question of who wrote the slip index. Being as there was no donor documentation accompanying the collection upon its donation, the current librarians cannot say whether the index was created by a third party prior to donation or if it was created by Mitchell staff upon receipt but before the items were recorded within the location books. There are signs that point both ways; on one hand, there are notes within the location books that seem to point to the index being created prior to donation, while on the other hand, the handwriting styles on a large majority of the slips seem to match that of some of the styles used in the location books. While this may seem a minor distinction, it does bring into question whether it can be safely assumed that the information about title or author contained in the location books matches that written on the slips. That is, the space given for the title alone is much larger on the slips than it is in the location books—in the location books, some information must be left out to fit within that space, while in the slips, some titles have been truncated for unknown reasons (unknown, because the issue of space was clearly not an issue in those cases). If these two repositories were written by different people, what is the likelihood that the same choice in truncation was made in both instances?

3.6.2 Sample size

The small sample size of interviewees from each user group selected here, with three patrons and two librarians, is also an important consideration to discuss. The overall staff in the special collections department at the Mitchell Library is approximately 10-15 librarians, with the two librarians in particular representing the staff with the most knowledge of the Stirling Maxwell collection. Thus, the small sample size of librarians chosen here can be used to create a persona of a special collections librarian who would be accessing this specific collection. In regard to the choice of three patrons, a review of enquiry numbers conducted at the end of October 2023 (Taylor, 2023) showed that an average week at the department would see approximately 87 letter enquiries (both by post and by email) and 541 counter enquiries (meaning an in-person discussion at the special collections department counter). Assuming that each type of enquiry is counted uniquely (meaning there is no overlap between the letter and counter enquiries and that each enquiry is posed by a different individual), this reveals an approximate total of 628 enquiries in an average week. It is more likely however that there is some degree of overlap between the two types of enquiries, so the actual statistic is probably lower than 628. This small end user population size in itself justified the small sample size of patron interviewees.

3.6.3 Collection access

Access to resources within the collection has been a significant hurdle for this work. Due to the current storage situation with the collection and the fact that most (if not all) of the books held by the department have not been catalogued past a slip index as mentioned in the introductory chapter, the test collection size of 509 items was not chosen in a truly mathematically random fashion as would be preferable. Instead, those 509 items were chosen based off the first 509 items that were able to be physically pulled from the RBC by the beginning of August 2023. Therefore, the selection of items in the test collection may not be representative of the entire collection in some respects, but as a truly random spread of items is rare within collections of any type in the field, the need for a random selection of items can be considered unnecessary in this case. The same selection process was followed when narrowing the test collection down into the sample collection for entry into the Koha system, but again is justified for the same reasons.

This access does not only pertain to the choice of the items within the test collection, but also applies in a much broader sense of the term. As discussed previously, the previous storage decisions made by librarians at the Mitchell Library between the donation(s) of the collection to the present day have left for a slightly chaotic distribution of the books within the RBC. The methods of documentation, organisation, and storage have changed significantly over the

decades, and most often these changes are not documented. This includes changes in locations of the items. Combined with the overall backlog of cataloguing meaning a lack of current catalogue entries in most of the Stirling Maxwell collection, these holes in documentation have resulted in difficulty providing access to the collection to the public. The work completed within this study aims to alleviate this current lack of access as well as provide a stronger framework for the Mitchell Library staff to follow in other, similar cases while items are being added to the current cataloguing software.

3.6.4 Online translation services

A note must be made about the prevalence of foreign-language titles in both the test collection and the collection as a whole. As discussed previously, the researcher is not fluent in any of the languages present in the collection and relied on Google Translate to understand the titles well enough to accurately apply the AACR2 capitalization rules as stipulated in the DCRM(B)3 (AACR, 2005; Committee, 2011). There have been numerous studies regarding the accuracy of Google Translate in various languages—most notably in the health profession—and while the software is regarded as a valid option, the consensus is that it pales in comparison to certified translators. For example, a study on the use of Google Translate in three languages (Arabic, Spanish, and simplified Chinese) to communicate directions for use and counselling points on the top 100 drugs in the United States found a significant number of errors:

For the top 100 drugs, 38 unique directions for use and 170 unique counseling points were identified for translation. For the 38 directions for use, 29 (76.3%) of the Arabic translations were accurate, 34 (89.5%) of the Chinese (simplified) translations were accurate, and 27 (71%) of the Spanish translations were accurate. For the 170 counseling points, 92 (54.1%) of the Arabic translations were accurate, 130 (76.5%) of the Chinese (simplified) translations were accurate, and 65 (38.2%) of the Spanish translations were accurate. Of the 247 inaccurate translations, 72 (29.1%) were classified as highly clinically significant or potentially life-threatening. (Cornelison *et al.*, 2021)

There is variation between accuracy of the three languages chosen in this study, but it is important to note that the percentages of inaccuracies are significant enough that they cannot be ignored in the context provided by Cornelison *et al.* (2021). For the purposes of this research however, the potential for inaccurate translations was considered negligible. The purpose of the translations here was solely to gain a general understanding of what was being said so the researcher could accurately apply the AACR2 capitalization rules as mentioned previously, rather than an accurate translation of important health information being used in healthcare. The lowered necessity for perfect accuracy in this case then excused the observed downsides of Google Translate as a translation option.

3.7 Summary

This chapter has discussed not only the exact steps followed in this research project, but also the reasoning behind those methodological choices. In addition to this is a discussion of limitations posed by the chosen methodology and how they were addressed or accounted for in the process of carrying out this research.

Chapter 4: Data

Following the information gleaned in the literature review contained within **Chapter 2** along with the methodology laid out in **Chapter 3**, this chapter lays out the data gathered in pursuit of answering RQ1 and RQ2 as defined in both the introduction and the methodology. The first of the two sets of data, consisting of Dataset One, contains the test collection and the sample collection selected from the test collection, as well as the two styles of bibliographic records for the sample collection based off of (1) the pre-existing slip index with its limited metadata and (2) the Excel spreadsheet with detailed metadata that was manually created for the Stirling Maxwell donation. The second set of data consists of Datasets Two and Three, which collectively contain an examination of user information behaviour through interviews of two sets of end users, which in itself sheds new light on information practices of patrons and librarians in the context of rare books and special collections. The aim of understanding these datasets is to provide a solid foundation for the analysis conducted in the case studies discussed in the next chapter after persona testing of the two styles of metadata for the sample collection.

4.1 Dataset One: Test Collection

As mentioned in **Chapter 3**, the first set of data for this project contains 509 unique items from the Stirling Maxwell collection currently held by the special collections department at the Mitchell Library. These items were chosen in a chronological manner rather than any traditional random sampling method as explained previously in **Chapter 3**—meaning that the items selected for the test collection were chosen based on the order in which they were retrieved from the Rare Books Cage. All these items were verified as currently physically existing within the library’s holdings and confirmed against the slip index as originating within the Stirling Maxwell collection. Where records were available the items were reconciled within the Mitchell Library location books, but this was most often not the case. Other verification methods included checking for one of various Maxwell bookplates as well as handwritten ownership inscriptions, the presence of both of which were noted in the Excel spreadsheet containing the collection.

4.1.1 Categories based on metadata standards

The initial categories used to organise the information contained in the Excel spreadsheet were taken from the DCRM(B)3. These categories are:

- “Title and Statement of Responsibility...
- Edition...
- Material (or Type of Publication) Specific Details...

- Publication, Distribution, Etc...
- Physical Description...
- Series...
- Note..." (Committee, 2011)

The category titled Material (or Type of Publication) Specific Details is not used for items such as those in the Stirling Maxwell donation, with the DCRM(B)3 stating, "No general use of this area is made for printed monographs" (Committee, 2011: p.75). This therefore leaves six categories remaining.

Some adjustments were made to these categories prior to proceeding with persona testing, as these categories were written with the use of more sophisticated cataloguing software in mind. Additionally, DCRM(B)3 "do[es] not address the construction and assignment of controlled headings used as main and added entries" (Committee, 2011: p.12), rather solely focused on the descriptive elements instead. Therefore, for this project, these categories were used as a jumping off point, combined with previous work on the collection, to create the categories noted in the following section.

4.1.2 Categories based on previous work

Referring back to the methodology, previous work done led to the use of the following categories in the Excel spreadsheet¹⁴ (Charney, 2020):

- ID Number*
- Found in Location Book? (y/n)
- Location Book Number
- Location Noted in Location Book
- Duplicate? (maybe or blank)*
- Current Location
- Authorised Author/Editor/Translator
- Transcribed Author/Editor/Translator
- See:
- Transcribed Title
- Translated Title
- Number of Volumes
- Transcribed Publication Statement
- Place of Publication
- Translated Place of Publication
- Transcribed Publication Date

¹⁴ The categories marked with * have been previously explained and commented upon in **Chapter 3**.

- Publication Date
- Language
- Estate Donated From*
- Description (Extent)
- Notes on Item
- Notes on Slip
- Notes in Location Book
- Notes in S.T.C. Catalogue (1964)
- Conservation Needed? (Y/N/M)
- Found in RBC? (Y/N)
- Upper Left Number*

These categories encompass the six areas required by the DCRM(B)3, with the latter set mapping to the former in the manner demonstrated by the below table:

DCRM(B)3 Area	Finding Aid Category
Title and Statement of Responsibility	Authorised Author/Editor/Translator
	Transcribed Author/Editor/Translator
	See:
	Transcribed Title
	Translated Title
Edition	Transcribed Title
	Translated Title
Publication, Distribution, Etc.	Transcribed Publication Statement
	Place of Publication
	Transcribed Publication Date
	Publication Date
Physical Description	Description (Extent)
Series	Number of Volumes
	ID Number
	Found in Location Book? (y/n)

Note	Location Book Number
	Location Noted in Location Book
	Duplicate? (maybe or blank)
	Current Location
	Language
	Estate Donated From
	Notes on Item
	Notes on Slip
	Notes in Location Book
	Notes in S.T.C. Catalogue (1964)
	Conservation Needed? (Y/N/M)
	Found in RBC? (Y/N)
	Upper Left Number

TABLE 4.1: MAPPING OF FINDING AID CATEGORIES TO DCRM(B)3 CATEGORIES

4.1.3 Distribution of items across various categories

This section describes the test collection in more detail based on the various descriptors used in the spreadsheet upon physical examination of the items selected.

4.1.3.1 Languages

Most of the 509 items contained within Dataset One were published in more than one main language. When recording the various languages, best attempts were made to correctly identify the language being used. Best attempts were also made to distinguish between the different levels of language usage based on prevalence within the text. These levels of usage were denoted by the use of the descriptors “[language]” or “[language] (minor)” within the spreadsheet.

In total, eight different languages were identified within the test collection, comprised of the following breakdown, ordered from most prevalent to least prevalent:

Language	Number of titles
English	206
Latin	191
French	90
Greek	26
Italian	22
German	1
Flemish	1
Hebrew	0
Total	537

TABLE 4.2: BREAKDOWN OF MAJOR LANGUAGES USED WITHIN TEXTS IN DATASET ONE

As expected by the fact that some of the items in the test collection contained more than one language in a major prevalence, the total number of titles is noted as 537, rather than the true number of items at 509 titles. This spread can be further broken down into the texts that used a minor amount of a language—this distinction was decided upon when a language was only observed in less than half of a text’s main body, in footnotes/indices, or in both places and totalling less than half of a text. This breakdown is given in the table below:

Language (minor prevalence)	Number of titles
Greek (minor)	47
Latin (minor)	23
Hebrew (minor)	3
French (minor)	1
English (minor)	0
German (minor)	0
Italian (minor)	0
Flemish (minor)	0

TABLE 4.3: BREAKDOWN OF MINORLY PREVALENT LANGUAGES USED WITHIN TEXTS IN DATASET ONE

While the presence of one language within a text did not preclude the presence of another, the presence of one language noted as used in the minor category did preclude the use of the major notation for the same language in the same text. Thus, these two tables can be combined into the following table demonstrating the full breakdown of the various languages used in the texts in the dataset:

Language (Major Usage)	Number of titles	Language (Minor Usage)	Number of titles
English	206	English (minor)	0
Latin	191	Latin (minor)	23
Greek	26	Greek (minor)	47
German	1	German (minor)	0
French	90	French (minor)	1
Italian	22	Italian (minor)	0
Flemish	1	Flemish (minor)	0
Hebrew	0	Hebrew (minor)	3

TABLE 4.4: FULL BREAKDOWN OF LANGUAGES PRESENT WITHIN TEXTS IN DATASET ONE

4.1.3.2 Estates

While there was one person who owned the complete collection prior to its arrival at the Mitchell Library, meaning there was only one estate in question if using the legal definition of the word, there was a need to note separate “estates” represented within the collection. This use of the word “estate” is therefore *not* the legal definition but instead represents the style of bookplate when a bookplate is present. It is not entirely clear when these bookplates were applied to the books, but these bookplates are currently being used as one form of definite confirmation of the books belonging to the Stirling Maxwell donation and therefore must be noted somewhere.

There are three main styles of bookplates present within the collection, each with their own text format and illustrative Maxwell family crest, each of which are pictured below:



FIGURE 4.1: THE THREE MAIN STYLES OF BOOKPLATES PRESENT WITHIN DATASET ONE

Each of the estates noted in the Excel spreadsheet are denoted by the content of and the style of the text present on each style of bookplate, the breakdown of which follows here:

Estate Donated From	Number of Titles
Maxwell of Polloc	212
Maxwell of Pollok, Bart ^t	38
Maxwell of Pollok, Bart	196

TABLE 4.5: BREAKDOWN OF ESTATES PRESENT WITHIN DATASET BASED ON BOOKPLATE

In addition to these three main bookplate styles, it was also noted if there were *no* bookplates present—totalling 55 items—and if no bookplates were present but some other form of identification *was*—totalling nine items. One item was noted as having two separate bookplate styles, so the total given from this breakdown comes to 510 items even though the true total of items in the dataset is only 509 items.

4.1.3.3 Handwritten annotations

Handwritten annotations in this sense refers strictly to the appearance of handwriting in the pages of the item. Many of the bookplates and versos of title pages included some combination of three handwritten characters (one example of which is pictured below in **Figure 4.2**), but as it is unclear exactly what this combination means or who actually added it, the presence of these was not noted with the descriptor. Instead, the presence and content of these combinations was recorded in the “Notes on Item” column. This column also holds, where possible based on the frequency of other handwritten annotations throughout the item, a transcription of said annotations.

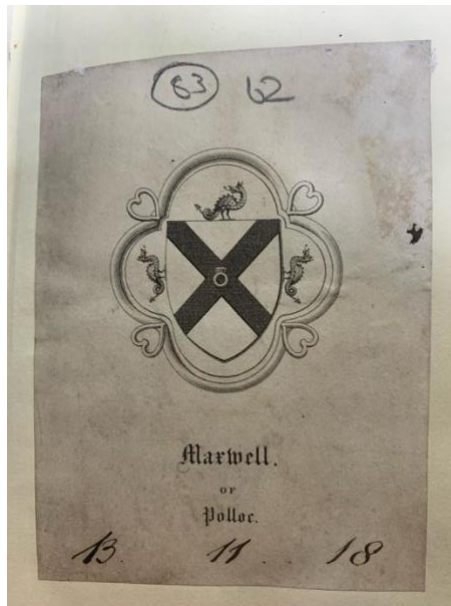


FIGURE 4.2: AN EXAMPLE OF A STIRLING MAXWELL BOOKPLATE WITH THREE-CHARACTER COMBINATION AT THE BOTTOM OF THE BOOKPLATE

Using this definition of the descriptor for this test collection thus means that 42 of the items in the test collection contained handwritten annotations of some sort.

4.1.3.4 Other special descriptors

Following the guidelines set forth in DCRM(B)3 (Committee, 2011), the breakdown of other descriptors present in the “Description (Extent)” column follows here:

Descriptor	Number of Titles
Printed annotations	284
Index	139
Indices	79
Table of contents	201
Illustrated (Ill.)	154

TABLE 4.6: BREAKDOWN OF OTHER SPECIAL DESCRIPTORS PRESENT IN DATASET ONE

4.1.4 DCRM(B)3 areas versus categories

As explained in **Table 4.1**, the six areas required by the DCRM(B)3 (Committee, 2011) cleanly map onto the categories created by the previous work on this collection. The following sections will discuss the mapping of both styles of metadata cataloguing onto the MARC fields used when creating the Koha library for the test collection. Not all of the MARC fields used are given in these sections despite being used in the creation of each record, as they are fields that do not have a partner within either the DCRM(B)3 or the finding aid categories, but are required by the Koha system to create a bibliographic record. A thorough discussion on all of the MARC fields discussed in the DCRM(B)3 can be found in **Chapter 2** of this thesis.

4.1.4.1 DCRM(B)3 areas encoded in MARC

The six areas defined in the metadata standards applicable to the Stirling Maxwell collection are clearly defined as the MARC fields used in the creation of

bibliographic records in Koha based off both the slip index and the spreadsheet. Using the information given in the DCRM(B)3 (Committee, 2011), the following table visualising the various MARC fields mapping onto the six areas is:

MARC Field	DCRM(B)3 Area
1XX – Main entry	Title and Statement of Responsibility
240 – Uniform title	
245 – Title statement ¹⁵	
246 – Varying form of title	
250 – Edition statement	Edition
260 – Publication, distribution, etc. (Imprint) ¹⁶	Publication, Distribution, etc.
300 – Physical description ¹⁷	Physical Description
4XX – Series statement	Series
500 – Source of title proper (if other than chief source)	Note
501 – With note	
502 – Dissertation note	
505 – Formatted contents note	
510 – Citation/references note	
655 – Genre/form headings ¹⁸	
7XX – Added entries	
8XX – Series added entries	

TABLE 4.7: COMPARING DCRM(B)3 AREAS TO MARC FIELDS USED FOR CORE-LEVEL RECORDS

Important to note in regard to the above table is that the Notes area is not limited solely to the 5XX fields specifically mentioned by the metadata standards, but that the Notes area does encompass all of the 5XX fields present in MARC.

4.1.4.2 Finding aid categories encoded in MARC

MARC Field	Finding Aid Categories
050, 082, 086, 090, etc. – Call numbers	ID Number
	Location Book Number

¹⁵ 245 \$a is mandatory

¹⁶ 260 \$c is mandatory

¹⁷ 300 \$a and \$c are mandatory

¹⁸ The preferred vocabulary used for this field is “the terminology used in controlled vocabularies issued by the RBMS Bibliographic Committee,” but other vocabularies are allowed (Committee, 2011: p.172)

	Upper Left Number
1XX – Main entry	Authorised Author/Editor/Translator
	Transcribed Author/Editor/Translator
	See:
240 – Uniform title	Transcribed Title
245 – Title statement	Transcribed Title
246 – Varying form of title	Translated Title ¹⁹
250 – Edition statement	Transcribed Title
	Translated Title
260 – Publication, distribution, etc. (Imprint)	Transcribed Publication Statement
	Place of Publication
	Transcribed Publication Date
	Publication Date
300 – Physical description	Description (Extent)
5XX – Notes	Language
	Estate Donated From
	Notes on Item
	Notes on Slip
	Notes in Location Book
	Notes in S.T.C. Catalogue (1964)
	Conservation Needed? (Y/M/N)
7XX – Added entries	Transcribed Publication Statement

TABLE 4.8: COMPARING FINDING AID CATEGORIES TO MARC FIELDS USED FOR CORE-LEVEL RECORDS

While Table 4.7 notes that the Notes area given by the current metadata standards encompasses field 655 and the entirety of the 8XX fields, those fields are not currently mapped to the Excel spreadsheet categories. This lack of mapping is due to the assignment of genre or form headings and any applicable 8XX fields falling outside of the scope of this project.

¹⁹ Translation in this usage refers to the requirement held by DCRM(B)3 that additional title access points be provided for alternative spellings of titles, amongst other similar requirements for additional title access points (Committee, 2011: p.179-185), meaning that in this case a translated title refers to an additional title access point with corrected and/or modernized spellings.

4.1.5 Sample collection entered into Koha

Following the selection and processing of the items contained in the test collection, a smaller collection of items was selected for a sample collection that was used during the persona testing discussed in the following chapter. A detailed breakdown of the selection process for this sample collection as well as its entry into Koha can be found in **Chapter 3**.

4.2 Dataset Two: Interview Transcripts and Codes

The first part of this set of data consists of the five total interview transcripts obtained through the interviews conducted with three members of the public (the “patrons”) and two special collections librarians based on the interview script developed and explained in **Chapter 3**. These transcripts were collected from automatic transcriptions created by Zoom and corrected for instances where the service misunderstood the speakers and for consolidation—consolidation here meaning the process of removing excess line breaks and insertions of timestamps by the online service. Once corrected and consolidated, the transcripts were then split into the two groups (patrons and librarians) and uploaded to NVivo in order to proceed with coding analysis.

The second half of the dataset involved in the process of answering RQ2 comprises the dataset of in vivo codes and literature codes gleaned while analysing the two groups of interview transcripts collected from interviewing patrons and librarians.

4.2.1 Codes collected through in vivo coding

During the process of in vivo coding of the patron interview transcripts, the following codes were created:

- Examples
- Historical Information
- Names
- Process Steps
 - Breadcrumbing
 - Staff Interaction
 - Stumbling Blocks
 - Wishes
- User Information
 - Occupation
 - Particular Interest
 - Pastime

The definitions of the codes were created during the in vivo coding process, with the five overarching codes being defined as:

- *Examples*: any time a patron spoke of a specific scenario they either had faced or hypothetically might face.

- *Historical Information*: any time a patron spoke on information about the history around what they were searching for in specific scenarios.
- *Names*: any time a patron mentioned the name of someone who had an effect on either their particular interest or on their searching process.
- *Process Steps*: any time a patron described the path they take when searching for information in the special collections department that does not fit in the narrower subcodes defined shortly.
- *User Information*: any time a patron spoke about themselves in a manner that does not fit the narrower subcodes defined shortly.

Each of these codes was directly involved in a separate area of persona/scenario creation, save for the two named “Historical Information” and “Names”—these two were used purely for contextualization when coding examples of information seeking behaviour. The code “Examples” was used for the creation of the scenarios used by the personas when testing the sample collections, as these were descriptions of exact scenarios faced by the patrons being interviewed.

For the purposes of persona creation, the overarching codes of “Process Steps” and “User Information” were the most useful. Within those two codes were subcodes that are defined as follows:

- *Breadcrumbing*: one patron described this process as following a series of clues leading through various items, like a trail of breadcrumbs left by those who came along before.
- *Staff Interaction*: this process was defined by the mention of any time the patrons needed to interact with staff members, whether it was by choice (i.e., the patron approaches staff willingly despite having the option not to) or not (i.e., the patron being forced to interact with staff in cases where interaction is required in order to access materials).
- *Stumbling Blocks*: this code was defined as any time the patrons mentioned running into issues of any sort (i.e., lack of access to materials, poor quality of available materials, etc.).
- *Wishes*: this code was defined by mentions of various features patrons mentioned wanting to have in a search engine.
- *Occupation*: this information was gathered to use in persona creation as a measure of the most common occupation of the interviewees.
- *Particular Interest and Pastime*: these two codes were differentiated from each other by the prior being a particular area of research interest, while pastime denotes any activities performed outside of research.

A table showing the quantified breakdown of the in vivo codes within the three patron interview transcripts follows below:

Code Name	Number of Files ²⁰	References ²¹
Examples	3	25
Historical Information	1	3
Names	1	7
Process Steps	3	29
Breadcrumbs	3	16
Staff Interaction	3	18
Stumbling Blocks	3	41
Wishes	3	7
User Information	0	0
Occupation	3	6
Particular Interest	3	11
Pastime	3	5

TABLE 4.9: QUANTIFIED BREAKDOWN OF IN VIVO CODES IN PATRON INTERVIEW TRANSCRIPTS

During the process of in vivo coding for the librarian interview transcripts, the following codes were created:

- Examples
- Issues
- Process Steps
 - Inter-agency Interaction
 - Inter-staff Interaction
 - Personal Interaction
- What They’re Looking For

²⁰ In this study, “files” refers to the interview transcripts, so this column references the number of transcripts in which a certain code appeared.

²¹ “References” refers to how many times a certain code appears in total across all three files. For example, the code “Examples” appears a total of 25 times across all three files.

The first two codes (“Examples” and “Issues”) and the overarching code of “Process Steps” are self-explanatory in terms of meaning and analysis, but definitions for the three sub-codes and final code are as follows:

- *Inter-agency Interaction*: this is defined as any point where an interviewee mentioned the need to interact with an agency outside of the Mitchell Library. This definition includes other branches of Glasgow Life (the parent company of the library), such as museums and other libraries as outside agencies, despite all being managed by the same parent company.
- *Inter-staff Interaction*: this is defined as any point where an interviewee mentioned the need to interact with any department or staff within the Mitchell Library, such as other special collections librarians, the cataloguing team, etc.
- *Personal Interaction*: this is defined as any point where an interviewee mentioned their interaction with patrons as related to research queries of any type.
- *What They’re Looking For*: this code is defined as points where the interviewees spoke specifically about the process of going from a patron’s initial query to what the patron is actually searching for, as it was noted that these two stages are often not the same.

A table showing the quantified breakdown of these in vivo codes within the two librarian interview transcripts is as follows:

Code Name	Number of Files	References
Examples	2	9
Issues	2	32
Process Steps	2	8
Inter-agency Interaction	1	3
Inter-staff Interaction	2	13
Personal Interaction	2	8
What They're Looking For	2	14

TABLE 4.10: QUANTIFIED BREAKDOWN OF IN VIVO CODES IN LIBRARIAN INTERVIEW TRANSCRIPTS

4.2.2 Codes collected from the literature

The major source of the codes used collected from the literature examined in

Chapter 2 is the study performed by H.L. Rhee, which expands upon work done by

Ellis as well as work done by Meho and Tibbo. This study named a series of ISBs

performed by historians as follows (Ellis, 1993; Meho and Tibbo, 2003; Rhee, 2012):

- Starting
- Chaining
- Browsing
- Differentiating
- Extracting
- Verifying
- Monitoring
- Accessing
- Networking
- Information-managing
- Orienting

- Constructing contextual knowledge
- Assessing

For the purposes of coding the interview transcripts collected at this stage of research using literature-based codes, these thirteen behaviours were chosen as the codes input into NVivo once in vivo coding had been completed. As discussed in the literature review and reiterated in the methodology, these behaviour codes were also applied to the librarian interview transcripts.

The result of this process can be seen in the following table, where an anonymised transcript excerpt from one patron is placed on the left, and the thirteen literature-based codes are placed on the right. Sections of the transcript are highlighted in various colours that correlate to colours used to highlight the codes present in the chosen excerpt, showing how these literature-based codes were applied to each transcript.

Sample Script Excerpt (Patrons)	Code Name
<p>Yeah, that's less of- that's of less interest. I- I'm more concerned with the what than the content, you know, and I mean the context, but I'm less interested in the ah- the bibliographical kind of detail. I know there are people that that's an important thing, you know, which edition it is cause obviously stuff does change with editions. That becomes a- that becomes a thing and- and- and whether you're looking at, for example, because the Mitchell has quite a lot of stuff where it's- it's documents or images that have been taken from a book, and then are photocopied at some point or [mumbled] well hold on second. a- a- is less clear than the original. So sometimes it's a question of- of making sure that what I'm looking at is in the original publication rather than just a photocopy that somebody says is from that publication, just to make sure. And of course, downstream as you get more and more stuff, which is, you know, has been PDF'd for example, and then transcribed, and it's like, "Well, hold on! I now- I now no longer have a direct link from the document I am reading to the document of what's to be taken from, and therefore I can't tell whether it actually is a correct, you know, version of that," which is why having so</p>	Accessing
	Assessing
	Browsing
	Chaining
	Constructing contextual knowledge
	Differentiating
	Extracting
	Information managing
	Monitoring
	Networking
	Orienting

<p>much stuff available online that you can download if you got the space in your laptop, you know, in the cloud, becomes important. But I do find that one of the- I mean, one of the perils of trying to work in Mitchell is- is this- still the split between card indexed and- and online material. And that's- that's this junction, and so, for example, and- and there are- there is stuff that is missing in their filing system. So, for example, and it's- sometimes it's- it's not. Is- is it because the evidence doesn't actually exist ostensibly, but can be inferred? So, for example, I had an occasion late last year to go in, and I was looking at a catalogue of ironmongery and it's undated, because the company that produced it- produced- kept on producing these catalogues, and didn't- and sometimes put edition numbers on them, but didn't actually put dates on them. But you can infer the publication date from certain references within the text or within the illustrations.</p>	Starting
	Verifying

TABLE 4.11: EXAMPLE APPLICATION OF LITERATURE-BASED CODES TO PATRON TRANSCRIPT EXCERPT

Specifically examining the frequency of appearance of these literature-based codes in the patron interview transcripts, the quantified breakdown follows:

Code Name	Number of Files	References
Accessing	3	21
Assessing	3	7
Browsing	2	4
Chaining	3	11
Constructing contextual knowledge	3	7
Differentiating	2	5
Extracting	3	3
Information managing	2	13
Monitoring	2	3
Networking	3	13
Orienting	3	6
Starting	3	11
Verifying	2	4

TABLE 4.12: QUANTIFIED BREAKDOWN OF LITERATURE CODES IN PATRON INTERVIEW TRANSCRIPTS

The same approach as above was applied to the librarian interview transcripts, which can again be shown below using the same style as in **Table 4.11** given previously. This excerpt is also an anonymised excerpt from a single librarian's transcript. The colours chosen in this table have no relation to the colours chosen

in the previous example—they are merely here to give a visual example of what types of data were chosen for each code.

Sample Script Excerpt (Librarians)	Code Name
<p>Exactly, because there- there's so much that it's just not on the catalogue. We have so many index slips in the building, you would not believe it. I mean, if- if those burn down, we're in serious trouble. We've got- we've got location books, we've got accession books, we've got index slips, we've got handwritten finding aids, we've got digitised finding aids, all these things in different formats and so on. So if you've been here a long time, you might not know exactly where something is, but you know where the finding aid is for it, and you know how to use that finding aid. The problem is- and for the customer- is that some of these finding aids are really antiquated, and if you don't have the member of staff interpreting them for you, it's like, "What does that mean?" It's like, you know, when you're presented with the index slips for Stirling Maxwell, a lot of it doesn't make sense, even to me. You can- you can say, "Well, yeah, I understand that. I understand that." But then you have to ask somebody else, "What does that mean?" Sometimes you don't know, because the people who did know are long gone. So we're always trying to improve that. But it's just so slow. It's so slow moving.</p>	Accessing
	Assessing
	Browsing
	Chaining
	Constructing contextual knowledge
	Differentiating
	Extracting
	Information managing
	Monitoring
	Networking
	Orienting
	Starting
	Verifying

TABLE 4.13: EXAMPLE APPLICATION OF LITERATURE-BASED CODES TO LIBRARIAN TRANSCRIPT EXCERPT

The frequency of these same codes in the librarian interview transcripts is broken down below:

Code Name	Number of Files	References
Accessing	2	13
Assessing	1	1
Browsing	0	0
Chaining	0	0
Constructing contextual knowledge	1	4
Differentiating	1	5
Extracting	0	0
Information managing	2	15
Monitoring	1	1
Networking	2	18
Orienting	1	4
Starting	2	7

Verifying	0	0
-----------	---	---

TABLE 4.14: QUANTIFIED BREAKDOWN OF LITERATURE CODES IN LIBRARIAN INTERVIEW TRANSCRIPTS

4.2.3 In vivo codes compared to literature codes

A major difference between the two families of codes used in the NVivo analysis of both groups of interview transcripts is the literature codes lacking terms for user information—the second subgroup of codes noted for the patron interview coding—or examples of scenarios in which patrons might use any form of ISB. The literature codes focus heavily on the process steps, which means that in the following section discussing the relative prevalence of each family of codes in the sample populations, only the codes relating to process steps will be compared.

An interesting result to note here is the melding of the process steps subgroup for librarians with the literature code called “networking,” meaning that three codes become one, and the use of the code “networking” for the purpose that the in vivo code “personal interactions” no longer fits—these items then usually fit more cleanly in instances of “differentiating,” as the discussions at those points were about drilling down to what a patron was truly asking for in their first inquiry. The loss of distinction here is notable in that understanding exactly where a librarian goes for information during the information-seeking process can shed light on where more detailed metadata can be of use for this particular user group.

4.2.4 Prevalence/non-prevalence of codes in sample population(s)

To further distill the information gleaned through in vivo and literature-based coding of the two sample populations, it is necessary to display a table relating the number of references to each code in the sample patron population and the sample librarian population. This table will then inform the following discussion on the creation of personas and scenarios in the next section.

Code Source	Code Name	Patron References	Librarian References
In Vivo Coding	Examples	25	9
	Historical Information	3	N/A
	Names	7	N/A
	Issues	N/A	32
	Process Steps	29	8
	Breadcrumbs	16	N/A
	Staff Interaction	18	N/A
	Stumbling Blocks	41	N/A
	Wishes	7	N/A
	Inter-agency Interaction	N/A	3
	Inter-staff Interaction	N/A	13
	Personal Interaction	N/A	8
	What They're Looking For	N/A	14
	User Information	0	N/A
	Occupation	6	N/A
	Particular Interest	11	N/A
	Pastime	5	N/A
		Total	168
Literature Coding	Accessing	21	13
	Assessing	7	1
	Browsing	4	0
	Chaining	11	0
	Constructing Contextual Knowledge	7	4
	Differentiating	5	5
	Extracting	3	0
	Information Managing	13	15
	Monitoring	3	1

	Networking	13	18
	Orienting	6	4
	Starting	11	7
	Verifying	4	0
	Total	108	68

TABLE 4.15: COMPILED BREAKDOWN OF ALL CODES PRESENT FOR SAMPLE POPULATION TRANSCRIPTS

The clear majority of references for each sample population fell under the codes “Stumbling Blocks” for the patron interview transcripts and “Issues” for the librarian interview transcripts, with 41 and 32 references, respectively. The three next most prevalent codes for the sample patron population are then:

1. Process Steps (29 references),
2. Examples (25 references), and
3. Accessing (21 references).

The three next most prevalent codes for the sample librarian population are then:

1. Networking (18 references),
2. Information Managing (15 references), and
3. What They’re Looking For (14 references).

What these two lists mean for the creation of the personas used in the next phase of research is discussed in the following section.

4.3 Dataset Three: Personas/Scenarios

The final dataset used to answer RQ2 consists of the four personas created after coding analysis of Dataset Two and the scenarios created after in vivo coding of the interview transcripts. While all personas are given names and photos, these do not represent actual people—the names were chosen using a random name generator (Campbell, 1996), and the photos were chosen from free images available within Canva’s Elements tool during the creation of the persona pages seen in the appendices (baseimage, n.d.; corelens, n.d.; diversifylens, n.d.; Monterde, n.d.). There is no otherwise identifying information contained in any of the personas created for this project, so all persona information created through the building process is contained in the four persona appendices mentioned below.

4.3.1 Librarian personas

Following the coding performed in the previous phase, two personas representing the librarians staffing the Mitchell Library. While both librarians interviewed were individuals employed at the library for at least several years, they did discuss scenarios regarding newer librarians searching for information. As such, one persona in this category was chosen to represent the senior library staff in the department (seen in **Appendix E**) while the other was chosen to represent a hypothetical junior librarian in the department (seen in **Appendix F**).

Returning to the four most prevalent codes uncovered for this sample population in the previous section, only three were used for a skeleton persona creation, while the other was used for scenario creation, with the remaining in vivo and literature codes forming the flesh applied to the skeleton persona. Specifically of the four codes mentioned above, the codes “Networking,” “Information Management,” and “What They’re Looking For” were used to inform the skills each librarian persona has expertise in—as both librarians interviewed heavily discussed the process of conducting reference interviews with patrons in order to understand what the patrons are looking for and the process of interacting with staff within the Mitchell Library and other institutions, the level of customer service skills for each librarian persona needed to be on the higher side of average. The connections skill was also informed by the inter-staff and inter-agency interactions, with the senior librarian having a higher level of this skill than the newly employed junior librarian.

4.3.2 Patron personas

Once again following the coding performed in the coding phase, the selected end user group of patrons accessing the Mitchell’s special collections and their behaviours were represented by one of the two patron personas created for testing. This patron persona, noted as the photographic historian persona is present in **Appendix G**.

Given the small sample size of patrons interviewed and the significant overlap in their information-seeking behaviours, as well as the style of information they seek (largely photographic evidence as opposed to textual evidence) it was not possible to create a robust second persona from this data. Therefore, the behaviours noted by Rhee (2012) and discussed in **Chapter 2** were used to create a purely literature-based patron persona. This persona, noted as the Victorian historian persona, can be found in **Appendix H**.

4.3.3 Scenarios

Returning again to the table discussing the relative prevalence of each code for both sample populations (**Table 4.15**), the specific codes used to create these scenarios each persona would face during testing are as follows:

1. Examples
2. Stumbling Blocks
3. Issues

The “Examples” code was the primary code informing the creation of these scenarios, as these were examples based in the sample patron population’s actual experiences with the special collections department at the Mitchell Library and the sample librarian population’s actual experiences with patron inquiries and the process of obtaining information for patrons, other staff, and other agencies. This resulted in the following list of scenarios:

1. A patron wanting to examine books with illustrations done by a particular illustrator published in the 17th through 18th centuries.
2. A patron wanting to examine books with ownership inscriptions (ex libris annotations, rather than other forms of handwritten annotations) to track the ownership history of books in the Stirling Maxwell collection.
3. A patron wanting to read books published in the 17th through 18th centuries about religious history and analysis.
4. A patron wanting to examine books published in the 17th through 18th centuries by the publishing house Elzevir.

These four scenarios were used for both sets of personas, with the patron personas approaching the scenarios from the perspective of the patron wanting to access the requested items and the librarian personas approaching the scenarios from the perspective of a librarian assisting the patron mentioned in each scenario. Two other scenarios were created in addition to the prior four, though these two scenarios were only for use by the librarian personas. These two librarian-specific scenarios are below:

1. A member of the conservation team requesting a list of items needing urgent conservation, along with a brief description of what conservation may be needed exactly.

2. A member of Glasgow Museums wanting a collection of originally bound items owned by the Stirling Maxwell family with handwritten annotations with the intent of displaying at Pollok House.

These scenarios were created based off of the presence of the codes “Inter-staff Interactions” and “Inter-agency Interactions” in the librarian interview transcripts to represent the additional types of scenarios that employees of the Mitchell Library will face apart from simply helping with patron inquiries. Therefore, while the patron personas created herein were only tasked with the initial four scenarios, the librarian personas were each tasked with a total of six scenarios.

4.4 Summary

This chapter has discussed the raw data collected from examination of the physical items available at the Mitchell Library as well as that collected through the interviews conducted of both end user groups and the coding of the resulting interview transcripts. The first section broke down the contents of the test collection and the sample collection, the latter of which was entered into the Koha system for the purposes of persona testing. The second section discussed the data collected and examined in the process of creating personas and scenarios for the next phase of testing and data collection. This discussion is necessary to set up the next chapter, where the results of the persona testing on the sample

collection in Koha will be discussed in depth and examined as they relate to the overarching research questions posed at the beginning of this project.

Chapter 5: Results

5.1 Before proceeding...

Following the data collected and laid out in the previous chapter, this chapter focuses on the results of the persona testing on both sample collections entered into the Koha system. This testing was broken down into two case studies, labelled Case Study One and Case Study Two, both of which will be defined and discussed in this chapter.

This testing was conducted in both case studies using the four user scenarios and two additional librarian scenarios created in the previous chapter. For ease of reference, these scenarios are repeated here, starting with the user scenarios:

- A patron wanting to examine books with illustrations done by a particular illustrator published in the 17th through 18th centuries.
- A patron wanting to examine books with ownership inscriptions (ex libris annotations, rather than other forms of handwritten annotations) to track the ownership history of books in the Stirling Maxwell collection.
- A patron wanting to read books published in the 17th through 18th centuries about religious history and analysis.

- A patron wanting to examine books published in the 17th through 18th centuries by the publishing house Elzeviers.
- A member of the conservation team requesting a list of items needing urgent conservation, along with a brief description of what conservation may be needed exactly.
- A member of Glasgow Museums wanting a collection of originally bound items owned by the Stirling Maxwell family with handwritten annotations with the intent of displaying at Pollok House.

This chapter lays out the individual results of each case study before proceeding to examine and compare the two datasets in-depth prior to the next chapter, where the results from this entire project are discussed at length with regards to what they mean for the Mitchell Library and possibly for other special collections departments outside of the Mitchell.

5.1.1 Distinction between Case Study One and Case Study Two

The distinction between these two case studies lies solely in the differing content of the sample collection being examined therein. Thus, the definitions of the two studies are as follows:

- *Case Study One (CS1)*: This case study involves the persona testing of the sample collection containing 34 bibliographic records created from the pre-existing slip index slips.

- *Case Study Two (CS2)*: This case study involves the persona testing of the sample collection containing 51 bibliographic records created using the more detailed entries from the Excel spreadsheet created by the author.

The content of the records involved in both case studies correspond with each other, meaning they are all recording the same items. The difference in number of records is simply due to the slip index treating multivolume publications as a single item and the Excel spreadsheet treating the same type of publications as individual items.

5.1.2 Distinction between search engines used by each persona type

Before discussing either case study in further detail, the difference between the two sides of Koha used by the two types of personas must be explained, as multiple references will be made to the format of the two search engine pages in both the process steps and the analyses later on.

When searching for items in the Koha online public access catalogue (OPAC), patrons are faced with the following advanced search page:

Home > Advanced search

Advanced search

Search for:

Keyword

and Keyword

and Keyword

Item type

Limit to any of the following:

Book

Availability:

Only items currently available for loan or reference

Sort by:

Relevance

Publication date range

For example: 1999-2001. You could also use "1987" for everything published in and before 1987 or "2008-" for everything published in 2008 and after.

Language

No limit

Audience

Any audience

Content

Any content

Format

Any format

Additional content types for books/printed materials

Any

FIGURE 5.1: ADVANCED SEARCH USING THE KOHA OPAC

The librarians have access not only to the OPAC, but the internal back end of the Koha library, which has a slightly different appearance.

On the internal site, the librarians have access to the following advanced search page:

Advanced search

Search Fewer options Clear fields Go to item search

Search for

Keyword [] Scan indexes:

and Keyword []

and Keyword [] [+]

Item type

Limit to any of the following:

Book

Limits

Year: [] (format: yyyy-yyyy)

Language: No limit [] Language of original: No limit []

Subtype limits

Any audience [] Any content [] Any format [] Additional content types []

Location and availability

Only items currently available:

Individual libraries: All libraries []

Sorting

Sort by: Relevance []

FIGURE 5.2: ADVANCED SEARCH USING THE INTERNAL SITE

Due to the difference in access between the two groups of users, the patron personas searched the OPAC, while the librarian personas performed their searches on the internal site when going through the scenarios listed above.

5.1.3 Evaluating persona search results

As explained in **Chapter 3**, the results of each search performed by the personas in each scenario will be evaluated in terms of recall and precision, which are calculated using the following formulae:

$$Recall \% = \frac{Number\ of\ Relevant\ Documents\ Retrieved}{Total\ Number\ of\ Relevant\ Documents\ Present\ in\ Collection} \times 100$$

$$\text{Precision \%} = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Number of Documents Retrieved by a Search}} \times 100$$

Both the recall and precision percentages for each persona type in each scenario will be compared between the two case studies and explored in detail in **Chapter 6**. The definition of a relevant item will vary by scenario, but there is no variation among these definitions between persona types or between the two case studies. Therefore, the definitions of what makes a relevant document herein are given below:

- *Scenario One:* A relevant document in this case is defined as an item noted in the bibliographic record as being illustrated (either using the full word “illustrated” or some abbreviation thereof) with a publication date falling between 1600-1799, inclusive.
- *Scenario Two:* A relevant document in this scenario is defined as an item noted in the bibliographic record as having a handwritten inscription denoting ownership of the item. This notation can be done through the explicit terms “ownership inscription”, “ex libris inscription”, any variation thereof, and/or the note of handwritten annotations containing text that makes up such an inscription.
- *Scenario Three:* A relevant document in this case is defined as an item noted in the bibliographic record as containing a subject heading that either broadens to or narrows to the term “religious history” and/or “religious

analysis”, or any such related terms given by LCSH, with a publication date falling between 1600-1799, inclusive. Also acceptable is an item containing a keyword that broadens to, narrows to, or is related to the terms “religious history” and/or “religious analysis” in the same publication date range.²²

- *Scenario Four:* A relevant document in this case is defined as an item noted in the bibliographic record as being published by the publishing house “Elzevier” or any historical name known to be related to this publishing house, again with a publication date falling between 1600-1799, inclusive.
- *Scenario Five:* A relevant document here is defined as an item noted in the bibliographic record as requiring urgent conservation using any combination of words/terms that mean such.
- *Scenario Six:* A relevant document here is defined as an item noted in the bibliographic record as remaining in its original binding and having handwritten annotations. Original binding here means anything that has not clearly been rebound since entering the Mitchell Library’s holdings. Handwritten annotations here include both the inscriptions mentioned in the definition of relevant items for Scenario Two as well as

²² As the bibliographic records created for these case studies did not include this information for either experiment, the information used to judge an item’s relevance was taken from external bibliographic records for each item obtained through a records search of the Library Hub Discover database created by Jisc (*Library Hub Discover Advanced Search*, n.d.). A thorough discussion of this is contained in **Chapter 6**.

other forms of handwritten annotations excluding any handwriting on the estate bookplates.

Both the recall and precision values will be displayed in a table for each scenario.

5.2 Case Study One

5.2.1 Sample collection

The sample collection examined in this first case study is the virtual Koha library created using 34 bibliographic records based on the information contained in the pre-made slip index for the Stirling Maxwell collection. This information was taken from the front and back of each slip, treating inscriptions written in various hands and inks as if they were all written at the same time by the same person for the purposes of metadata entry *only*. This means that all of the information written on the front and back of each slip was treated as valid and accurate and entered as such. Once the records were fully created and stored within Koha, any notes or inscriptions written on the slips that fell into either the chosen 024 Standard Identifier or 500 General Note MARC fields were then treated as if of unclear origin in either time (meaning it is uncertain *when* in the collection's lifecycle they were made) or person (meaning it is uncertain if the same person who wrote the author and title information *also* wrote the extra inscriptions). One example of such an inscription visible in **Figure 5.3** below is the pencilled checkmark made in the lower

left corner, which was recorded in the 500 General Note MARC field as seen in

Figure 5.5.

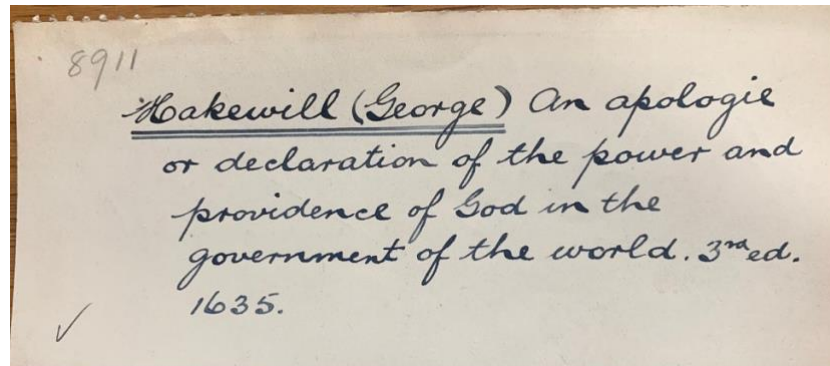


FIGURE 5.3: AN EXAMPLE OF THE FRONT OF A SLIP FOR AN ITEM SELECTED FOR CS1'S SAMPLE COLLECTION

Information given on the front of a slip typically includes what—at the time of the slip index's creation—is the authorised access point for the author, a portion of the title and a four-digit number, as well as the publication year. Occasionally, as in the example given in **Figure 5.3**, edition information can be given. Not pictured above, but also occasionally present is a publication place or information about if a slip refers to an item published in more than one volume. Other symbols, such as the checkmark in the lower left corner in the above figure, are undefinable in that the meaning of these symbols has either never been written down or has been recorded but since forgotten. The presence of these undefined symbols is still recorded in the Koha records created for CS1.

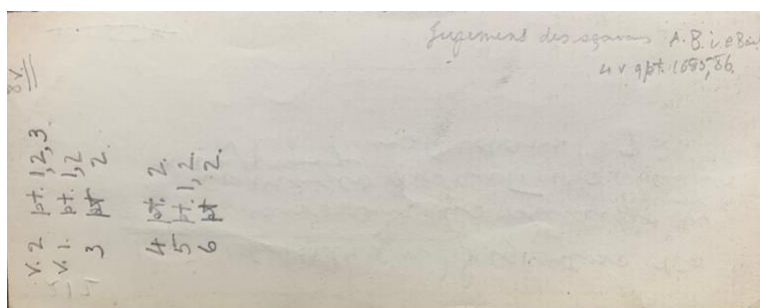


FIGURE 5.4: AN EXAMPLE OF THE BACK OF A SLIP FOR A SEPARATE ITEM SELECTED FOR CS1'S SAMPLE COLLECTION

Figure 5.4 shows the backside of a slip representing a separate item present in the CS1 sample collection. In some cases, such as this one, information about if the item is published in more than one volume (specifically if multiple of those volumes had been donated as part of the collection) will appear here rather than on the front of the slip. Other information that may appear here includes information about the author's name—often accompanied by initials that appear to describe the source from which this information is taken—publication information, or notes about the physical state of the copy in question held as part of the collection, but this is not the limit of the information that can appear here.

Returning to the example slip shown in **Figure 5.3**, the information contained on that slip was entered into Koha using the MARC field definitions given by the Library of Congress (*MARC21 Format for Bibliographic Data*, 2000) and guidelines given by the DCRM(B)3 for minimal-level records (Committee, 2011). This results in the MARC record displayed via the OPAC in **Figure 5.5** below:

[Normal view](#)
[MARC view](#)
[ISBD view](#)

An apologie or declaration of the power and providence of God in the government of the world. (Record no. 24)

[\[view labeled \]](#)

```

000 00567nam a22001697a 4500
003 0St
005 20250628231027.0
008 250414b ||||| |||| 00| u eng d
024 7 0 _2Stirling Maxwell Slip Index Upper Left Number
    _a8911
040 _cal
    _edcrmb
100 _aHakewell, George
    _cDoctour of Divinitie and Archdeacon of Surrey
    _926
245 _aAn apologie or declaration of the power and providence of God in the
    government of the world.
250 _a3rd ed.
260 _c1635
500 _aNotes on front of slip read: "[checkmark]"
942 _2ddc
    _cBK
    _n0
999 _c24
    _d24

```

FIGURE 5.5: MARC RECORD DISPLAY IN OPAC FOR THE RECORD CREATED FROM FIGURE 5.3

As the items entered into Koha were done so as if the records would be used by a physical library, authority records were also created based not only off of the information contained on the slips but also off information contained in the spreadsheet. These authority records were then linked to the bibliographic records used for both case studies, so for the purposes of CS1 the information held in Field 100 \$c (“Titles and words associated with a name”) was ignored, as this information was given by the spreadsheet.

This MARC record results in the following view in the OPAC upon searching for and selecting the item:

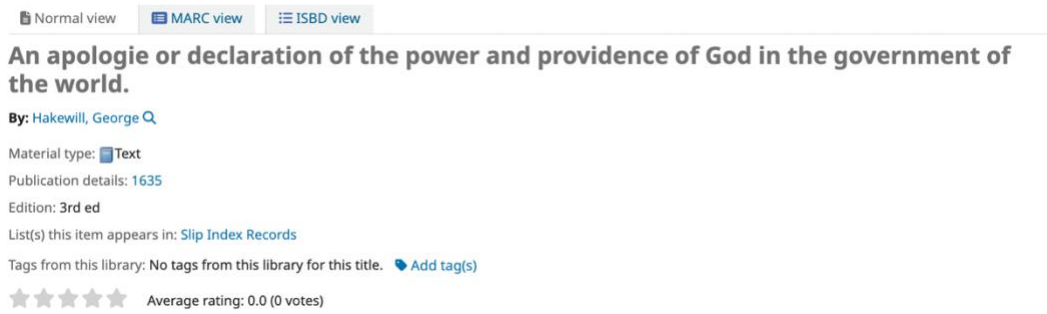


FIGURE 5.6: NORMAL OPAC VIEW OF THE BIBLIOGRAPHIC RECORD CREATED FROM FIGURE 5.3

This OPAC view is what appears for the patrons of the library when searching the online Koha catalogue.

On the internal Koha site, the record for the item displayed in **Figure 5.3** appears as below:

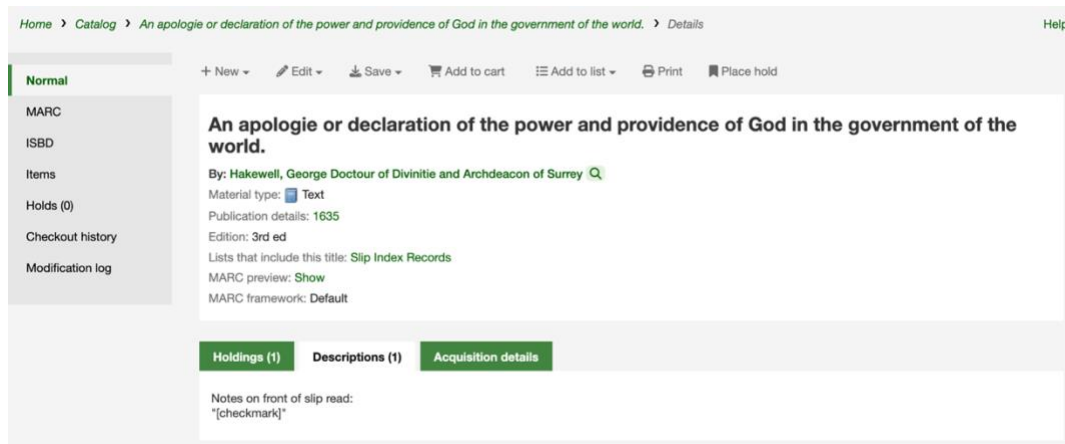


FIGURE 5.7: NORMAL INTERNAL VIEW OF THE BIBLIOGRAPHIC RECORD CREATED FROM FIGURE 5.3

5.2.2 Persona testing results

This section will discuss the results of the persona testing, first broken down by user group persona type and then broken down by the scenario being tested.

5.2.2.1 Patron persona results

Both patron personas approached the scenarios in very similar manners, using similar search phrases and using similar filtering options. The searching done by these personas was performed in the Koha OPAC system, as its search engine allows for much more granularity than Excel's simple Find function, seen in more detail in **Chapter 3**.

Scenario One

The first scenario faced by the two patron personas involved finding any items in the sample collection that are illustrated by a particular illustrator in the 17th through 18th centuries.

As the prompt does not specify an illustrator by name, the first approach taken by both personas was to search for items that are marked as “illustrated” in the catalogue. This was done by entering “illustrated” into the advanced search under the field “Keyword,” with “1600-1799” entered into the “Publication date range” field to narrow the search down to the dates specified in the prompt, as seen in the figure below:

Home > Advanced search

Advanced search

Search for:

	Keyword	illustrated	+ -
and	Keyword		+ -
and	Keyword		+ -

Search Fewer options New search

Item type

Limit to any of the following:

Book

<p>Availability:</p> <p><input type="checkbox"/> Only items currently available for loan or reference</p>	<p>Sort by:</p> <p>Relevance</p>	<p>Publication date range</p> <p>1600-1799</p> <p><small>For example: 1999-2001. You could also use ~1987* for everything published in and before 1987 or "2008*" for everything published in 2008 and after.</small></p>	<p>Language</p> <p>No limit</p>
<p>Audience</p> <p>Any audience</p>	<p>Content</p> <p>Any content</p>	<p>Format</p> <p>Any format</p>	<p>Additional content types for books/printed materials</p> <p>Any</p>

Search Fewer options New search

FIGURE 5.8: INITIAL SEARCH PARAMETERS FOR PATRON PERSONAS IN SCENARIO ONE

This search found zero results, meaning nothing in the catalogue for this sample collection had “illustrated” within the metadata entered into the records and fit within the specified publication date range. Rather than this result stymying the patron personas, they both returned to the advanced search options and added another keyword—“ill”—with the stipulation that it be included in the search as an “or” option instead of an “and”, to catch instances where the full term “illustrated” was not used within the stored metadata. This search also returned zero results.

The final modification to the search terms performed by both personas involved deleting the information entered into the “Publication date range” entirely, now attempting to find any illustrated items at all contained in the sample

collection. This search therefore was comprised of the terms “illustrated” or “ill.” Once again, however, this search returned no results.

At this point, a note was made by Joseph Callow (Photographic Historian Persona, **Appendix F**) as to one way the “Publication date range” field could be manipulated in an additional way to the previous search patterns, and one final modification to the search parameters was made with this persona. This change aligns with the persona’s noted preference for searching online databases prior to stepping foot in the department, compared to the second patron persona’s preference for either reaching out to the librarians prior to arriving at the department and/or browsing the shelves once they arrive. This preference of Mr. Callow’s therefore leads to a better familiarity with online database search engines and their quirks than Dr. Childs’ (Victorian Historian Persona, **Appendix G**).

With this familiarity in mind, the final adjustment consisted of Mr. Callow changing the entry in the “Publication date range” field from “1600-1799” to “-1799,” precipitated by the note underneath the field reading “You could also use “-1987” for everything published in and before 1987” (**Figure 5.8**). Despite this last modification, this final search also returned zero results.

A breakdown of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“illustrated” AND “1600-1799”	16	0	0	0%	0%
“illustrated” OR “ill” AND “1600-1799”	16	0	0	0%	0%
“illustrated” OR “ill”	16	0	0	0%	0%
“illustrated” OR “ill” AND “-1799”	16	0	0	0%	0%
Total²³	16	0	0	0%	0%

TABLE 5.1: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO ONE

After these four search attempts, both personas believed there were no items in the collection that they needed under this scenario and turned their attentions to other collections held elsewhere.

Scenario Two

The patron personas were tasked with searching for items held in the sample collection that had ownership (also known as ex libris) inscriptions handwritten in the items in this scenario.

The first approach taken by both personas was to enter the phrase “ownership inscriptions” into the “Keyword” field, as this would be the natural first search phrase used in response to this scenario. As there were no dates specified in the task instruction, there was no need to fill in any other field to narrow down the search, unlike in the other scenarios faced by the patron personas.

²³ “Total” in this usage refers to the number of *unique* results returned throughout the entire scenario, which means it may not be the mathematical sum of the entire column above.

This search returned zero results, again meaning that none of the metadata recorded within the collection in this case study contains comments about ownership inscriptions using those exact words. Both personas then returned to the advanced search page and searched for the keywords “handwritten inscriptions,” as they both recognized that the specific phrasing used in descriptions of items may use a slightly expanded vocabulary than expected—this understanding comes from both personas’ familiarity with the searching process when using online databases. Once again, this search returned no results.

A final attempt was made to search for any items that may meet this task, this time by using only the word “handwritten” in the “Keyword” field. This search again returned no results, and both personas thus finished this task believing that no items were held in the sample collection that were asked for in this scenario.

A breakdown of the search results and their resulting recall and precision percentages follows:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“ownership inscriptions”	9	0	0	0%	0%
“handwritten inscriptions”	9	0	0	0%	0%
“handwritten”	9	0	0	0%	0%
Total	9	0	0	0%	0%

TABLE 5.2: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO TWO

Scenario Three

For the third scenario, the patron personas were tasked with finding books published in the 17th through 18th centuries about religious history and analysis.

This search process proved to be more involved than the previous two for both personas. The first attempt was done by entering “religious history” into the “Keyword” field, along with “1600-1799” in the “Publication date range” field. This search returned no results, so both personas next attempted to change the information entered into the “Publication date range” field by deleting the date range—leaving the search phrase in the “Keyword” field as-is—to see if that would make any notable difference in the search results. This change did not make a difference, as the search returned no results again.

The next search was done by changing the “Keyword” field to the “Subject” field by way of the dropdown menu seen below:

The screenshot displays an advanced search interface. At the top, there is a breadcrumb trail: "Home > Advanced search". Below this, the section is titled "Advanced search". The search criteria are as follows:

- Search for: Subject (dropdown) with the value "religious history" entered in the text field.
- Operator: "and" (dropdown) followed by Keyword (dropdown) with an empty text field.
- Operator: "and" (dropdown) followed by Keyword (dropdown) with an empty text field.

Buttons for "Search", "Fewer options", and "New search" are located below the search criteria. Below the search criteria, there is a section for "Item type" with a "Limit to any of the following:" section containing a checkbox for "Book".

Below the item type section, there are several filter panels:

- Availability:** A checkbox for "Only items currently available for loan or reference".
- Sort by:** A dropdown menu set to "Relevance".
- Publication date range:** An empty text input field. Below it, a note reads: "For example: 1999-2001. You could also use *-1987* for everything published in and before 1987 or *2008*" for everything published in 2008 and after."
- Language:** A dropdown menu set to "No limit".
- Audience:** A dropdown menu set to "Any audience".
- Content:** A dropdown menu set to "Any content".
- Format:** A dropdown menu set to "Any format".
- Additional content types for books/printed materials:** A dropdown menu set to "Any".

Buttons for "Search", "Fewer options", and "New search" are located below the filter panels.

FIGURE 5.9: SEARCH FIELD CHANGE USED IN SCENARIO THREE

The search term “religious history” was left as-is in the “Subject” field, and the “Publication date range” field was left blank for this attempt. No results were returned for this search either, meaning that either there were no items with that specific subject information contained in their metadata in the sample collection, or there were no items that matched that search regardless of whether that information had been recorded within the metadata or not. The same search was performed again, this time with “1600-1799” in the “Publication date range” field, again returning no results.

Both personas then elected to change the search phrase from “religious history” to the simpler term “religion” in the “Subject” field with “1600-1799” in the “Publication date range” field. This search again returned no results, so the

“Subject” field was changed back to the “Keyword” field, still with the same information in the “Publication date range” field and the search was attempted once more. This search also returned no results.

Finally, the previous search was repeated, this time with no information contained in the “Publication date range” field in an attempt to broaden the search slightly. This search returned one result, seen in the figure below:

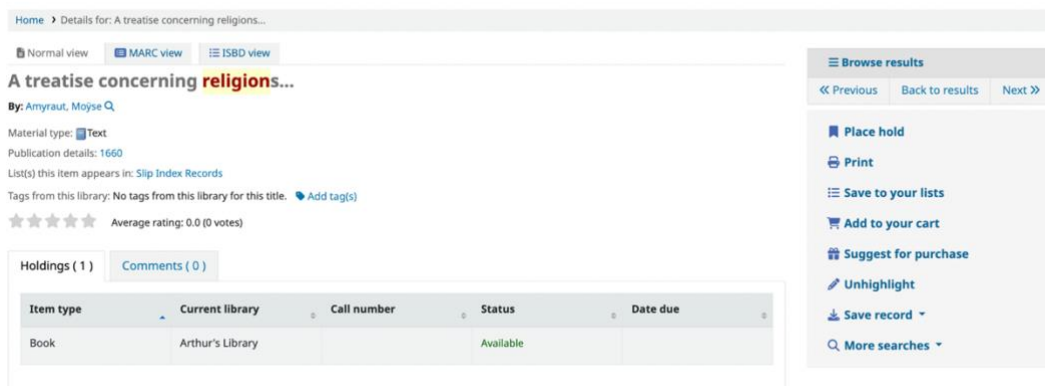


FIGURE 5.10: SOLE SEARCH RESULT RETURNED FOR SCENARIO THREE

Both personas did note the fact that the publication date noted as 1660 mentioned for this item means it theoretically should have been retrieved by the search before this one—where the “religion” was in the “Keyword” field and “1600-1799” was in the “Publication date range” field—but did continue to use the “Publication date range” field for the remainder of this case study and the next one as they both wanted to cover as much area as possible. This note is discussed in more detail in the analysis section at the end of this chapter.

A breakdown of the search results and their resulting recall and precision percentages follows:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“religious history” (keyword) AND “1600-1799”	8	0	0	0%	0%
“religious history” (keyword)	8	0	0	0%	0%
“religious history” (subject)	8	0	0	0%	0%
“religious history” (subject) AND “1600-1799”	8	0	0	0%	0%
“religion” (subject) AND “1600-1799”	8	0	0	0%	0%
“religion” (keyword) AND “1600-1799”	8	0	0	0%	0%
“religion” (keyword)	8	1	1	12.5%	100.0%
Total	8	1	1	12.5%	100.0%

TABLE 5.3: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO THREE

Scenario Four

The fourth scenario given to the patron personas was to find books in the sample collection that were published by the publishing house Elzevir in the 17th through 18th centuries.

As the publishing house was specified in this task, the first search both personas attempted was to use “Elzevir” in the “Keyword” field, as well as “1600-1799” in the “Publication date range” field. This search returned no results, so— remembering the note made in Scenario Three about the appearance of an item

that should have been retrieved by the search form having a publication date range specified but was not—the patrons removed the information in the “Publication date range” field and repeated the search. This search returned one result, seen in the figure below:

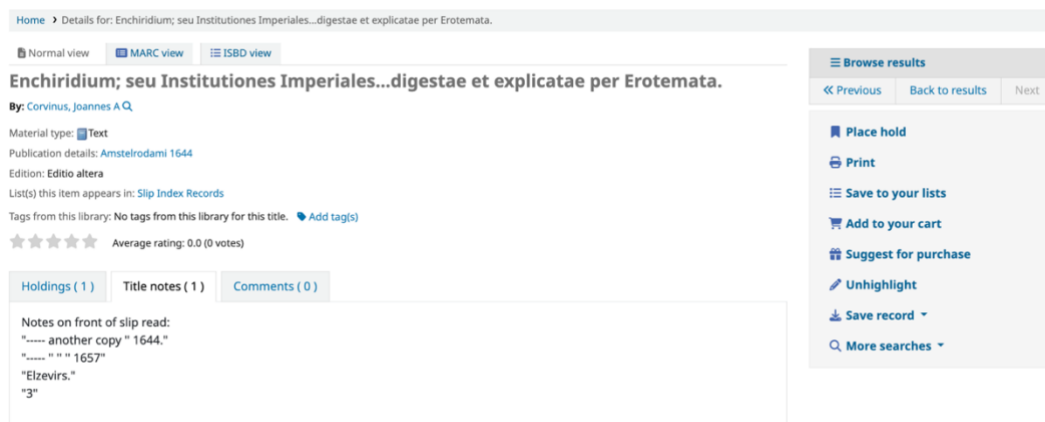


FIGURE 5.11: SOLE SEARCH RESULT RETURNED FOR SECOND SEARCH IN SCENARIO FOUR

Once again it was noted that the previous use of “Publication date range” field had erroneously excluded this item when it should have been included, but an additional note was also made that the mention of Elzevir recorded for this item is a handwritten note on the front of the slip. It was therefore unclear to both personas if this item was definitively published by Elzevir or not, so while it was included as an item that meets the criteria laid out in this scenario, both personas were unsure of the veracity of that information.

Two additional searches were performed by the patron personas in this case. The first search again left the “Publication date range” field blank but changed the “Keyword” field to the “Publisher” field using the same dropdown menu as the

change made in the previous scenario. The search term “Elzevir” was continued in the “Publisher” field. This search resulted in zero returns. The final search was performed with “Elzevir” in the “Publisher” field and “1600-1799” in the “Publication date range” field, which again returned no results.

A breakdown of the search results and their resulting recall and precision percentages is laid out below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“Elzevir” (keyword) AND “1600-1799”	2	0	0	0%	0%
“Elzevir” (keyword)	2	1	1	50.0%	100.0%
“Elzevir” (publisher)	2	0	0	0%	0%
“Elzevir” (publisher) AND “1600-1799”	2	0	0	0%	0%
Total	2	1	1	50.0%	100.0%

TABLE 5.4: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO FOUR

5.2.2.2 Librarian persona results

Akin to the patron personas’ approach to their tasks, the two librarian personas performed their searches for each task in similar manners as each other. The only significant difference between the two librarian personas noted consistently for all six scenarios was that the primary step taken by the junior librarian persona was to approach other members of staff in the department for assistance with starting the searching process. Because the focus of this research is primarily on the level of detail in collected metadata, however, as well as the fact that the junior librarian

persona was still directed to search on the internal site, this step was regarded as insignificant. As noted previously, both personas' searches were performed on the internal site rather than the OPAC, with the results laid out below.

Scenario One

For this scenario, the librarian personas were tasked with aiding a patron interested in books held in the sample collection with illustrations done by a particular illustrator in the 17th through 18th centuries.

Similar to the searching undertaken by the patron personas, the librarians both identified that the likeliest option for a search term to identify items in the sample collection lay in the publication dates (again identified as 1600-1799) and a variation of the word "illustrated." Thus, the first search attempt consisted of using the term "illustrated" in the first "Keyword" field, along with "1600-1799" in the "Year" field. This search returned no results. An additional step taken by the librarian personas here was to test if the inclusion created by toggling the option to limit the item types in the search results to anything labelled a "Book" (which is denoted in the MARC records in 942 \$c as "BK" for all items in the sample collection). This again returned no results.

The third search attempt involved the removal of the date range entered in the "Year" field, which again returned no results. A fourth attempt with the same

parameters as the third plus the toggled “Book” option also returned no results for either persona.

Due to a higher familiarity with the language used by cataloguers at the library, the next search replaced “illustrated” with “ill” in the “Keyword” field, combined with the identified date range in the “Year” field. This returned no results, as did a sixth search with the toggled “Book” option. Removing the date range and repeating the search without the toggled “Book” option did result in two returns. These returns did not, however, contain any information about if the records returned were books containing any illustrations. Instead, as seen in **Figure 5.12** below, the match that the Koha system made was with the presence of the letters “ill” in the title and/or author names for the two results, ultimately leading to a false positive and not a true result.



FIGURE 5.12: SEARCH RESULTS RETURNED FOR SEVENTH SEARCH ATTEMPT

Interestingly, a similar quirk with the Koha search engine appears in the above figure as it did in the patron personas during Scenario Three—the publication dates noted for the two results (1666 and 1607) do fall within the specified date range in

the fifth and sixth search attempts. This therefore means that these two results should also be present in the results for those two search attempts but were not. This quirk was noted by both librarian personas and—as seen in the other scenarios—did have an effect on the search process for the rest of the scenarios.

A final search was performed with the keyword “ill” and the toggled “Book” option, returning the same two results as the previous search, again meaning this search resulted in a false positive rather than a true result.

As the seventh and eighth search terms resulted in two false positives that are not true matches for the prompt, the “Relevant Results Returned” values for those queries in **Table 5.5** (below) remain at zero, leaving both searches with a recall value of 0%.

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“illustrated” AND “1600-1799”	16	0	0	0%	0%
“illustrated” AND “1600-1799” AND “Book”	16	0	0	0%	0%
“illustrated”	16	0	0	0%	0%
“illustrated” AND “Book”	16	0	0	0%	0%
“ill” AND “1600-1799”	16	0	0	0%	0%
“ill” AND “1600-1799” AND “Book”	16	0	0	0%	0%
“ill”	16	2	0	0%	0%
“ill” AND “Book”	16	2	0	0%	0%
Total	16	2	0	0%	0%

TABLE 5.5: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO ONE

Scenario Two

This scenario involved helping a patron wanting to see books with ownership or ex libris inscriptions held within the sample collection.

Again, due to a higher familiarity with the language used by cataloguers, the senior librarian persona immediately started her search with “handwritten annotations” entered into the first “Keyword” field. The junior librarian persona performed this same search shortly after, having had to consult with another member of staff for help on the particular language. This search returned no results, so a second attempt was made with the same keyword phrase and the toggled “Book” option, which again returned no results. No attempt at including a date range in either search was performed, as the scenario does not specify a date the patron was interested in. At this point, both personas performed no more searches, as due to their knowledge of the language used when cataloguing these books, they did not believe there were any more iterations to try.

Table 5.6 below shows the recall and precision of the two searches in this scenario:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“handwritten annotations”	9	0	0	0%	0%

“handwritten annotations” AND “Book”	9	0	0	0%	0%
Total	9	0	0	0%	0%

TABLE 5.6: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO TWO

Scenario Three

This task consisted of aiding a patron interested in books published in the 17th through 18th centuries focused on religious history and analysis.


Once again, as there is a date range specified, every search term will be both accompanied by the phrase “1600-1799” in the “Year” field and not accompanied by that phrase to cover all bases. This was done due to the search engine quirk noted in the first librarian persona scenario in that specifying a date range in the “Year” field may or may not result in the *exclusion* of relevant search results rather than their *inclusion*.

The first search attempt was performed by entering the phrases “religious history” and “analysis” into the first two “Keyword” fields, linked by the “and” Boolean operator option, as well as the specified date range. This search returned no results, as did the same search phrases accompanied by the toggled “Book” option. A third attempt with only the two keyword phrases was performed, again with no results. Using the two keyword phrases and the toggled “Book” option also returned nothing.

Returning to the use of the specified date range, the next step for the librarians was to revert to a much broader subject term than the relatively narrow “religious

history”, as both librarian personas are knowledgeable in the common use of Library of Congress Subject Headings (LCSH). Thus, the two keyword phrases were replaced with the one phrase “religion”. Searches were performed using this keyword with and without the specified date range as well as with and without the toggled “Book” option.

The seventh search attempt, consisting only of the keyword “religion”, returned one result, shown below in **Figure 5.13**. The only note about the subject of this item is that the title includes the word “religion,” which appears to be what the search engine matched with.

2. **A treatise concerning religions...**
by Amyraut, Moÿse.
Material type:  Text; Format: print
Publication details: 1660

[Holds \(0\)](#) | [Add to cart](#) | [Edit record](#) | [Edit items](#)

FIGURE 5.13: SEARCH RESULT RETURNED FOR SEVENTH SEARCH ATTEMPT

Pulling up the MARC record for this item shows no mention of the nature of its contents, as shown in the figure below:

MARC preview

```
000 00373nam a22001457a 4500
003 0St
005 20250628232604.0
008 250414b ||||||| |||| 00| u eng d
024 7 0 _2 Stirling Maxwell Slip Index Upper Left Number
    _a 9281
040  _c al
    _e dcrmb
100  _a Amyraut, Moÿse
    _9 4
245  _a A treatise concerning religions...
260  _c 1660
942  _2 ddc
    _c BK
    _n 0
999  _c 4
    _d 4
```

FIGURE 5.14: MARC RECORD FOR SOLE RESULT FOR SEVENTH SEARCH ATTEMPT

Thus, as the title states this item is on religions, this item was tentatively deemed by the librarian personas to be a relevant search result for the patron, but a note was made regarding the need to perhaps examine the item in further detail to ascertain if it was a true match. Again, the sole result for the seventh search attempt has a publication date that falls within the specified date range and should have been returned in the fifth and sixth attempts as well. An eighth search was performed with the keyword “religion” and the toggled “Book” option, which again returned the same single result as the seventh search.

A final search was performed by changing the “Keyword” field to the “Subject” field and the search with the widest number of results from the previous searches was

performed again—in this case being the subject “religion” without the specified date range or the toggled “Book” option. The reasoning given behind this search being performed in this manner was that using the broader subject term in a search specifically looking for the *subject* of an item instead of an item that has a chosen keyword present anywhere in the record could prove to be a more helpful search. In this case, however, the broad subject search performed here returned no results, and the librarian personas were satisfied with the searches they had performed.

A breakdown of the search results and their resulting recall and precision percentages is laid out below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“religious history” (keyword) AND “analysis” (keyword) AND “1600-1799”	8	0	0	0%	0%
“religious history” (keyword) AND “analysis” (keyword) AND “1600-1799” AND “Book”	8	0	0	0%	0%
“religious history” (keyword) AND “analysis” (keyword)	8	0	0	0%	0%
“religious history” (keyword) AND “analysis” (keyword) AND “Book”	8	0	0	0%	0%
“religion” (keyword) AND “1600-1799”	8	0	0	0%	0%
“religion” (keyword) AND “1600-1799” AND “Book”	8	0	0	0%	0%

“religion” (keyword)	8	1	1	12.5%	100.0%
“religion” (keyword) AND “Book”	8	1	1	12.5%	100.0%
“religion” (subject)	8	0	0	0%	0%
Total	8	1	1	12.5%	100.0%

TABLE 5.7: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO THREE

Scenario Four

The personas were asked to help a patron wanting to examine any books in the sample collection that were published by the publishing house Elzevir in the 17th through 18th centuries.

Similar to the process the two patron personas used in this scenario, the librarians returned to the search engine and changed the first “Keyword” field to “Publisher,” where they input the name “Elzevir.” They also changed the second “Keyword” field to “Publication date (yyyy-yyyy)” and input the specified date range of “1600-1799” in this field. This field was used instead of the “Year” field used in the previous three scenarios to see if the same quirk affecting that field would affect this field as well, which it unfortunately did. This search returned no results. Removing the specified date range also returned no results.

The senior librarian persona was hesitant to continue with the searching process at this point, as while she is more familiar with the library’s holdings overall as well as the cataloguing terms used than the junior librarian persona, she believed a third search would cast too broad a net and result in more work for her further on. She did go ahead with one more search however, in an attempt to be as thorough as possible for the patron she was tasked with assisting.

When changing the “Publisher” field back to “Keyword” and repeating the search with the same search term, both personas were given one result, shown below:



FIGURE 5.15: NORMAL INTERNAL VIEW OF SOLE SEARCH RESULT FOR THIRD SEARCH ATTEMPT

The match the search engine made in the case of this result lies in the item description, where the cataloguer has noted that there are notes on the slip this entry is based on with the word “Elzevirs.” There is no mention of if this relates to a publisher or how exactly the term “Elzevirs” relates to the item, as well as the difficulty noting *which* of the items recorded on the slip this note relates to; another one of the notes on the slip seems to indicate two copies of this item published in different years. Similar to the record returned and discussed in the previous scenario, this record was included in the results as a possible positive result, but both personas noted that the item(s) indicated could not be truly ruled in or out until they were able to physically examine the item(s) represented by this record.

Both personas ended their online search process here, as the third search was deemed so broad as to be inexhaustible. A table recording the searches performed and their results follows below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“Elzevir” (publisher) AND “1600-1799” (publication date)	2	0	0	0%	0%
“Elzevir” (publisher)	2	0	0	0%	0%
“Elzevir” (keyword)	2	1	1	50.0%	100.0%
Total	2	1	1	50.0%	100.0%

TABLE 5.8: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO FOUR

Scenario Five

In this scenario, the personas were approached by a member of the in-house conservation team who asked for a list of items needing urgent conservation, as well as a brief description of the conservation needed for these items.

The search term chosen in this case was once again dictated by the knowledge of terms used by the cataloguing team for this collection. As they were being asked to identify items needing conservation, the first keyword chosen was simply “conservation.” This term, when used *alone*, would theoretically pull up *all* items in the collection, as the cataloguing staff record three levels of conservation needs: does not need conservation, may/might need conservation, and needs conservation. Therefore, a second part of the search phrase needed to be used to only return results that were in the third category, instead of all three at once. For this second portion, the first “and” field was set to “not” and filled with “does not

need” to exclude any results for the term “conservation” that include the exact term “does not need”. It was then deemed necessary to add a third and fourth portion to the phrase to exclude results noted in the “may/might need conservation” category, as the request in this scenario asks for items needing *urgent* conservation—in other words, falling solely into the third category of conservation status. The second and third “and” field was set to “not” and filled with “may need” and “might need” to thus exclude items falling into the second category. This first search returned no results.

The senior librarian persona stopped at this stage, believing there were no items in this sample collection needing urgent conservation—removing any “not” key phrase used would only return results that ultimately did not fit the needs expressed by the conservation staff, so performing another search was deemed extraneous and unnecessary for her.

In contrast, the junior librarian persona performed three more searches, preferring to cast a broader net than the senior librarian persona due to her relative unfamiliarity with the collection’s contents. The second search she performed removed the fourth portion of the initial search phrase, again returning no results. The third search she performed removed the third portion of the search phrase, still with no results. The final search she performed removed the second portion of the initial search phrase as well, which once again returned no results, as noted in the table below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“conservation” NOT “does not need” NOT “may need” NOT “might need” ²⁴	7	0	0	0%	0%
“conservation” NOT “does not need” NOT “may need”	7	0	0	0%	0%
“conservation” NOT “does not need”	7	0	0	0%	0%
“conservation”	7	0	0	0%	0%
Total	7	0	0	0%	0%

TABLE 5.9: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO FIVE

Following this fourth search, both librarian personas were satisfied that they had thoroughly searched the sample collection for any items fitting the needs specified by the conservation team and finished their searches there.

Scenario Six

For this task, the personas were approached by a member of Glasgow Museums staff who asked for a list of items that were in their original bindings and had handwritten annotations for the purpose of displaying these items at Pollok House.

As with the previous scenario, a significant amount of thought needed to be given to the exact search phrase used here. The first portion was deemed to be the same phrase used in the second scenario: “handwritten annotations.” This was chosen as the first portion because the second portion needed to have the Boolean operator “not” for similar reasons as those noted in the previous scenario.

²⁴ This is the sole search for this scenario performed by both personas.

Cataloguing staff use two types of phrases to denote the binding of an item in the sample collection for this case study: “has been rebound” to indicate an item that is no longer in its original binding, and “bound in ____” to indicate the type of binding if it is not immediately clear that the item has been rebound since its original publishing. As neither of these options specifically denotes if an item is in its true, original binding, and one option is a variable phrase, the decision was made to treat items not marked as being rebound as if they were therefore in their original binding. Therefore, the second portion of the search phrase was written to exclude items marked with “has been rebound”, as theoretically only originally bound items would remain in the search results. To do this, the second field was set with the Boolean operator “not” and filled with the phrase “has been rebound”. This search returned no results.

Again, as with the previous scenario, the senior librarian persona stopped her search process here while the junior librarian persona continued with her search by casting a slightly different net. She set the Boolean operator back to “and” and changed the second portion to the phrase “bound in”, which once again returned no results. After this result, the junior librarian persona also halted her search process.

A breakdown of the search results and their resulting recall and precision percentages is laid out below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“handwritten annotations” NOT “has been rebound” ²⁵	3	0	0	0%	0%
“handwritten annotations” AND “bound in”	3	0	0	0%	0%
Total	3	0	0	0%	0%

TABLE 5.10: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO SIX

5.3 Case Study Two

5.3.1 Sample collection

The sample collection examined in this case study consists of the virtual Koha library created using 51 bibliographic records based on the information contained in the test collection Excel spreadsheet created after examination of the items noted in the slip index as being part of the collection. A further discussion of the selection process for this sample collection is found in **Chapter 3**. The 51 records used correspond to the 34 records used for CS1—the difference in the number of records between each sample collection is present because of the prior decision made in the creation of the Excel spreadsheet to catalogue each volume of a multivolume publication in discrete records, rather than the slip index’s and DCRM(B)3’s use of one entry per publication, regardless of if there are more than one volume in the publication.

²⁵ This is again the sole search performed in this scenario by both personas.

Returning to **Figure 5.3** at the beginning of this chapter as an example, further examination of the noted item resulted in the Excel spreadsheet entry²⁶ seen in

Table 5.11 below:

Field Name ²⁷	Data Recorded
ID Number	
Found in Location Book?	Yes
Location Book Number	834425
Location Noted in Location Book	[redacted]
Duplicate? (maybe or blank)	
Current Location	
Author/Editor/Translator	Hakewell, George
Transcribed Author/Editor/Translator	By George Hakewill Doctour of Divinitie and Archdeacon of Surrey
See:	
Title	An apologie or declaration of the pover and providence of God in the government of the world. Consisting in an examination and censvre of the common errovr tovching natvres perpetuall and universall decay, divided into six bookes. Whereof the first treates of this pretended decay in generall, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary bodies, man onely excepted. The third of the pretended decay of mankinde in regard of age and duration, of strength and stature, of arts and wits. The fourth of this pretended decay in matter of manners, together with a large prooffe of the future consummation of the world from the testimonie of the gentiles, and the use which we are to draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression. By George Hakewill Doctour of Divinitie and Archdeacon of Surrey. The third edition revised, and in sundry passages and whole sections augmented by the authour; besides the addition of two entire bookes not formerly published. [pious invocation]
Translated Title	
Number of Volumes	
Transcribed Publication Statement	Oxford, printed by William Turner printer to the famous vniversitie. Anno dom.
Place of Publication	Oxford
Translated Place of Publication	

²⁶ Some information has been redacted upon request by Mitchell Library staff.

²⁷ The meanings of each of these fields is discussed in **Chapter 3**.

Transcribed Publication Date	1635
Publication Date	1635
Language	English; Latin (minor);
Estate Donated From	Maxwell of Polloc
Description	[46], 606p., [10], 378p., [41]; ill. title page; table of contents; indices; printed annotations; tables; handwritten annotations
Notes on Item	preserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "[redacted]" and "834425"; pages from another publication included in binding before title page and after final page; (brown ink) "J Maxwell [???" on page before ill. title page; some sort of security sticker inner back cover; has been rebound? with new spine?; "[redacted]" embossed on front cover; "Apologie" "Hakewill" and "1635" embossed on spine;
Notes on Slip	Pencilled number upper left corner; front (pencilled) checkmark lower left corner
Notes in Location Book	"MAXWELL OF POLLOCK" in location column; author name noted as "Hakewill"
Notes in S.T.C. Catalogue (1964)	HAKEWILL (George) / Second copy [An apologie or declaration of the power and providence of God in the government of the world. Third ed. 2 pts. Oxford, W. Turner (for R. Allott, London) 1635.]. Heraldic bookplate, Maxwell of Polloc. 'J. Maxwell' / [redacted]
Conservation Needed?	Maybe
Found in RBC?	Yes
Upper Left Number	8911

TABLE 5.11: EXCEL SPREADSHEET ENTRY FOR ITEM NOTED IN FIGURE 5.3

Once this information was fully collected, this information was translated into the applicable MARC fields as previously explained in **Chapter 4** and entered into the Koha system. The record for the item described in **Table 5.11** then appears as the MARC record seen below in **Figure 5.16**:

Normal view MARC view ISBD view

An apologie or declaration of the povver and providence of God in the government of the world. (Record no. 75)

[view labeled]

```

000 04414nam a22002894a 4500
003 05t
005 20250512133250.0
008 250422b ||||| |||| 00] u eng d
024 7 0_25t r Ling Maxwell Slip Index Upper Left Number
      _a0911
024 7 0_2 Mitchell Library Location Books
      _a834425
040 _cal
      _a dc rmb
100 _aHakewell, George
      _cDoctour of Divinitie and Archdeacon of Surrey
      _926
245 _aAn apologie or declaration of the povver and providence of God in the government of the world.
      _bConsisting in an examination and censure of the common error touching natures perpetuall and universall decay, divided into six bookes. Whereof the first treats of this
      pretended decay in general, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary
      bodies, man only excepted. The third of the pretended decay of mankind in regard of age and duration, of strength and stature, of arts and wits. The fourth of this
      pretended decay in matter of manners, together with a large prooffe of the future consummation of the world from the testimonie of the gentiles, and the use which we are to
      draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression.
      _cBy George Hakewill Doctour of Divinitie and Archdeacon of Surrey.
246 _aAn apology or declaration of the power and providence of God in the government of the world.
      _bConsisting in an examination and censure of the common error touching natures perpetual and universal decay, divided into six bookes. Whereof the first treats of this
      pretended decay in general, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary
      bodies, man only excepted. The third of the pretended decay of mankind in regard of age and duration, of strength and stature, of arts and wits. The fourth of this
      pretended decay in matter of manners, together with a large proof of the future consummation of the world from the testimony of the gentiles, and the use which we are to
      draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression.
250 _aThe third edition revised,
      _band in sundry passages and whole sections augmented by the authour; besides the addition of two entire bookes not formerly published.
260 _aOxford
      _bTurner, William
      _c1635
300 _a[46], 606p., [10], 378p., [41]
      _bill title page, table of contents, indices, printed annotations, tables, handwritten annotations
500 _aPublication imprint reads: "Oxford, printed by William Turner printer to the famous unversitie. Anno dom. 1635"
546 _aEnglish, Latin (minor)
      _bRoman alphabet
556 _aAssociated slip in slip index has added notes: pencilled number upper left corner; front (pencilled) checkmark lower left corner Associated entry in location books has
      added notes: "MAXWELL OF POLLOCK" in location column; author name noted as "Hakewill" Associated entry in S.T.C. Catalogue (1964) reads: "HAKEWILL (George) / Second copy
      [An apologie or declaration of the power and providence of God in the government of the world. Third ed. 2 pts. Oxford, W. Turner (for R. Allott, London) 1635.]. Heraldic
      bookplate, Maxwell of Polloc. 'J. Maxwell' / S.T.C. 12613 (292)"
562 _aPreserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "STC 12613fo" and "834425"; pages from
      another publication included in binding before title page and after final page; (brown ink) "J Maxwell [???]" on page before ill. title page; some sort of security sticker
      inner back cover; has been rebound? with new spine?; "STC 12613" embossed on front cover; "Apologie" "Hakewill" and "1635" embossed on spine;
      _bPreserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "STC 12613fo" and "834425"
563 _ahas possibly been rebound with new spine
700 _aTurner, William
      _cprinter to the famous unversitie
      _eprinter
      _996
942 _2ddc
      _cBK
      _nb
999 _c75
      _d75

```

FIGURE 5.16: MARC RECORD DISPLAY IN OPAC FOR ITEM DESCRIBED IN TABLE 5.11

This MARC record then results in the following OPAC display that is visible to patrons:

[Normal view](#)
[MARC view](#)
[ISBD view](#)

An apologie or declaration of the povver and providence of God in the government of the world. Consisting in an examination and censure of the common error touching natvres perpetuall and universall decay, divided into six bookes. Whereof the first treates of this pretended decay in generall, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary bodies, man onely excepted. The third of the pretended decay of mankinde in regard of age and duration, of strength and stature, of arts and wits. The fourth of this pretended decay in matter of manners, together with a large proofe of the future consummation of the world from the testimonie of the gentiles, and the use which we are to draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression. By George Hakevill Doctour of Divinitie and Archdeacon of Surrey.

By: [Hakewell, George Doctour of Divinitie and Archdeacon of Surrey](#)

Contributor(s): [Turner, William printer to the famous univrsitie](#) [\[printer\]](#)

Material type: [Text](#)

Publication details: [Oxford Turner, William 1635](#)

Edition: [The third edition revised, and in sundry passages and whole sections augmented by the authour; besides the addition of two entire bookes not formerly published](#)

Description: [\[46\], 606p., \[10\], 378p., \[41\] ill. title page, table of contents, indices, printed annotations, tables, handwritten annotations](#)

Other title:

[An apology or declaration of the power and providence of God in the government of the world. Consisting in an examination and censure of the common error touching natvres perpetual and universal decay, divided into six bookes. Whereof the first treat of this pretended decay in general, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary bodies, man onely excepted. The third of the pretended decay of mankind in regard of age and duration, of strength and stature, of arts and wits. The fourth of this pretended decay in matter of manners, together with a large proofe of the future consummation of the world from the testimony of the gentiles, and the use which we are to draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression](#)

Lists) this item appears in: [Spreadsheet Records](#)

Tags from this library: No tags from this library for this title. [Add tag\(s\)](#)

Average rating: 0.0 (0 votes)

[Holdings \(1\)](#)
[Title notes \(5\)](#)
[Comments \(0\)](#)

Publication imprint reads:

 "Oxford, printed by William Turner printer to the famous univrsitie. Anno dom. 1635"

English, Latin (minor) Roman alphabet

Associated slip in slip index has added notes:

 pencilled number upper left corner; front (pencilled) checkmark lower left corner

Associated entry in location books has added notes:

 "MAXWELL OF POLLOCK" in location column; author name noted as "Hakewill"

Associated entry in S.T.C. Catalogue (1964) reads:

 "HAKEWILL (George) / Second copy [An apologie or declaration of the power and providence of God in the government of the world. Third ed. 2 pts. Oxford, W. Turner (for R. Allott, London) 1635.]. Heraldic bookplate, Maxwell of Polloc. J. Maxwell / S.T.C. 12613 (292)"

preserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "STC 12613fo" and "834425"; pages from another publication included in binding before title page and after final page; (brown ink) J Maxwell [???] on page before ill. title page; some sort of security sticker inner back cover; has been rebound? with new spine?; "STC 12613" embossed on front cover; "Apologie" "Hakewill" and "1635" embossed on spine; preserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "STC 12613fo" and "834425" has possibly been rebound with new spine

FIGURE 5.17: NORMAL OPAC VIEW OF THE BIBLIOGRAPHIC RECORD CREATED FROM TABLE 5.11

On the internal site, the normal display for the bibliographic record is as seen in the figure below:

Home > Catalog > An apology or declaration of the power and providence of God in the government of the world. Consisting in an examination and censure of the common error touching natvres perpetual and universal decay, divided into six books. Whereof the first treats of this pretended decay in general, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary bodies, man only excepted. The third of the pretended decay of mankind in regard of age and duration, of strength and stature, of arts and wits. The fourth of this pretended decay in matter of manners, together with a large proove of the future consummation of the world from the testimonie of the gentiles, and the use which we are to draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression. > Details

Normal MARC ISBD Items Holds (0) Checkout history Modification log

→ New → Edit → Save → Add to cart → Add to list → Print → Place hold

An apology or declaration of the power and providence of God in the government of the world. Consisting in an examination and censure of the common error touching natvres perpetual and universal decay, divided into six books. Whereof the first treats of this pretended decay in general, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary bodies, man only excepted. The third of the pretended decay of mankind in regard of age and duration, of strength and stature, of arts and wits. The fourth of this pretended decay in matter of manners, together with a large proove of the future consummation of the world from the testimonie of the gentiles, and the use which we are to draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression. By George Hakevill Doctour of Divinitie and Archdeacon of Surrey.

By: [Hakevill, George Doctour of Divinitie and Archdeacon of Surrey](#) [Q](#)
 Contributor(s): [Turner, William printer to the famous universitie](#) [Q](#)
 Material type: [Text](#)
 Publication details: Oxford: Turner, William 1635
 Edition: The third edition revised, and in sundry passages and whole sections augmented by the author, besides the addition of two entire bookes not formerly published
 Description: [46], 606p., [10], 378p., [4] il. title page, table of contents, indices, printed annotations, tables, handwritten annotations
 Other title: An apology or declaration of the power and providence of God in the government of the world. Consisting in an examination and censure of the common error touching natvres perpetual and universal decay, divided into six books. Whereof the first treats of this pretended decay in general, together with some preparatives thereunto. The second of the pretended decay of the heavens and elements, together with that of the elementary bodies, man only excepted. The third of the pretended decay of mankind in regard of age and duration, of strength and stature, of arts and wits. The fourth of this pretended decay in matter of manners, together with a large proove of the future consummation of the world from the testimonie of the gentiles, and the use which we are to draw from the consideration thereof. The fifth and sixth are spent in answering objections made since the second impression
 Action note: may need conservation
 Lists that include this title: [Spreadsheet Records](#)
 MARC preview: [Show](#)
 MARC framework: [Default](#)

Holdings (1) **Descriptions (7)** **Acquisition details**

Publication imprint reads:
 "Oxford, printed by William Turner printer to the famous universitie. Anno dom. 1635"
 Sir William Stirling Maxwell, 9th Baronet's son, Sir John Maxwell Stirling-Maxwell, 10th Baronet donation
 English, Latin (mixed) Roman alphabet
 Associated slip in slip index has added notes:
 pencilled number upper left corner; front (pencilled) checkmark lower left corner
 Associated entry in location books has added notes:
 "MAXWELL OF POLLOCK" in location column; author name noted as "Hakevill"
 Associated entry in S.T.C. Catalogue (1964) reads:
 "HAKEWILL (George) / Second copy [An apology or declaration of the power and providence of God in the government of the world. Third ed. 2 pts. Oxford, W. Turner (for R. Allott, London) 1635.]. Heraldic bookplate, Maxwell of Polloc. J. Maxwell / S.T.C. 12613 290?"
 preserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "STC 12613b" and "B3442S"; pages from another publication included in binding before title page and after final page; (brown ink) "J Maxwell [??]" on page before il. title page; some sort of security sticker inner back cover; has been rebound with new spine"; "STC 12613" embossed on front cover; "Apologie" "Hakevill" and "1635" embossed on spine; preserved bookplate inner cover with crest and printed "Maxwell of Polloc" and handwritten "B.1.14"; (pencilled) inner cover "STC 12613a" and "B3442S"
 has possibly been rebound with new spine
 may need conservation

FIGURE 5.18: NORMAL INTERNAL VIEW OF THE BIBLIOGRAPHIC RECORD CREATED FROM TABLE 5.11

When comparing the level of detail present in the MARC record shown in **Figure 5.16** to that of the MARC record created only with the slip index information shown in **Figure 5.5**, it is clear that there is significantly more detail in the later version. This detail is not only present within the same fields as used by the former version (i.e., the title, publication information, etc.) but can also be seen in the presence of MARC fields not found in the earlier, simpler record created for CS1. The impact of this increase in granularity is shown throughout the persona testing discussed in the following section and discussed in even further detail in the following chapter.

5.3.2 Persona testing results

This section will discuss the results of the persona testing in this second case study, organised in a similar manner to the results of CS1 laid out previously.

5.3.2.1 Patron persona results

Once again, both patron personas approached the four scenarios in strikingly similar manners as each other, with any variations being noted in the scenario(s) in which they occurred. The personas were instructed to approach the scenarios in this case study as if they had not previously performed the searches for the same scenarios in the previous case study. As in CS1, the searching conducted by the patron personas was conducted in the OPAC for the same reasons as the previous case study, referencing **Figure 5.1** throughout.

Scenario One

This scenario presented the patron personas with a request to find books held in the Stirling Maxwell sample collection that were illustrated by a particular illustrator and published in the 17th-18th centuries.

The first step completed by both personas was to break down the prompt into concise search words and/or phrases. In this case, the most important points they mentioned were:

1. The books being searched for needed to be illustrated, and
2. The books needed to be published between the 17th and 18th centuries.

Given that the wording of the prompt requested books illustrated “by a particular illustrator” but that no illustrator was explicitly named, a note was made here that the personas would both normally include the illustrator’s name in their searches to potentially further narrow the results down. However, as no name was given in the prompt, this could not be done.

Therefore, the first search attempt used the word “illustrated” in the “Keyword” field, and the dates “1600-1799” in the “Publication date range” field within the advanced search page (**Figure 5.1**). This search returned no results.

The information in the “Keyword” field was left as-is, while the date range information was changed from “1600-1799” to “-1799,” once again noting the phrasing below the field (**Figure 5.1**). This also returned no results.

The third search only utilised the word “illustrated” in the “Keyword” field with no publication date range information, again returning no results.

Both personas then included a second keyword in their search attempts using the Boolean operator “or” to link them in the search options, this time using the abbreviation “ill” to potentially find any results. This combination (“illustrated” OR “ill”) was repeated with the same publication date combinations as the first three search attempts. The first two attempts each returned no results, but the third attempt with this combination of keywords returned 28 results.

Due to the use of the abbreviation “ill”, both personas did need to comb through these results to potentially eliminate any false positive results that did not specifically state the items on record were illustrated or were not published within the requested date range. After this double checking was completed, four false positive results were noted and removed, leaving the final tally at 24 confirmed relevant items. Both personas noted that the four items deemed false positives had the letters “ill” either in the title of the items or in the author’s name, explaining the reason the search engine had added them to the total number of results.

Again, the personas noted that the 24 confirmed items in the final search did have publication dates falling within the previously specified date range used in the prior five search attempts, yet it was only when this range was removed from the search options that Koha’s engine returned any results whatsoever. This was again noted as a quirk of the system for both personas for the remaining three scenarios.

A breakdown of the search results and their resulting recall and precision percentages follows:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“illustrated” AND “1600-1799”	24	0	0	0%	0%
“illustrated” AND “-1799”	24	0	0	0%	0%
“illustrated”	24	0	0	0%	0%
“illustrated” OR “ill” AND “1600-1799”	24	0	0	0%	0%

“illustrated” OR “ill” AND “-1799”	24	0	0	0%	0%
“illustrated” OR “ill”	24	28	24	100.0%	85.71%
Total	24	28	24	100.0%	85.71%

TABLE 5.12: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO ONE

Scenario Two

For this scenario, the patron personas were asked to find items within the collection that have ownership (or ex libris) inscriptions while ignoring any other forms of handwritten annotations.

Both personas landed on the keyword phrase “ownership inscriptions” as the first search term to use in the “Keyword” field of the search engine, and as no publication dates were specified for this scenario, this field was left blank. This search produced no results, so a slight change to the search phrase was made. This change removed the word “inscriptions” to leave the search word as “ownership”, as this was deemed the most important descriptor out of the two words used previously. This search returned a single result, depicted below in

Figure 5.19:

Normal view MARC view ISBD view

C. Julii Caesaris quae exstant ex viri docti accuratissima recognitione accedit nunc vetus interpres Craecus librorum VII de bello Gallico, ex bibliotheca P. Petavii Praeterea notae, adnotationes, commentarii partim veteres, partim novi. Adhaec indices rerum, et locorum utiles. Quae altera pagina plenius exhibet...Nunc auctior et comtior. Editio olim adornata opera et studio Gothofredi Jungermanni, Lipsiensis.

By: [Caesar, Caius Julius](#)

Contributor(s): [Jungermann, Gottfried](#) | [Zunner, Johann David \[publisher\]](#) | [Humm, Paul \[printer\]](#)

Material type: Text

Publication details: [Frankfurt Zunner, Johann David Humm, Paul 1669](#)

Description: [12], 450p., [24], 112 columns, [8], 1050 columns, [11], 3 leaves of plates: map ill., maps, index, ill. title page, handwritten annotations

Other title:
C. Julii Caesaris quae exstant ex viri docti accuratissima recognitione accedit nunc vetus interpres Craecus librorum VII de bello Gallico, ex bibliotheca P. Petavii Praeterea notae, adnotationes, commentarii partim veteres, partim novi. Adhaec indices rerum, et locorum utiles. Quae altera pagina plenius exhibet...Nunc auctior et comtior

List(s) this item appears in: [Spreadsheet Records](#)

Tags from this library: No tags from this library for this title. [Add tag\(s\)](#)

★ ★ ★ ★ ★ Average rating: 0.0 (0 votes)

Holdings (1) Title notes (5) Comments (0)

Publication imprint reads:
"M DC LXXIX. Fracofvrti sumptibus Johannis Davidis Zunneri. Typis Pauli Hummii."

Latin, Greek; Roman alphabet, Greek alphabet

Associated slip in slip index has added notes:
pencilled number upper left corner; back (pencilled) "Jungermann (Gottfried)"

bookplate inner cover with crest and printed "Maxwell of Pollok, Bart." and handwritten "T.2.6."; bookplate attached across from bookplate with Glasgow coat of arms and printed "Glasgow Corporation Public Libraries The Mitchell Library"; (pencilled) "Arms of Amsterdam on cover a school pr[i]z[os?]e" above bookplate, two title pages; inner page entirely in Latin with handwriting suggesting ownership in Amsterdam? bookplate inner cover with crest and printed "Maxwell of Pollok, Bart." and handwritten "T.2.6."

original binding

FIGURE 5.19: SOLE SEARCH RESULT RETURNED FOR SECOND ATTEMPT IN SCENARIO TWO

The reason this result was returned for this attempt lies in the title notes seen in the figure, where it is noted that there is “handwriting suggesting ownership in Amsterdam?” (Figure 5.19). There is no note regarding the person(s) who may have inscribed this however, so both personas deemed this result to be a tentative positive result that required examination of the physical item itself, as it could not be immediately ruled in or out as a positive result for the given scenario.

A third search attempt was performed using the specific phrase “ex libris” instead of the previous two attempts. Both personas explained this choice as being made because they were unsure of how such inscriptions would be noted in the

catalogue, and they wanted to cover their bases. This search returned four results, one of which is pictured below:

The screenshot shows a library catalog record for the title "Justini historiarum...libri XLIV. ex Trogo Pompeio". At the top, there are tabs for "Normal view", "MARC view", and "ISBD view". The author is listed as "By: Trogus Pompeius". Below that, contributors are listed: "Contributor(s): Tonson, Jacob [printer] | Watts, John [printer] | Maittaire, Michael [editor]". The material type is "Text". Publication details are "London Tonson, Jacob Watts, John 1713". The description is "[10], 281p., [50] col. ill. title page, indices, table of contents, handwritten annotations". It also mentions "List(s) this item appears in: Spreadsheet Records". There are no tags from the library for this title. The average rating is 0.0 (0 votes). Below the main record, there are tabs for "Holdings (1)", "Title notes (5)", and "Comments (0)". The "Title notes" tab is selected, showing several notes. The first note is "Publication imprint reads: 'Londini: ex officinâ Jacobi Tonson, & Johannis Watts. M DCC XIII. Cum privilegio.'" The second note is "Latin Roman alphabet". The third note is "Associated slip in slip index has added notes: pencilled number upper left corner; front (pen) '-----another copy', (pencilled) '2' upper right corner". The fourth note is "Associated location book entry has added notes: noted in cancelled[?] section of 'Sculer Collection' in LB90; DOR noted as '1/5/78'". The fifth note is a long, detailed note about the bookplate and spine, mentioning "Glasgow coat of arms", "Glasgow Corporation Public Libraries The Mitchell Library", "Y122", "899124", "1/5/78", "Maxwell of Pollok, Bart.", "Bart.", "83[circled] 13", "Trogo Pompeius B.M.C.", "Trogi Historiarum", and "1713". The final note is "has been rebound".

FIGURE 5.20: ONE RESULT RETURNED IN THIRD ATTEMPT FOR SCENARIO TWO

As seen in **Figure 5.20**, the match made by the Koha search engine lies within the title notes for the item. Herein is a note of a handwritten note reading “Ex Libris Joannis Maxwell” (**Figure 5.20**), which is the exact type of item requested by the scenario prompt. Examining the other three results returned for this attempt showed that all four results were indeed positive results and fit the requirements given to the personas.

A final search attempt was performed using the phrase “handwritten annotations” in place of the previously used keywords. The personas noted that the reason this

search phrase was used was to identify any items held in the sample collection with any handwritten notes at all, as there may be some items that are noted as having handwritten annotations but no specific ownership or ex libris inscriptions noted in those exact words. This search returned 31 results, but again the results required further examination to ascertain if any were false positives or needed physical examination to definitively rule them in or out. Examining the title notes for each of the 31 results given for this search attempt identified that 21 results had no mention of any sort of ownership notations—the handwritten annotations recorded were described more as notes than any sort of declaration of ownership—with the remaining ten results having handwritten notes that fit the requirements of this scenario’s prompt. The ten results did include the results recorded in attempts two and three, as seen in the table below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“ownership inscriptions”	10	0	0	0%	0%
“ownership”	10	1	1	10.0%	100.0%
“ex libris”	10	4	4	40.0%	100.0%
“handwritten annotations”	10	31	10	100.0%	32.26%
Total	10	31	10	100.0%	32.26%

TABLE 5.13: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO TWO

Scenario Three

This scenario asked the patron personas to look for books in the sample collection that were on religious history and analysis and published in the 17th-18th centuries.

Both personas quickly identified that the key search phrases to use for the initial search were the phrases “religious history” and the given publication date range of “1600-1799”, as these two terms would interact with enough specificity to potentially produce any item(s) that fit the prompt specifications. This search returned no results, and—having remembered the earlier quirk with the publication date range noted in Scenario One—both personas repeated the search with the publication date range information removed. This again returned no results. A third attempt was made with the same keyword, this time with the publication date range noted as “-1799”, which again produced no results.

The next three search attempts were performed with the “Keyword” field option changed to a “Subject” field, as seen in **Figure 5.9**, and the same combinations of search phrase and publication date ranges. Each search produced no results.

The next searches changed the search phrase from “religious history” to “religion”, as both personas have seen this term used to describe the subject of items held elsewhere that fit the needs of the prompt they were given. Three subject searches were performed with this phrase and the three variations on the date range, each returning no results.

In a final attempt to find any items suiting this scenario, the personas changed the search types from subject searches back to keyword searches, and repeated the same combination of the keyword “religion” and the three options for publication

date ranges, each with varying results. The first of these three, with the keyword “religion” and date range “1600-1799” returned no results, as did the search with the same keyword and date range “-1799”. By contrast, the final search—this time only using the keyword “religion”—returned two results.

A closer examination of the two results returned for this final search revealed that the match made by Koha’s search engine was a simple match with the word “religion” in both titles, as seen in the following two figures:

Normal view MARC view ISBD view

L'estat de l'Eglise, avec le discours des temps depuis les apostres iusques au present. Augmenté & reueu tellement en cest edition, que ce qui concerne le siege romain, & autres royaumes depuis l'Eglise primitiue iusques à ceux qui regnent auiourdhuy, y est en brieues annales proposé. Item vn traité de la religion & republique des iuifs, depuis le retour de l'exil de Babylone, iusques au dernier saccagement de Ierusalem. Iean Crespin, a l'Eglise de Iesvs Christ, S.

By: [Crespin, Jean](#)

Contributor(s): [Vignon, Eustace \[publisher\]](#)

Material type: [Text](#)

Publication details: [\[s.l.\] Vignon, Eustace 1581](#)

Description: 80p. in various pagings, 583p. index, table of names, printed annotations

Other title:
L'estat de l'Eglise, avec le discours des temps depuis les apostres jusques au present. Augmenté & revue tellement en cest edition, que ce qui concerne le siege romain, & autres royaumes depuis l'Eglise primitive jusques à ceux qui regnent aujourd'hui, y est en briefes annales proposé. Item un traité de la religion & republique des juifs, depuis le retour de l'exil de Babylone, jusques au dernier saccagement de Jerusalem

List(s) this item appears in: [Spreadsheet Records](#)

Tags from this library: No tags from this library for this title. [Add tag\(s\)](#)

★ ★ ★ ★ ★ Average rating: 0.0 (0 votes)

Holdings (0) Title notes (5) Comments (0)

Publication imprint reads:
"M. D. LXXXI. Chez Evstace Vignon"

French Roman alphabet

Associated slip in slip index has added notes:
pencilled number upper left corner

bookplate inner cover with Glasgow coat of arms and printed "Glasgow Corporation Public Libraries The Mitchell Library" and (pencilled) "Y138"; preserved bookplate with crest and printed "Maxwell of Pollok, Bart." and handwritten "B.8.10" and (pencilled) handwritten "83[circled] 282"; (brown ink) "D__[r/v?_]7" on reverse of title page; (pencilled) "Crespin (Jean) B.M.C." on reverse of title page; has been rebound; "L'Etat de L'Eglise" "Crespin" and "1581" embossed on spine; preserved bookplate with crest and printed "Maxwell of Pollok, Bart." and handwritten "B.8.10" and (pencilled) handwritten "83[circled] 282"

has been rebound

FIGURE 5.21: FIRST RESULT RETURNED FOR FINAL SEARCH IN SCENARIO THREE

Normal view MARC view ISBD view

A treatise concerning religions, in refutation of the opinion which accounts all indifferent. Wherein is also evinc'd the necessity of a particular revelation, and the verity and preeminence of the Christian religion above the Pagan, Mahometan, and Jewish rationally demonstrated. Rendred into English out of the French copy of Moyes Amyraldvs late Professor of Divinity at Saumur in France.

By: [Amyraut, Moÿse](#)

Contributor(s): [Simons, M](#) [printer] | [Nealand, Will](#) [bookseller]

Material type: Text

Publication details: [London Simons, M. 1660](#)

Description: [20], 543p. table of contents, four unnumbered pages inserted between p.510 and p.511

List(s) this item appears in: [Spreadsheet Records](#)

Tags from this library: No tags from this library for this title. Add tag(s)

★ ★ ★ ★ ★ Average rating: 0.0 (0 votes)

Holdings (1) Title notes (5) Comments (0)

Publication imprint reads:
"1660. London printed by Will.Nealand book-seller in Cambridge and are to be sold there and at the sign of the Crown in Duck-lane."

English Roman alphabet

Associated slip in slip index has added notes:
pencilled number upper left corner

Associated location book entry has added notes:
author name noted as "Amyraut, Moÿse"; "MAXWELL OF POLLOCK" in location column

bookplate inner cover with Glasgow coat of arms and printed "Glasgow Corporation Public Libraries The Mitchell Library" and (pencilled) handwritten "A3037" and class number "834584"; carbon paper being used as bookmark; preserved bookplate with crest and printed "Maxwell of Pollok, Bart." and (pencilled) handwritten "374[circled] 4709"; (brown/black ink) "D__[r/v?].9" on reverse of title page; Mitchell library stamp with (pencilled) "18" [Initials?]" and "290AMY" on reverse of title page; (pencilled) "Amyraut (Moÿse) B.M.Cata." on reverse of title page; has been rebound; "A Treatise Concerning Religions" and "1660" embossed on spine; bookplate inner cover with Glasgow coat of arms and printed "Glasgow Corporation Public Libraries The Mitchell Library" and (pencilled) handwritten "A3037" and class number "834584"; carbon paper being used as bookmark; preserved bookplate with crest and printed "Maxwell of Pollok, Bart." and (pencilled) handwritten "374[circled] 4709"

has been rebound

FIGURE 5.22: SECOND RESULT RETURNED FOR FINAL SEARCH IN SCENARIO THREE

The first of the two items, *L'estat de l'Eglise*, is noted as being published in 1581 (Figure 5.21), which sets it firmly outside of the date range specified by the prompt, so this result was eliminated as a false positive. While these are detailed records for each item, the lack of LCSH applied to the two items means that this final result was deemed only a tentatively positive result for this scenario, as examination of the physical book and its contents would be needed to either rule it as a true positive or rule it out as a false positive.

A table of the search results and their resulting recall and precision percentages follows:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“religious history” (keyword) AND “1600-1799”	10	0	0	0%	0%
“religious history” (keyword)	10	0	0	0%	0%
“religious history” (keyword) AND “-1799”	10	0	0	0%	0%
“religious history” (subject) AND “1600-1799”	10	0	0	0%	0%
“religious history” (subject) AND “-1799”	10	0	0	0%	0%
“religious history” (subject)	10	0	0	0%	0%
“religion” (subject) AND “1600-1799”	10	0	0	0%	0%
“religion” (subject) AND “-1799”	10	0	0	0%	0%
“religion” (subject)	10	0	0	0%	0%
“religion” (keyword) AND “1600-1799”	10	0	0	0%	0%
“religion” (keyword) AND “-1799”	10	0	0	0%	0%
“religion” (keyword)	10	2	1	10.0%	50.0%
Total	10	2	1	10.0%	50.0%

TABLE 5.14: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO THREE

Scenario Four

The final scenario presented to the patron personas for CS2 was to find any items published by the publishing house Elzevir in the 17th-18th centuries.

The first search attempt was to use the keyword “Elzevir” with the specified date range, which produced no results. Again, the date range was changed from “1600-

1799” to “-1799” and repeated, again producing no results. When the date range was completely removed however, the search returned three results, one of which is shown in the figure below:

The screenshot shows a library catalog record for the book "Enchiridivm; seu institutiones imperiales, insertis latioribus materiis, theoricè & practicè digestae, & explicatae per erotemata. Johannis Arn. Corvini JC." The record is displayed in a clean, modern interface with tabs for "Normal view", "MARC view", and "ISBD view". The title is prominently displayed in a bold, dark font. Below the title, the author "By: Corvinus, Joannes A Q" and the contributor "Contributor(s): Elzevir [publisher]" are listed. The material type is identified as "Text". Publication details include "Amsterdam Elzevir 1644". The edition is noted as "Editio altera, emendator, & uberioribus additionibus auctior". The description is "[20], 714p. index, handwritten annotations". The other title is "Enchiridivm; seu institutiones imperiales, insertis latioribus materiis, theoricè & practicè digestae, & explicatae per erotemata". The record also shows it appears in "Spreadsheet Records", has no tags from the library, and an average rating of 0.0 (0 votes). At the bottom, there are tabs for "Holdings (1)", "Title notes (5)", and "Comments (0)". The "Title notes" tab is selected, showing a detailed description of the book's physical characteristics and associated notes, including the publication imprint "Amstelrodami, apud Ludovicum Elzevirium, c15 15C XLIV.", the language "Latin; Roman alphabet", and various handwritten and printed markings on the bookplate and spine.

FIGURE 5.23: ONE RESULT FOR THIRD SEARCH IN SCENARIO FOUR

As shown in **Figure 5.23**, the item is clearly marked as being published by Elzevir, easily marking this as a true positive result in this search. The other two results for this search are marked in the same manner, again meaning they too are true positive results.

Out of thoroughness, both personas additionally changed the “Keyword” field to the “Publisher” field and repeated the search without the publication date range, again returning the same three results as the third search attempt.

A table of the search results and their resulting recall and precision percentages follows:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“Elzevir” (keyword) AND “1600-1799”	3	0	0	0%	0%
“Elzevir” (keyword) AND “-1799”	3	0	0	0%	0%
“Elzevir” (keyword)	3	3	3	100.0%	100.0%
“Elzevir” (publisher)	3	3	3	100.0%	100.0%
Total	3	3	3	100.0%	100.0%

TABLE 5.15: RECALL AND PRECISION OF PATRON SEARCH RESULTS FOR SCENARIO FOUR

5.3.2.2 Librarian persona results

This section will discuss the search results for the two librarian personas in each scenario laid out at the beginning of this chapter. Similar to the patron personas in the previous section, the librarian personas were instructed to ignore any notes made during CS1, thus approaching this case study with a blank slate. These personas utilised Koha’s internal search engine for their six scenarios, referencing **Figure 5.2** throughout.

Overall, both personas approached each scenario in a similar manner and utilised similar search terms when looking for the items specified in each scenario, with

any differences noted when they arise. Generally, the junior librarian persona did need to begin her search by consulting more senior staff, due to her relative unfamiliarity with the cataloguing verbiage used, but would follow that step by searching similarly to the senior librarian persona.

Scenario One

The first scenario faced by the librarian personas was a patron request for books held in the sample collection that were illustrated by a particular illustrator and published in the 17th-18th centuries.

As no illustrator was named specifically, neither persona noted the ability to use the illustrator's name as a keyword in the internal search engine. If they had been able to do so, they also noted that this would have been their first search, as both of them are familiar with the cataloguing vocabulary used in the records. Instead, the two search terms chosen to use in Koha were the specified publication date range ("1600-1799") and some variation of the word "illustrated." Using those two terms exactly returned no results, as did toggling the option to limit item types to "Book" and repeating the search with the same search words.

Due to the format for the year being noted as being "yyyy-yyyy" (**Figure 5.2**), the librarians did not attempt to use the same adjustment for publication date as the patron personas did (changing the range from "1600-1799" to "-1799"). Instead, the publication date range was removed from the search, but this also produced

no results, as did using the keyword “illustrated” with the “Book” option toggled on.

Again citing their familiarity with the vocabulary used by the cataloguers, the librarian personas changed the keyword used from “illustrated” to “ill” and combined this new keyword with the specified date range. This combination produced no results, as did this combination and the toggled “Book” option.

Unlike the previous searches, using the keyword “ill” with no publication date range specified returned 28 results. Due to the broadness of the search, the need to examine each of these records for their relevance to the desired information was noted by both personas. Upon examination, these 28 records were then narrowed down to 24 results identified as true positives. The four results that were removed were identified as false positives, as the matches made by the search engine were simply in the inclusion of the letters “ill” in either the titles or the name of the author. Once again, the Koha search engine appears to have a quirk that prevents it from returning results that fall within a specified date range when using the “Year” field in the format it instructs the user to follow, which was noted by both personas and remembered moving forward.

This final search was deemed as so broad that any further attempts—including toggling the “Book” option—would only result in either the same number of results

returned or an increased number of false positives, and both personas declared this the end of the search process for this scenario.

A table of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
"illustrated" AND "1600-1799"	24	0	0	0%	0%
"illustrated" AND "1600-1799" AND "Book"	24	0	0	0%	0%
"illustrated"	24	0	0	0%	0%
"illustrated" AND "Book"	24	0	0	0%	0%
"ill" AND "1600-1799"	24	0	0	0%	0%
"ill" AND "1600-1799" AND "Book"	24	0	0	0%	0%
"ill"	24	28	24	100.0%	85.71%
Total	24	28	24	100.0%	85.71%

TABLE 5.16: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO ONE

Scenario Two

The prompt for Scenario Two was to assist a patron with finding items held in the sample collection that had ownership or ex libris inscriptions on them as opposed to more general handwritten annotations such as margin notes.

The first step the senior librarian persona undertook was to search the catalogue using the exact keyword phrase "ex libris", as this was the most specific phrase within the prompt that could potentially uncover results. The junior librarian also

started with this phrase, after a consultation with other staff. Given that there was no publication date range specified, in comparison to Scenario One, the two personas did not see any reason to specify any years in the “Year” field of the advanced search. This first search returned four results, as seen in **Figure 5.24** below:

Results	
<input type="checkbox"/>	<p>1. È PALAIA DIATHÈChÈ KATA TOYS EBDOMÈKONTA Vetus Testamentum Graecum ex versione septuaginta interpretum. Juxta exemplar Vaticanvm Romae editum.</p> <p>by Field, John academic printer [printer].</p> <p>Material type: Text; Format: print</p> <p>Publication details: Cambridge Field, John 1665</p> <p>Description: 19p., [1], 755p. index, handwritten annotations bound with "The Psalms of David in metre".</p> <p>Other title: È PALAIA DIATHÈChÈ KATA TOYS EBDOMÈKONTA Vetus Testamentum Graecum ex versione septuaginta interpretum. Juxta exemplar Vaticanum Romae editum..</p> <p style="text-align: right;">Holds (0) Add to cart Edit record Edit items</p>
<input type="checkbox"/>	<p>2. Enchiridium; seu institutiones imperiales, insertis latioribus materiis, theoricè & practicè digestae, & explicatae per erotemata. Johannis Arn. Corvini JC.</p> <p>by Corvinus, Joannes A Elzevir [publisher].</p> <p>Material type: Text; Format: print</p> <p>Publication details: Amsterdam Elzevir 1644</p> <p>Description: [20], 714p. index, handwritten annotations.</p> <p>Edition: Editio altera; emendator, & uberioribus additionibus auctior.</p> <p>Other title: Enchiridium; seu institutiones imperiales, insertis latioribus materiis, theoricè & practicè digestae, & explicatae per erotemata..</p> <p style="text-align: right;">Holds (0) Add to cart Edit record Edit items</p>
<input type="checkbox"/>	<p>3. Justini historiarum...libri XLIV. ex Trogo Pompeio</p> <p>by Trogus Pompeius Tonson, Jacob [printer] Watts, John [printer] Maittaire, Michael [editor].</p> <p>Material type: Text; Format: print</p> <p>Publication details: London Tonson, Jacob Watts, John 1713</p> <p>Description: [10], 281p., [50] col. ill. title page, indices, table of contents, handwritten annotations.</p> <p style="text-align: right;">Holds (0) Add to cart Edit record Edit items</p>
<input type="checkbox"/>	<p>4. Dictionnaire françois et italien: recueilli premierement par I. Antoine Phenice: & nouvellement reueu & augmenté d'une infinité de mots & manieres de parler, tirees de Bocace, Petrarque, Dante, Ariosto, & autres fameux auteurs italiens...Cest œuvre est comme nouveau, tres-utile & necessaire à tous ceux qui sont studieux de ces deux langues Par Pierre Canal</p> <p>by Canal, Pierre Chouet, Jacques [publisher].</p> <p>Material type: Text; Format: print</p> <p>Publication details: [s.l.] Chouet, Jacques 1598</p> <p>Description: 1 v. (unpaged) handwritten annotations.</p> <p>Other title: Dictionnaire françois et italien: recueilli premierement par I. Antoine Phenice: & nouvellement revue & augmenté d'une infinité de mots & manieres de parler, tirees de Bocace, Petrarque, Dante, Ariosto, & autres fameux auteurs italiens...Cest œuvre est comme nouveau, tres-utile & necessaire à tous ceux qui sont studieux de ces deux langues.</p> <p style="text-align: right;">Holds (0) Add to cart Edit record Edit items</p>

FIGURE 5.24: SEARCH RESULTS FOR FIRST SEARCH IN SCENARIO TWO

The match that Koha’s search engine made with these four results is not immediately apparent, so both personas noted the need to enter into each record to both verify that the system had in fact made a match to the exact phrase as well

as ensuring that the “ex libris” phrase in each record referred to an actual ownership inscription, rather than a note made by some anonymous third party. Once inside the record, Koha no longer highlights the phrase(s) it has made a match to, so the entire record had to be examined for each result. Upon this examination, all four records were confirmed to have true ex libris inscriptions and were added to the confirmed list of results. This search was then repeated with the “Book” option toggled on, returning the same four results.

Here is where the two personas differed on their next step in response to this prompt.

The junior librarian decided to cast a wider net with her next search, this time using the search phrase “handwritten annotations” in an attempt to uncover any handwritten inscriptions that may indicate ownership that did not use the exact phrase “ex libris” in doing so. An example of an ex libris inscription in comparison to one without this phrase is seen in **Figures 5.25** and **5.26** below:

Ē PALAIA DIATHĒChĒ KATA TOYS EBDOMĒKONTA Vetus Testamentum Graecum ex versione septuaginta interpretum. Juxta exemplar Vaticanvm Romae editum.

Contributor(s): [Field, John academic printer \[printer\]](#) [Q](#)
Material type: [Text](#)
Publication details: [Cambridge Field, John 1665](#)
Description: 19p., [1], 755p. index, handwritten annotations bound with "The Psalms of David in metre"
Other title: Ē PALAIA DIATHĒChĒ KATA TOYS EBDOMĒKONTA Vetus Testamentum Graecum ex versione septuaginta interpretum. Juxta exemplar Vaticanvm Romae editum
Uniform titles: [Bible Greek](#)
Action note: [does not need conservation](#)
Lists that include this title: [Spreadsheet Records](#)
MARC preview: [Show](#)
MARC framework: [Default](#)

[Holdings \(1\)](#) [Descriptions \(7\)](#) [Acquisition details](#)

Publication imprint reads:
"M DC LXV Cantabrigiae, excusum per Joannem Field, typographum academicum."
Sir William Stirling Maxwell, 9th Baronet's son, Sir John Maxwell Stirling-Maxwell, 10th Baronet donation
Greek, Latin (preface only) Greek alphabet, Latin alphabet
Associated slip in slip index has added notes:
pencilled number upper left corner; front (pencilled) struck out "Old Testament", "1,2" beside part; back (pencilled) "p.p. 19, 755, 516, 273"; "Bible" - sharing a slip with the following entry
Associated location book entry has added notes:
"MAXWELL OF POLLOCK" in location column
(pencilled) handwritten "834484" "Y123" "MAXWELL OF POLLOCK" "WING B2719"; [\(brown ink\) "Ex Libris Georgii Maxwell. 1745" on title page](#) has been rebound; "Vetus Testamentum Graecum" and "1665" embossed on spine; (pencilled) handwritten "834484" "Y123" "MAXWELL OF POLLOCK" "WING B2719"
has been rebound
does not need conservation

FIGURE 5.25: EXAMPLE OF A RECORD WITH AN "EX LIBRIS" INSCRIPTION NOTED

Sectionum conicarum libri V. Auctore Roberto Simson, in Academia Glasguensi Matheseos Professore.

By: [Simson, Robert Simson, Robert, M.D., Prof. Mathematics, Glas. Univ](#) [Q](#)
Contributor(s): [T. & W. Ruddimann \[publisher\]](#) [Q](#)
Material type: [Text](#)
Publication details: [Edinburgh T. & W. Ruddimann 1735](#)
Description: [1], iv-viii, 204p., 36 leaves of plates: ill. errata on p.viii, ill., ill. figures, printed annotations, handwritten annotations
Action note: [does not need conservation](#)
Lists that include this title: [Spreadsheet Records](#)
MARC preview: [Show](#)
MARC framework: [Default](#)

[Holdings \(1\)](#) [Descriptions \(7\)](#) [Acquisition details](#)

Publication imprint reads:
"MDCCLXXXV. Edinbvrgi, apud T. & W. Ruddimannos."
Sir William Stirling Maxwell, 9th Baronet's son, Sir John Maxwell Stirling-Maxwell, 10th Baronet donation
Latin Roman alphabet
Associated slip in slip index has added notes:
pencilled number upper left corner; front (pen) "Prof. Mathematics, Glas. Univ."
bookplate inner cover with Glasgow coat of arms and printed "Glasgow Corporation Public Libraries The Mitchell Library" and (pencilled) handwritten "Y148"; preserved bookplate with crest and printed "Maxwell of Polloc" and handwritten "E.2.2." and (pencilled) handwritten "83[circled] 4814"; [\(brown ink\) "J. Maxwell" on title page](#); has been rebound; "Sectionum Conicarum" "Simson" and "1735" embossed on spine; preserved bookplate with crest and printed "Maxwell of Polloc" and handwritten "E.2.2." and (pencilled) handwritten "83[circled] 4814"
has been rebound
does not need conservation

FIGURE 5.26: EXAMPLE OF A RECORD WITH AN OWNERSHIP-INDICATING INSCRIPTION NOTED

This search on its own returned 31 results, with the Koha search engine matching to either the phrase in its entirety or only the word "annotations". This persona then needed to go through each of these results to rule out any false positives where either there were no handwritten annotations noted (noted primarily in results with the match made only to the word "annotations" rather than the full phrase) or where there were no inscriptions that indicated any form of

ownership. After following this process, the 31 results were narrowed down to ten results, which included the same four results as the initial searches. This search was repeated with the “Book” option toggled on, with the same number of initial results and the same number of final results returned. Following this step, the junior librarian deemed the searching process finished, as she believed she had found all items in the sample collection that fit the requirements of the prompt.

In contrast, the senior librarian’s next step after the initial two searches—while also a broader search than the first two—was still more specific than the junior librarian persona’s third attempt. The senior librarian used the same search phrase, but searched for the exact phrase instead of any combination of the two words by placing the phrase in quotation marks like she had the “ex libris” phrase. The purpose of this search was the same as the junior librarian persona’s purpose, but the senior librarian wanted to eliminate any results without the exact phrase “handwritten annotations,” as she knew that any sort of inscription like those being requested would be recorded in an item with that descriptor and not the descriptor “printed annotations,” which would trigger the search engine to match a result if she did not use quotation marks. This search returned the same 31 results as the junior librarian’s search, including some items that did not have the exact phrase in the record. It is unclear why these results were included in the search results for this attempt given the bid to eliminate records without the exact phrase, but rather than immediately continuing with the same process as the

junior librarian persona had with her third and fourth searches, the senior librarian discarded this search, returned to the search page, and separated the two words of the phrase into two separate keywords (“handwritten” and “annotations”) and linked them with the Boolean operator “and” in another effort to try and find records with that exact phrasing. This attempt also returned the same 31 results with the same issue with the search engine not excluding any items without the exact phrasing of “handwritten annotations”, so the senior librarian was essentially forced to proceed with the same winnowing process. This list of 31 results was narrowed down to the same ten results as the junior librarian persona. The senior librarian did not repeat this final search with the “Book” option toggled on, as she believed it would only possibly eliminate any items that may fit the requirements as opposed to increasing the number of positive results returned. The quirk with Koha’s search engine and its perceived inability to search for exact phrasing matches was noted by the senior librarian persona and remembered while proceeding with the remaining scenarios.

A table of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“ex libris”	10	4	4	40.0%	100.0%
“ex libris” AND “Book”	10	4	4	40.0%	100.0%

“handwritten annotations” ²⁸	10	31	10	100.0%	32.26%
“handwritten annotations” AND “Book” ²⁹	10	31	10	100.0%	32.26%
“handwritten annotations” ³⁰	10	31	N/A	N/A	N/A
“handwritten” AND “annotations” ³¹	10	31	10	100.0%	32.26%
Total	10	31	10	100.0%	32.26%

TABLE 5.17: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO TWO

Scenario Three

For this scenario, the two librarian personas were asked to assist a patron looking for books about religious history and analysis that were published in the 17th through 18th centuries.

Both personas approached this search process in the same manner. They both changed the first two “Keyword” fields to “Subject” fields and entered “religious history” in the first field and “analysis” in the second field. Given the specified time period for publication, they entered “1600-1799” into the “Year” field. This search returned no results, so the search was repeated without the specified date range, as both personas returned to the note made in Scenario One about the first of Koha’s search engine quirks. This search returned no results either.

²⁸ This step was only performed by the junior librarian.

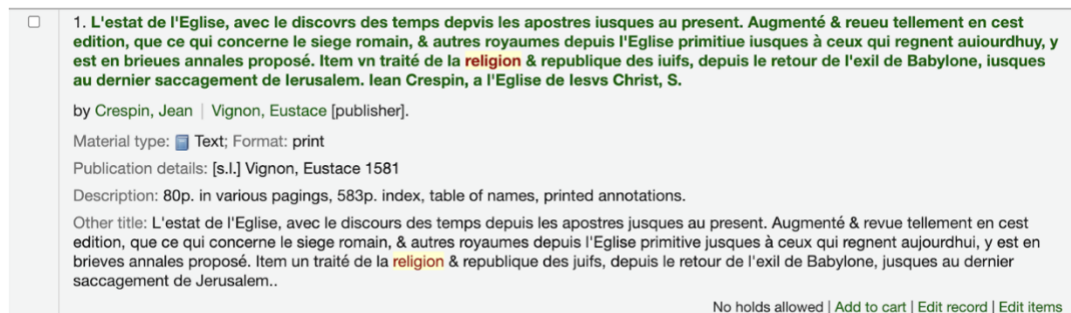
²⁹ This step was only performed by the junior librarian.

³⁰ This search was only performed by the senior librarian and was immediately deemed an unhelpful set of results, so no attempt at narrowing down the search results occurred here. As this number includes results that do not fit with the prompt requirements, this number is *not* considered when entering the total number of unique records returned for all searches performed by both librarian personas.

³¹ This step was only performed by the senior librarian.


The next search attempt was done with the subject information changing from “religious history” and “analysis” to the simpler term of “religion” and repeated both with and without the publication date range. Both attempts returned no results.

The “Subject” field was returned to the “Keyword” field, and the keyword “religion” was searched for with the date range “1600-1799”, which again returned no results. When the keyword “religion” was used with no specified date range however, two results were returned. The text within the record that Koha matched the query to was not noted as a subject for either record but was present in the title of both records, as one example shows in **Figure 5.27** below:



1. L'estat de l'Eglise, avec le discovrs des temps depuis les apostres iusques au present. Augmenté & reueu tellement en cest edition, que ce qui concerne le siege romain, & autres royaumes depuis l'Eglise primitiue iusques à ceux qui regnent auiourdhy, y est en brieues annales proposé. Item vn traité de la religion & republique des iuifs, depuis le retour de l'exil de Babylone, iusques au dernier saccagement de Ierusalem. Iean Crespin, a l'Eglise de Iesvs Christ, S.

by Crespin, Jean | Vignon, Eustace [publisher].

Material type:  Text; Format: print

Publication details: [s.l.] Vignon, Eustace 1581

Description: 80p. in various pagings, 583p. index, table of names, printed annotations.

Other title: L'estat de l'Eglise, avec le discours des temps depuis les apostres jusques au present. Augmenté & revue tellement en cest edition, que ce qui concerne le siege romain, & autres royaumes depuis l'Eglise primitive jusques à ceux qui regnent aujourd'hui, y est en brieves annales proposé. Item un traité de la religion & republique des juifs, depuis le retour de l'exil de Babylone, jusques au dernier saccagement de Jerusalem..

No holds allowed | [Add to cart](#) | [Edit record](#) | [Edit items](#)

FIGURE 5.27: AN EXAMPLE OF ONE RECORD RETURNED FOR FIFTH SEARCH ATTEMPT IN SCENARIO THREE

Before proceeding with further examination of this record, both personas noted that the publication date of 1581 (**Figure 5.27**) means that this item—regardless of subject *or* content—does not suit the requested date range, so it was thus eliminated as a false positive.

A detailed examination of the only other record, shown in **Figure 5.28** below, shows that there are no LCSH applied to the item denoting the content of the item. This lack of LCSH means that, while this item ostensibly seems to be about religion, there is no way to confirm or deny if the subject matter contained within matches the request from the patron without examining the item itself. Therefore, further examination of the physical item is needed prior to returning to the patron in order to either deem this result a true or false positive.

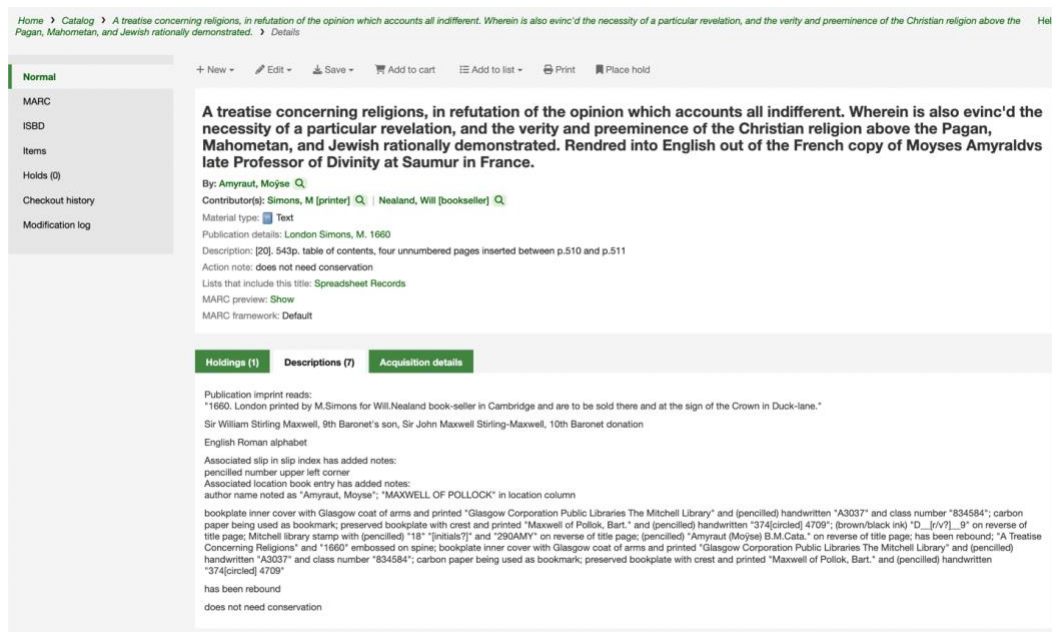


FIGURE 5.28: NORMAL VIEW OF LAST RESULT FOR FIFTH SEARCH IN SCENARIO THREE

Following this last search, both personas deemed any further searching unnecessary, as this search was already as broad as both personas could imagine a search being for this request.

A table of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“religious history” (subject) AND “analysis” (subject) AND “1600-1799”	10	0	0	0%	0%
“religious history” (subject) AND “analysis” (subject)	10	0	0	0%	0%
“religion” (subject) AND “1600-1799”	10	0	0	0%	0%
“religion” (subject)	10	0	0	0%	0%
“religion” (keyword) AND “1600-1799”	10	0	0	0%	0%
“religion” (keyword)	10	2	1	10.0%	50.0%
Total	10	2	1	10.0%	50.0%

TABLE 5.18: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO THREE

Scenario Four

The two librarian personas were tasked with assisting a patron in finding books published by Elzevir in the 17th through 18th centuries for this scenario.

Given the name of the publisher being sought and being familiar with the cataloguing format, both librarians approached this search in the same way. The “Keyword” field was changed to the “Publisher” field and filled out with “Elzevir”, while the “Year” field was filled with the date range “1600-1799”. This search returned no results, so it was repeated with the date range information removed. Without the date range, the search returned three results. All of these records included Elzevir listed specifically as the publisher, and all three items were published in the requested date range, thus confirming all three results as true positives for the patron. One final search was performed with the “Publisher”

field changed back to the “Keyword” field so the search engine would be looking for “Elzevir” as a keyword rather than specifically in the publisher role, but this search returned no additional results aside from the three returned previously.

Following this third search, the librarian personas deemed that they had found all items in the sample collection that fit the requirements of the prompt, and concluded their searches there.

A table of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“Elzevir” (publisher) AND “1600-1799”	3	0	0	0%	0%
“Elzevir” (publisher)	3	3	3	100.0%	100.0%
“Elzevir” (keyword)	3	3	3	100.0%	100.0%
Total	3	3	3	100.0%	100.0%

TABLE 5.19: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO FOUR

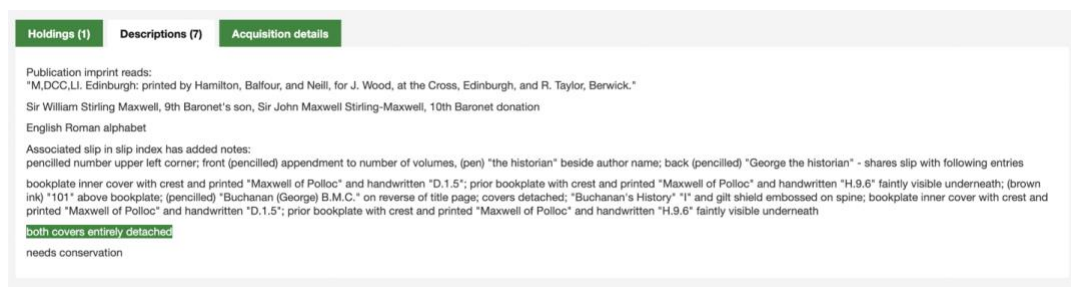
Scenario Five

The fifth scenario the librarian personas faced was a request from a member of the in-house conservation team looking for any items needing urgent conservation, as well as a description of what exactly the items are documented as needing.

While the junior librarian persona documented her first step in the searching process was to consult a senior member of staff for the exact verbiage used by the

cataloguers, once this was done both personas proceeded in the same manner. Both personas were now aware of the fact that the cataloguers for the sample collection use one of three phrases to describe the physical status of an item in regard to its conservation needs and the urgency required: “does not need conservation”, “may/might need conservation”, and “needs conservation”. Being that the request for this scenario specifies that the conservation team is looking for items needing *urgent* conservation, both personas decided they were looking for items labelled as “needs conservation”.

The first search performed by both personas was a simple keyword search for the phrase “needs conservation”. This search returned 12 results, but each result had to be examined in detail to ensure that the search engine had not made an erroneous match of some sort. This examination confirmed that these results were all in fact items deemed in need of urgent conservation, and each one had a note regarding the reason it was sorted into this category by the cataloguer. An example of such a note can be seen highlighted in the following figure:



The image shows a screenshot of a library catalog record. At the top, there are three tabs: "Holdings (1)", "Descriptions (7)", and "Acquisition details". The "Acquisition details" tab is selected. Below the tabs, the text reads: "Publication imprint reads: 'M,DCC,LI. Edinburgh: printed by Hamilton, Balfour, and Neill, for J. Wood, at the Cross, Edinburgh, and R. Taylor, Berwick.'" followed by "Sir William Stirling Maxwell, 9th Baronet's son, Sir John Maxwell Stirling-Maxwell, 10th Baronet donation" and "English Roman alphabet". A section titled "Associated slip in slip index has added notes:" contains a long, detailed note about bookplate inner covers, crests, and handwritten numbers. The phrase "both covers entirely detached:" is highlighted in green, and the phrase "needs conservation" is also highlighted in green below it.

FIGURE 5.29: ITEM DESCRIPTION CONTAINING INFORMATION ON CONSERVATION NEEDS

Following this first search, both librarian personas decided to perform one more search to ensure all items held in the sample collection meeting the needs of the request would be included. For this second search, the first keyword was winnowed down from “needs conservation” to simply the word “conservation”, which was then joined by the three key phrases “does not need”, “might need”, and “may need”, all of which were linked to the first keyword with the Boolean operator “not” to eliminate items falling into the two categories the personas were asked to ignore. This search returned the same 12 results as the initial search, thus satisfying both personas’ beliefs that their initial search had in fact returned all items in the sample collection that fit the request of the conservation team member.

A table of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“needs conservation”	12	12	12	100.0%	100.0%
“conservation” NOT “does not need” NOT “may need” NOT “might need”	12	12	12	100.0%	100.0%
Total	12	12	12	100.0%	100.0%

TABLE 5.20: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO FIVE

Scenario Six

Scenario Six asked the librarian personas to find items in the sample collection that were both in their original binding and contained handwritten annotations. For

the purposes of this scenario, “original binding” is defined as the binding material used to bind the item prior to its arrival at the Mitchell Library.

Both personas quickly identified the need to use a multipart search phrase for this scenario, as there are two different components that the item(s) needed to have to be a true positive result: original binding and handwritten annotations. This differed from the items requested in the previous scenario, where the request did technically ask for two components—needing urgent conservation and a description of the exact conservation needs—as the second of those two components was deemed as so unique for each item as to be unusable as a search term. Here, the first search term has already been identified in Scenario Two as “handwritten annotations”.

For the second portion of the search phrase, both personas turned to their prior knowledge of the cataloguing verbiage used. When a cataloguer determines that an item has been rebound—most often by either an in-house member of staff or an outside bindery requisitioned by the library—after its donation, this is noted with the exact phrase “has been rebound”. Where either the cataloguer cannot make this determination or the item is definitively in the binding with which it arrived at the library, the descriptor used is instead the exact phrase “bound in ____”, where the blank is filled in with a brief description of the material used.

Therefore, the search phrase used by both personas was made of two key phrases linked by the Boolean operator “not”: “handwritten annotations” as the first, and “has been rebound” as the second. This first search returned eight results, all of which required more detailed record examination to ensure their fitness for this prompt. This examination narrowed down the list of results to only three items that contained handwritten annotations and explicitly did not use the phrase “has been rebound”. The other five results were included because of the same simple matching issue noted in Scenario Two with the phrase “handwritten annotations”, meaning that these items were matched purely on the presence of *only* the word “annotations” instead of the full phrase.

Here, the personas differed from each other on their next step. The senior librarian believed this was the end of the search and that any other search would provide more work in terms of combing through more item records for verification only to end up with the same number of true positives, and the junior librarian wanted to try one more attempt to find any other items.

This second search once again used the key phrase “handwritten annotations”, but the Boolean operator was changed from “not” to “and” so the second key phrase could be changed to the confirmatory phrase “bound in”. This search returned five results, which narrowed down to only two of the same results as the first search. The junior librarian then re-performed the first search to compare the results to see what she had missed, and found that the third of the results found in

the initial set in fact used the phrase “original binding” instead of the accepted “bound in ____”. Armed with this knowledge, the junior librarian persona then performed a third search with the second phrase changed from “bound in” to “original binding” to see if any other items in the collection had been mistakenly catalogued as such. This search only returned the same result as the third one returned for the first search, and the junior librarian persona then decided that there were no other items held in the sample collection that fit the requirements of the prompt.

A table of the search results and their resulting recall and precision percentages is below:

Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
“handwritten annotations” NOT “has been rebound” ³²	3	8	3	100.0%	37.5%
“handwritten annotations” AND “bound in”	3	5	2	66.67%	40.0%
“handwritten annotations” AND “original binding”	3	1	1	33.33%	100.0%
Total	3	8	3	100.0%	37.5%

TABLE 5.21: RECALL AND PRECISION OF LIBRARIAN SEARCH RESULTS FOR SCENARIO SIX

³² This is the sole search for this scenario performed by both personas.

5.4 Summary

A comprehensive table containing the recall and precision values for all of the patron persona searches in all scenarios across both case studies is shown below, barring any footnotes discussed earlier in the chapter:

Case Study and Scenario	Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
CS1 Scen. 1	“illustrated” AND “1600-1799”	16	0	0	0%	0%
	“illustrated” OR “ill” AND “1600-1799”	16	0	0	0%	0%
	“illustrated” OR “ill”	16	0	0	0%	0%
	“illustrated” OR “ill” AND “-1799”	16	0	0	0%	0%
	Total	16	0	0	0%	0%
CS2 Scen. 1	“illustrated” AND “1600-1799”	24	0	0	0%	0%
	“illustrated” AND “-1799”	24	0	0	0%	0%
	“illustrated”	24	0	0	0%	0%
	“illustrated” OR “ill” AND “1600-1799”	24	0	0	0%	0%
	“illustrated” OR “ill” AND “-1799”	24	0	0	0%	0%
	“illustrated” OR “ill”	24	28	24	100.0%	85.71%
	Total	24	28	24	100.0%	85.71%
CS1 Scen. 2	“ownership inscriptions”	9	0	0	0%	0%
	“handwritten inscriptions”	9	0	0	0%	0%
	“handwritten”	9	0	0	0%	0%

	Total	9	0	0	0%	0%
CS2 Scen. 2	“ownership inscriptions”	10	0	0	0%	0%
	“ownership”	10	1	1	10.0%	100.0%
	“ex libris”	10	4	4	40.0%	100.0%
	“handwritten annotations”	10	31	10	100.0%	32.26%
	Total	10	31	10	100.0%	32.26%
CS1 Scen. 3	“religious history” (keyword) AND “1600-1799”	8	0	0	0%	0%
	“religious history” (keyword)	8	0	0	0%	0%
	“religious history” (subject)	8	0	0	0%	0%
	“religious history” (subject) AND “1600-1799”	8	0	0	0%	0%
	“religion” (subject) AND “1600-1799”	8	0	0	0%	0%
	“religion” (keyword) AND “1600-1799”	8	0	0	0%	0%
	“religion” (keyword)	8	1	1	12.5%	100.0%
	Total	8	1	1	12.5%	100.0%
CS2 Scen. 3	“religious history” (keyword) AND “1600-1799”	10	0	0	0%	0%
	“religious history” (keyword)	10	0	0	0%	0%
	“religious history” (keyword) AND “-1799”	10	0	0	0%	0%
	“religious history” (subject) AND “1600-1799”	10	0	0	0%	0%
	“religious history”	10	0	0	0%	0%

	(subject) AND “-1799”					
	“religious history” (subject)	10	0	0	0%	0%
	“religion” (subject) AND “1600-1799”	10	0	0	0%	0%
	“religion” (subject) AND “-1799”	10	0	0	0%	0%
	“religion” (subject)	10	0	0	0%	0%
	“religion” (keyword) AND “1600-1799”	10	0	0	0%	0%
	“religion” (keyword) AND “-1799”	10	0	0	0%	0%
	“religion” (keyword)	10	2	1	10.0%	50.0%
	Total	10	2	1	10.0%	50.0%
CS1 Scen. 4	“Elzevir” (keyword) AND “1600-1799”	2	0	0	0%	0%
	“Elzevir” (keyword)	2	1	1	50.0%	100.0%
	“Elzevir” (publisher)	2	0	0	0%	0%
	“Elzevir” (publisher) AND “1600-1799”	2	0	0	0%	0%
	Total	2	1	1	50.0%	100.0%
CS2 Scen. 4	“Elzevir” (keyword) AND “1600-1799”	3	0	0	0%	0%
	“Elzevir” (keyword) AND “-1799”	3	0	0	0%	0%
	“Elzevir” (keyword)	3	3	3	100.0%	100.0%
	“Elzevir” (publisher)	3	3	3	100.0%	100.0%
	Total	3	3	3	100.0%	100.0%

TABLE 5.22: RECALL AND PRECISION OF ALL PATRON SEARCH RESULTS FOR CS1 AND CS2

A comprehensive table containing the recall and precision values for all of the librarian persona searches in all scenarios across both case studies is shown here, barring any footnotes discussed earlier in the chapter:

Case Study and Scenario	Search	Number of Relevant Items in Sample Collection	Number of Results Returned	Relevant Results Returned	Recall (%)	Precision (%)
CS1 Scen. 1	"illustrated" AND "1600-1799"	16	0	0	0%	0%
	"illustrated" AND "1600-1799" AND "Book"	16	0	0	0%	0%
	"illustrated"	16	0	0	0%	0%
	"illustrated" AND "Book"	16	0	0	0%	0%
	"ill" AND "1600-1799"	16	0	0	0%	0%
	"ill" AND "1600-1799" AND "Book"	16	0	0	0%	0%
	"ill"	16	2	0	0%	0%
	"ill" AND "Book"	16	2	0	0%	0%
	Total	16	2	0	0%	0%
CS2 Scen. 1	"illustrated" AND "1600-1799"	24	0	0	0%	0%
	"illustrated" AND "1600-1799" AND "Book"	24	0	0	0%	0%
	"illustrated"	24	0	0	0%	0%
	"illustrated" AND "Book"	24	0	0	0%	0%
	"ill" AND "1600-1799"	24	0	0	0%	0%
	"ill" AND "1600-1799" AND "Book"	24	0	0	0%	0%
	"ill"	24	28	24	100.0%	85.71%
	Total	24	28	24	100.0%	85.71%
CS1 Scen. 2	"handwritten annotations"	9	0	0	0%	0%
	"handwritten annotations" AND "Book"	9	0	0	0%	0%
	Total	9	0	0	0%	0%

CS2 Scen. 2	“ex libris”	10	4	4	40.0%	100.0%
	“ex libris” AND “Book”	10	4	4	40.0%	100.0%
	“handwritten annotations”	10	31	10	100.0%	32.26%
	“handwritten annotations” AND “Book”	10	31	10	100.0%	32.26%
	“handwritten annotations”	10	31	N/A	N/A	N/A
	“handwritten” AND “annotations”	10	31	10	100.0%	32.26%
	Total	10	31	10	100.0%	32.26%
CS1 Scen. 3	“religious history” (keyword) AND “analysis” (keyword) AND “1600-1799”	8	0	0	0%	0%
	“religious history” (keyword) AND “analysis” (keyword) AND “1600-1799” AND “Book”	8	0	0	0%	0%
	“religious history” (keyword) AND “analysis” (keyword)	8	0	0	0%	0%
	“religious history” (keyword) AND “analysis” (keyword) AND “Book”	8	0	0	0%	0%
	“religion” (keyword) AND “1600-1799”	8	0	0	0%	0%
	“religion” (keyword) AND “1600-1799” AND “Book”	8	0	0	0%	0%
	“religion” (keyword)	8	1	1	12.5%	100.0%
	“religion” (keyword) AND “Book”	8	1	1	12.5%	100.0%
	“religion” (subject)	8	0	0	0%	0%
	Total	8	1	1	12.5%	100.0%

CS2 Scen. 3	“religious history” (subject) AND “analysis” (subject) AND “1600-1799”	10	0	0	0%	0%
	“religious history” (subject) AND “analysis” (subject)	10	0	0	0%	0%
	“religion” (subject) AND “1600-1799”	10	0	0	0%	0%
	“religion” (subject)	10	0	0	0%	0%
	“religion” (keyword) AND “1600-1799”	10	0	0	0%	0%
	“religion” (keyword)	10	2	1	10.0%	50.0%
	Total	10	2	1	10.0%	50.0%
CS1 Scen. 4	“Elzevir” (publisher) AND “1600-1799” (publication date)	2	0	0	0%	0%
	“Elzevir” (publisher)	2	0	0	0%	0%
	“Elzevir” (keyword)	2	1	1	50.0%	100.0%
	Total	2	1	1	50.0%	100.0%
CS2 Scen. 4	“Elzevir” (publisher) AND “1600-1799”	3	0	0	0%	0%
	“Elzevir” (publisher)	3	3	3	100.0%	100.0%
	“Elzevir” (keyword)	3	3	3	100.0%	100.0%
	Total	3	3	3	100.0%	100.0%
CS1 Scen. 5	“conservation” NOT “does not need” NOT “may need” NOT “might need”	7	0	0	0%	0%
	“conservation” NOT “does not need” NOT “may need”	7	0	0	0%	0%
	“conservation” NOT “does not need”	7	0	0	0%	0%

	“conservation”	7	0	0	0%	0%
	Total	7	0	0	0%	0%
CS2 Scen. 5	“needs conservation”	12	12	12	100.0%	100.0%
	“conservation” NOT “does not need” NOT “may need” NOT “might need”	12	12	12	100.0%	100.0%
	Total	12	12	12	100.0%	100.0%
CS1 Scen. 6	“handwritten annotations” NOT “has been rebound”	3	0	0	0%	0%
	“handwritten annotations” AND “bound in”	3	0	0	0%	0%
	Total	3	0	0	0%	0%
CS2 Scen. 6	“handwritten annotations” NOT “has been rebound”	3	8	3	100.0%	37.5%
	“handwritten annotations” AND “bound in”	3	5	2	66.67%	40.0%
	“handwritten annotations” AND “original binding”	3	1	1	33.33%	100.0%
	Total	3	8	3	100.0%	37.5%

TABLE 5.23: RECALL AND PRECISION OF ALL LIBRARIAN SEARCH RESULTS FOR CS1 AND CS2

This chapter has thus laid out the detailed results of both CS1 and CS2 in preparation for the analysis discussed in the following chapter. The results were first organised by case study, then broken down into the results for each persona type when working through each of the scenarios laid out at the beginning of this chapter. Following this in the next chapter is a thorough discussion of the meaning of these results.

Chapter 6: Discussions

This chapter lays out what exactly the results shown in the previous two chapters mean in terms of not only this thesis, but also the field of special collection librarianship as a whole. In order to thoroughly discuss these points, this chapter will lay out an analysis of the data contained in **Chapter 4** regarding the investigations into answering RQ1 and RQ2, before moving to an analysis of the data contained in **Chapter 5** gleaned in the process of answering RQ3. The discussion around RQ3 will first involve a comparison between the two sets of results gleaned from CS1 and CS2 (shown in detail in **Chapter 5**), as well as their meaning. Following these analyses, the chapter will then proceed to a discussion of the reflections and limitations encountered throughout the entire project, before concluding with a brief possible generalisation of this research's results.

6.1 Remembering the overarching research goals

It is important at this point to reiterate the research goals and aims of this project prior to further discussion and analysis. As a reminder, the ultimate goal of *any* librarian is to provide access to the items under their supervision to both external and internal users. In the current digital age, this facilitation can be a difficult process for rare books and special collections librarians in particular, due to the

unique nature of the items in the field. Where mainstream items can come with metadata and bibliographic records ready-made as part of the Cataloguing in Publication (CIP) process, the same cannot be said for items like those in the Stirling Maxwell collection studied herein. Berger (2014) and Cullingford (2022) make this point very clear: the users who access rare books and special collections use the items held within not only as carriers of information but also as informational objects in themselves. As such, librarians placed in charge of the handling and management of these collections must therefore have a higher level of specialist knowledge in order to catalogue these items correctly and thoroughly, which then facilitates better access to patrons and links items both within a collection and to other collections held in the library's inventory, as discussed in the literature review contained in **Chapter 2**. The importance of this specialist knowledge has been demonstrated clearly in **Chapters 4** and **5**, the meaning of which will be discussed thoroughly in this chapter.

6.2 Question One discussion

This section discusses the results of the research conducted in answering RQ1, defined in **Chapter 1**. This analysis is based on the literature review detailed in **Chapter 2, Section 2.1** as well as the hands-on work completed with the Stirling Maxwell donation in **Chapter 4, Section 4.1**.

6.2.1 Defining rare books and special collections

Chapter 2, Section 2.1 discusses the nebulous definitions of the terms “rare books”, “special collections”, and “archives” as departments in relation to the work performed in this project. The consensus in the cited literature is that the definition of rare books and special collections—while indistinct by necessity—lies primarily in the value, vulnerability, and cohesiveness of a given collection, as well as in the cultural heritage information contained within (Berger, 2014; Cullingford, 2022). These departments do not only hold the centuries-old leatherbound, crumbling books one usually thinks of when thinking of a certain department, as noted by Berger: “It is clear that if an institution has something it does not know what to do with and, for any reason, cannot dispose of, special collections is a good dumping ground” (2014: p.3). Thus, there is the need to use a slightly different definition of the descriptor “rare” when speaking of these items—their “rarity” lies not in the scarcity of an item, but in the *uniqueness* of the item instead.

Take for example, Anne Boleyn’s *Book of Hours*, published in Bruges sometime around 1410-1450—*Hours* were “personal prayer books which were popular in England from the 13th Century until the [R]eformation. Short services to the Virgin Mary were read at eight fixed ‘Hours’ during the day. ‘Hours’ also contained a calendar of church festivals, psalms, prayers to favourite saints and services for the dead” (‘Castle Objects of the Month: Anne Boleyn’s Books of Hours’,

2021). Given their popularity—specifically in England—at the time, scarcity is not the most useful metric by which to measure this particular copy’s rarity. Instead, the more nuanced definition of “rarity” that relies on an item’s uniqueness is the better of the two, as “This illuminated manuscript is the only book known where her personal annotation survives intact” (‘Castle Objects of the Month: Anne Boleyn’s Books of Hours’, 2021). Without this unique aspect, it is fair to say that the book would be viewed as less rare than it is with the annotation—still special, just less so. It is a similar case with some items in the Stirling Maxwell collection examined in **Chapter 4, Section 4.1**, where 42 items of the 509 items in Dataset One contained handwritten annotations. Similarly, the vast majority of the items in this dataset contained Maxwell bookplates of one sort or another, an aspect that once again makes these items unique compared to duplicate copies owned by some other family or library.

One example of an item that could be compared to Anne Boleyn’s Book of Hours that is in the Stirling Maxwell collection is a copy of *Les aventures de Telemaque, fils d’Ulysse* that contains the ex libris inscription shown in the picture below:

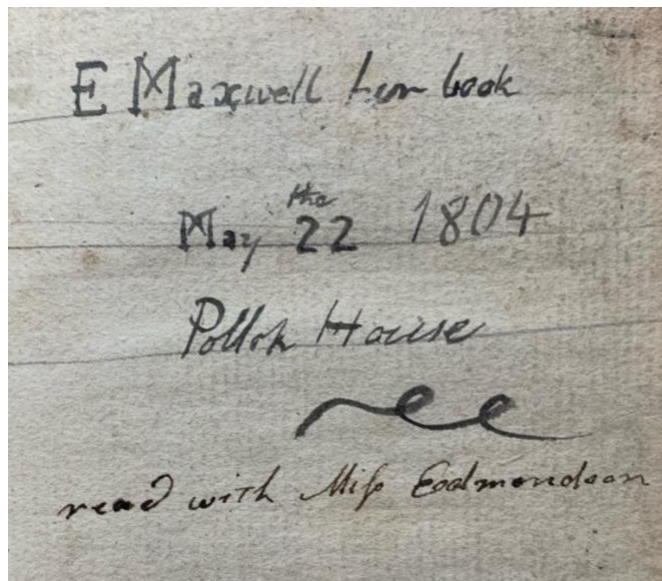


FIGURE 6.1: EX LIBRIS INSCRIPTION PRESENT IN THE COLLECTION

This inscription is not by Sir William himself, but by another Maxwell. When examining the truncated family tree below, there is an “E Maxwell” who would have been alive in 1804: his mother, Elizabeth Maxwell.

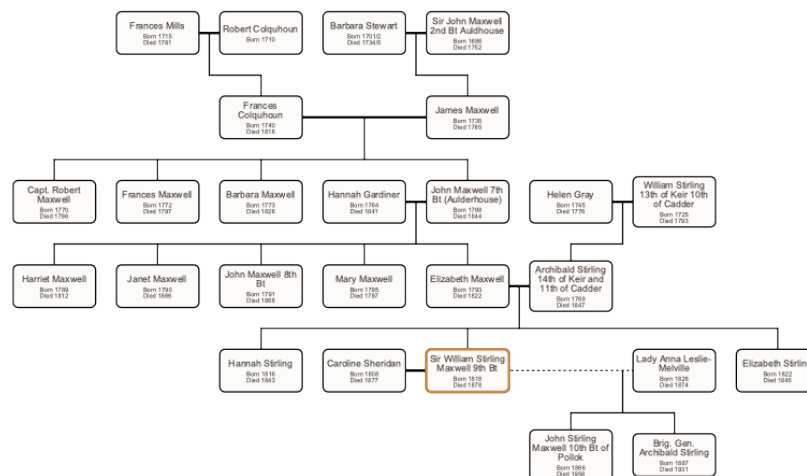


FIGURE 6.2: TRUNCATED STIRLING MAXWELL FAMILY TREE

As Pearson notes, “Handwritten transcriptions on title pages, flyleaves or elsewhere form the most frequently encountered evidence of provenance” (2019:

p.20), meaning that it can conclusively be said that this book was once owned by Sir William's mother due to the inscription, and that it was passed down to him at some point due to the fact that it ended up in his library. The note beneath the ex libris inscription in different handwriting is an annotation. Annotations like this one "provide direct evidence that books *were* used, that they did not sit on shelves unread" (Pearson, 2019: p.4). Ultimately, much can be learned about Sir William and his mother through this simple inscription in a single book in his collection:

The mere fact of ownership of a book, whether annotated or not, provides evidence of past interest. Knowing the contents of someone's library has an obvious relevance to the study of the life and thinking of that individual (Pearson, 2019: p.4).

It is factors such as this one that set rare books and special collections items apart from mainstream publishing, therefore requiring detailed documentation that is not only available to the library patrons for perusal, but also to the library staff and stakeholders. In the example of Anne Boleyn's 1410-1450 *Book of Hours*, the item is used as an attractor for visitors because of its rarity and uniqueness as well as its historical, cultural heritage value to the community ('Castle Objects of the Month: Anne Boleyn's Books of Hours', 2021). This then may lead to increased funding for the organization caring for the item, whether that is in the form of donations, revenue from on-site purchases by the public, or even options like grants or increased budget from the city council or parent organization. With

funding constraints being such a prevalent and pressing issue for libraries of any type (Lynch, Tomas and Hattenstone, 2024; Stenstrom and Haycock, 2015; Oder, 2009), as laid out in **Chapter 2, Section 2.4.3.2**, having items that can provide this type of attraction for visitors to the library can then be incredibly beneficial.

6.2.2 Defining the rarity of this collection

Following the identification of what is meant by the department name of rare books and special collections comes the justification for the Stirling Maxwell collection being handled by this type of department over any other at the Mitchell Library.

As stated above, there are 42 out of 509 items in Dataset One that have handwritten annotations of some sort. If the handwritten annotations in Anne Boleyn's 1410-1450 *Book of Hours* are to be treated as a significant marker of that book's uniqueness, rarity, and historical importance ('Castle Objects of the Month: Anne Boleyn's Books of Hours', 2021), then it stands to reason that they should be treated as such in the Stirling Maxwell collection. As for the items that do not have handwritten annotations, the vast majority do have Maxwell family bookplates, which could be compared to the "personal annotation" ('Castle Objects of the Month: Anne Boleyn's Books of Hours', 2021) contained in the *Book of Hours* as well—the bookplates prove that the items were in fact owned by the Stirling Maxwell family. These two factors then form the bulk of the argument for the collection's rarity and cohesiveness, two factors given as necessary for the definition of rare

books and special collections above. The third factor that then must be considered as one that solidifies the position of this collection as part of special collections is the cultural heritage aspect.

Sir William Stirling Maxwell's extensive involvement in Glaswegian politics and education during his lifetime has been detailed by Charney (2020), making him a highly influential individual of the time period, especially for the community and locale the Mitchell Library serves. This builds on the foundation laid by Berger (2014) and Cullingford (2022) for the definition of such items in particular, providing practical examples of the distinctions they discuss. It is then this combination of the definitive rarity and cohesiveness of the items in the collection, as well as their prior ownership by Stirling Maxwell, that unequivocally defines this collection as one of rare books and special collections.

6.3 Question Two discussion

This section discusses the results of the research conducted in answering RQ2, defined in **Chapter 1**. This analysis is based on the literature review detailed in **Chapter 2, Section 2.3** as well as the hands-on work with interviews, coding, and persona/scenario creation in **Chapter 4, Sections 4.2-4.3**.

6.3.1 Differences between literature and in vivo code prevalence

While **Chapter 4, Sections 4.2-4.3** discussed how the data obtained in the process of answering RQ2 was used to create the personas and scenarios used in the final phase of this project, with analysis as to the results' meaning in terms of the ongoing project alone, this section will instead detail the wider meaning of the results obtained during the interview and coding process.

Returning to **Table 4.15**, which breaks down the number of references per sample population for codes created both by in vivo coding and by literature-based codes, will provide the data needed to understand the following analysis.

The difference in number of references per population between the two code sources is small, but there is still a noticeable decrease in total number of references for both populations when moving from the in vivo codes to the literature-based codes. What is important to consider however, is that the literature-based codes can be essentially subsumed into the in vivo code category named "Process Steps", rather than each code being a direct comparison to another code in the opposite category. This fact thus renders the difference between code sources and their prevalence in the sample population transcripts unimportant.

6.3.2 Comments on the literature

The prevalence of the codes obtained from the work done by Rhee (2010; 2012) as an expansion on Meho and Tibbo (2003), which was in turn an expansion of Ellis (1993), reinforces the previous work in the historian ISB field with some suggestions.

The first suggestion is that the codes prevalent in the literature for the ISBs of historians *are* also able to be applied to further research regarding the ISBs of special collections librarians primarily handling historical materials such as the Stirling Maxwell collection. In the small sample population of Mitchell Special Collections librarians interviewed in this project, nine out of 13 behaviours from the literature regarding historians were noted as being present—it is a very real possibility that with a wider and larger sample of similar librarians, more of the 13 behaviours could be noted.

The next suggestion, following directly from the first, is that if applying these behaviours as codes to special collections librarians handling historical items, one of the behaviours should be split into smaller, less nebulous behaviours. As discussed in **Chapter 4, Section 4.2.3**, the merging of the three codes composing the “Process Steps” in vivo code group³³ for the librarian transcripts into the single code named “Networking” from the literature-based codes results in the loss of

³³ Each of these three codes are defined in **Chapter 4, Section 4.2.1**.

critical nuance of special collections librarians' ISBs. Without the nuance provided by the sub-codes "Inter-agency Interaction", "Inter-staff Interaction", and "Personal Interaction", uncovering a librarian's exact process for searching for information becomes harder for the analyst doing the coding. Thus, returning the nuance to this list will streamline the analysis process, even if it only does so by some small measure.

Finally, returning to the list of the 13 behaviours used for literature-based codes, a 14th behaviour should be added: breadcrumbing. While this behaviour may be similar at first glance to the *chaining* behaviour, there is one critical difference found when interviewing the historians sample population at the Mitchell Library. Ellis' definition of *chaining* as the act of "follow[ing] chains of citations or other forms of referential connection between material" (Ellis, 1993: p.482) is the closest of the three definitions provided by all the cited researchers to something akin to the process described by one patron interview as bread-crumbling, there is still enough of a difference to validate the addition of the new behaviour. *Chaining*, especially as defined by Rhee (2012) and Meho and Tibbo (2003), refers most clearly to citations of some sort, or footnotes. Breadcrumbing is a process that does not use academic citations or footnotes—instead, the "breadcrumbs" are more colloquial, or even visual, references that the historians follow through the metaphorical forest of information. Rather than a solid, clear link between citations like that provided through chaining, breadcrumbing is a more meandering

process. For example, a photographic historian, while looking at a photo tentatively dated somewhere in the late 1870s, may notice a particular wrought-iron railing style in the background that was not being produced until the late 1880s. This then can recontextualise the whole photo, but could also then lead to more similar information crumbs that provide better depth of knowledge to the historian looking at it. Thus, the suggestion is that chaining be defined as following solid, ready-made links like academic citations, while breadcrumbing be defined as following scattered colloquial or photographic “breadcrumbs” to more information.

Future research into breadcrumbing could delve into the level of intention compared to that seen in chaining behaviours through the use of qualitative methods such as further interviews with wider sample sizes. A larger sample size could effectively narrow the proposed definition of breadcrumbing as well as that of chaining.

6.3.3 Benefits of the process

The other outcome of this phase of the project lies in the development of a methodology by which other libraries with rare books and special collections departments can then evaluate the needs of their own users as they work to either create or update their own collection management policies. **Chapter 2, Section 2.4.2.1** lays out exactly why librarians *must* understand the needs of their users in

the process of CMP creation and re-evaluation—a CMP that does not accommodate for the collection user’s needs is one that will at best annoy the users, and at worst, drive users away from the collection in droves. Just as explained in **Section 6.2.1**, it is better for libraries to *attract* and *retain* visitors and users, especially when it comes to justifying higher funding needs to stakeholders.

This benefit does not apply solely to what Cullingford (2022) and Crockett (2016) consider external users—it *also* applies to the internal users: the librarians and other library staff as well. Just as a CMP that does not accommodate for the needs of the external users will undoubtedly result in frustration at best and outright avoidance at worst, these results will also be seen to some extent in the internal user group. Thus, it is necessary that these user needs and information-seeking behaviours are accounted for in addition to those of the external user group(s) in the CMP creation and re-evaluation process.

Library staff tasked with CMP creation for a special collections department like that at the Mitchell Library in Glasgow can therefore follow the methodology used herein to conduct their own user needs analysis. This procedure, drawing heavily from the literature in its development, is simple enough to be conducted by a single staff member. It removes some of the barriers noted by Clayton and Gorman (2006), Oder (2009), Stenstrom and Haycock (2015), Lynch, Tomas and Hattenstone (2024), and Hibner and Kelly (2013), amongst others discussed in **Chapter 2, Section 2.4.3**, which then increases the likelihood of the necessary

CMP process taking place in more libraries around the world. After all, CMPs aid librarians in performing the core task of their profession: organising knowledge.

6.4 Question Three discussion

This section discusses the results of the research conducted in answering RQ3, defined in **Chapter 1**. This analysis is based primarily on the raw data detailed in **Chapter 5**.

6.4.1 Comparison of case studies

To thoroughly analyse and compare the results of both case studies, the results of each scenario for each persona type—both for recall and for precision—in each case study must be displayed in the table below for ease of comparison:

Scenario Number	Persona Result Recall ³⁴ (%) and Precision ³⁵ (%)							
	Patron Personas Recall (CS1)	Patron Personas Precision (CS1)	Patron Personas Recall (CS2)	Patron Personas Precision (CS2)	Librarian Personas Recall (CS1)	Librarian Personas Precision (CS1)	Librarian Personas Recall (CS2)	Librarian Personas Precision (CS2)
One	0%	0%	100.0%	85.71%	0%	0%	100.0%	85.71%
Two	0%	0%	100.0%	32.26%	0%	0%	100.0%	32.26%
Three	12.5%	100.0%	10.0%	50.0%	12.5%	100.0%	10.0%	50.0%
Four	50.0%	100.0%	100.0%	100.0%	50.0%	100.0%	100.0%	100.0%
Five	N/A	N/A	N/A	N/A	0%	0%	100.0%	100.0%

³⁴ “Recall” here is again defined as the value given by the formula:

$$\text{Recall \%} = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Number of Relevant Documents Present in Collection}} \times 100$$

³⁵ “Precision” here is again defined as the value given by the formula:

$$\text{Precision \%} = \frac{\text{Number of Relevant Documents Retrieved}}{\text{Total Number of Documents Retrieved by a Search}} \times 100$$

Six	N/A	N/A	N/A	N/A	0%	0%	100.0%	37.5%
-----	-----	-----	-----	-----	----	----	--------	-------

TABLE 6.1: COMPARING RECALL (%) AND PRECISION (%) PER SCENARIO PER PERSONA TYPE FOR BOTH CS1 AND CS2

By and large, the recall of returned results for each of the six scenarios in CS1 is incredibly low, aside for Scenario Four. For both persona groups, the sole result returned for Scenarios Three and Four were deemed only “possible” positives, as in each case, the link between the note in the bibliographic record and the initial scenario request was unclear past being a simple match made by the search engine. In Scenario Three, this fogginess lay in linking the title of a book that contains the word “religion” to an initial request for books on religious history and analysis in the 17th through 18th centuries—as seen in **Figures 5.10** and **5.13-4**, the sparse information included in the record does not include any LCSHs or information about the subject matter of the book, meaning a judgement on relevance in the search result had to be made based solely on the simple match of the word “religion” in the title with the same word in the keyword search. For Scenario Four, the uncertainty is in the fact that the item description for the item having a handwritten note on the slip the record is based on, with no knowledge of when this note was made, who made the note, or if the note is referring to Elzevir as the publisher of the item on the slip, or any other possibility. As a result, while this level of metadata detail did mean there was a narrowing down of items contained in the sample collection that may match the request, it is clear that the more detailed metadata collected and analysed for CS2 is significantly more helpful

across both persona group types, as demonstrated in the increase in recall values in **Table 6.1**.

Precision, on the other hand, can be seen as a slightly less useful measurement of the effectiveness of an information retrieval system when viewed on its own than recall. After all, the ability to recall more information can be quickly and easily displayed as in the table above, or in a graph showing the visible increase in recall. In contrast, precision is a much more subtle measurement that can seem confusing to those not familiar with the practice of evaluating information retrieval methods—in **Table 6.1** alone, while two out of four scenarios faced by the patron personas and four out of six scenarios faced by the librarian personas show an increase in precision, one scenario shows no change whatsoever, while the other remaining scenario shows a significant decrease. Referring back to Buckland, as discussed in **Chapter 5**, precision is merely a measure of how *pure* the results retrieved by a given system are—in effect, out of all the results called up in response to a query, how many of those are truly relevant to the intent of the query (Buckland, 2017)? As discussed in **Chapters 3** and **5**, the ideal, perfect retrieval system would have 100% recall and 100% precision, but this is a rare result in real-world, practical applications. More often, “efforts to avoid nonrelevant items in order to achieve high precision tend also to increase the number of relevant items that are not retrieved (lower recall)” (Buckland, 2017: p.156), and vice versa; an

increase in recall is most often accompanied by decreases in precision as more results are retrieved.

This is not to say that precision is a useless evaluation method here. As a real-world application of an information retrieval system, the variation noted in the above table is an *expected* result. Some of the scenarios' results—for example, Scenarios One and Two—still show an increase in precision as well as recall, which demonstrates that the metadata framework being tested within is a more useful framework than the existing cataloguing framework (the slip index) in a more holistic, thorough way than the sole use of recall could. The precision results simply need to be taken more in an individual manner than an overall manner like the recall results of this research can.

Examining the data laid out in **Table 6.1** in a scenario-by-scenario manner is necessary for a true understanding of the exact magnitude of the difference between CS1 and CS2.

6.4.1.1 Scenario One

Scenario One shows a clear distinction between the two case studies, with both persona types jumping from a 0% recall in CS1 with no results returned out of the sample collection's 16 relevant items, to a 100.0% accuracy in CS2 with 24 results returned out of the sample collection's 24 relevant items. This result is particularly notable through the lens of the Photographic Historian Persona created in **Chapter**

4, as they represent the group of patrons that access the Mitchell Library's special collections for the purpose of researching photographs and illustrations. In CS1, even with the same items being catalogued in each sample collection, the lack of detailed metadata recorded in the catalogue would definitively lead that patron to believe that there were no items held in the sample collection that fit their need for illustrated items, which is demonstrably false—the same request of the sample collection used in CS2 proves that there *are* items in the collection that fit the patron's needs.

In terms of precision, there is also a significant increase between the two case studies upon application of the more detailed metadata framework, with CS1 having a precision value of 0% and CS2 having a value of 85.71% for both persona types. In CS1, there were no results retrieved *at all* for any query, so by virtue of there being no relevant results among the non-existent pool of overall results returned, there can only be a 0% precision or “completely imprecise” rating. On the other hand, with the increase in recall present in CS2 comes an increase in precision—because more items were returned per query in this case study than the previous one, it therefore mathematically creates an equation that provides a non-zero result. While the precision is still lower than the recall in the second case study (85.71% vs. 100.0%), remember that, by definition, a real-world information retrieval system is going to be less precise as it returns more results. The 85.71% value is still indicative of a highly precise system, or the significantly increased

ability of the Koha search engine to pick out results when given the more detailed metadata contained in the CS2 sample collection compared to the bare-bones metadata contained in the CS1 sample collection.

6.4.1.2 Scenario Two

Scenario Two is another example of the stark difference between the two case studies, with both persona types having a 0% recall in CS1, compared with a 100.0% recall in CS2. Once again, the lack of detailed metadata contained in the sample collection for CS1 would effectively lead both patrons and librarians to the erroneous belief that none of the 34 items fit the needs of the prompt (out of the nine items previously defined as relevant), when CS2 proves that 10 out of 10 relevant items fit the needs of the prompt and are held in the sample collection.

In terms of precision, this scenario performed similarly to the previous one, with an increase in precision accompanying the increase in recall. Again, the precision for both persona types in CS1 is at 0%—coming from the search results returning no results at all, let alone relevant results—whereas CS2 has a precision of 32.26%. The increase between the two case studies for this scenario is smaller than the increase presented for Scenario One but still demonstrates that the metadata framework in the second case study results in a more precise output for the Koha search engine than the initial framework used in CS1.

6.4.1.3 Scenario Three

Out of the six scenarios encountered by both persona types in both case studies, Scenario Three appears on the surface to have fared worse in CS2 than in CS1, with a lower recall in the former than in the latter. This reported decrease in recall is slightly misleading, as in actuality, there is no change in the raw number of relevant results returned between the two case studies—each case study was able to return only one relevant result from both sample collections. Instead, the apparent decrease in recall comes solely from the fact that CS2 has a larger number of relevant records contained in its sample collection than CS1, which results in a smaller result when calculating recall for the scenario. In addition, it is important to note that the sole result returned across both case studies is only a *potential* result rather than a *verified* result that fits the needs of the prompt—meaning that the available metadata for both case studies does not contain enough detail for any persona to judge the item’s relevance by the catalogue record alone. As noted in **Chapter 5**, this is because neither sample collections’ records contain any LCSH or any other form of subject notation recorded in the MARC coding. Without this specific metadata, both persona types noted that the only way they could possibly see if the results returned fit their needs in this scenario was via a simple match between their keyword search for the word “religion” and the word “religion” in the title. Additionally, there are further relevant items in both sample collections, but the lack of LCSH or some similar

form of subject notation means that Koha's search engine—or indeed, any other search engine—is unable to properly identify any items that may fit. A further discussion regarding the clear need to include LCSH, or some other form of subject notation, in special collections records can be found further along in this chapter.

The precision similarly decreases when comparing CS1 to CS2, though in this case it is not as misleading as the recall change for this scenario. Here, the basic level of metadata detail used to catalogue the sample collection for CS1 meant that the only relevant result returned in the search queries for both persona groups was in fact the only result returned overall. The more detailed metadata framework used in CS2, however, as it included the use of the entire title as printed on the item's title page, meant that the only search returning anything at all in this scenario returned *two* results, with only one of those results being relevant to the query. This does mean that the framework used in the second case study resulted in a less precise result demonstrated on the Koha search engine, as seen in **Table 6.1**.

A further discussion regarding subject headings and notation when it comes to is found further along in this chapter, but the important point at this moment is that while the other five scenarios used in this project display an increase in recall, this scenario displays the need for even *further* detail in metadata used to catalogue rare books and special collections.

6.4.1.4 Scenario Four

Scenario Four returns to the same pattern as Scenarios One and Two, with a marked increase in the number of verified results returned for both persona types. In CS1, publication information outside of year and location was not included in the metadata, while records used in CS2 also included the name(s) of the item's publisher(s) as part of the more detailed metadata employed for the sample collection. As such, it was significantly easier for both persona types to not only find the true number of items held in the sample collection published by the chosen publisher, but also to verify that the results returned were accurate to the needs of the prompt—when the same searches were conducted in CS1, only a single result was returned out of the two relevant items in the sample collection, and it was noted that it could only be considered a *potential* result. This is because the match made by the search engine lay in a handwritten note of unknown provenance on the slip that only consisted of the name “Elzevirs” with no other explanation of what this meant in relation to the item being catalogued, meaning the personas would require further examination of the physical item (another step) to achieve true verification of the result. On the other hand, given the more detailed metadata employed in CS2, the need to complete an additional verification step for the three results returned is eliminated entirely, as the information needed is already noted and catalogued in a clear and concise manner. Additionally, the results from CS2 reveal that there are an additional two

items held in the sample collection that fit the prompt specifications, where CS1's potential result is the only one returned for both persona types' queries.

This scenario displays the exceedingly rare occurrence of a metadata framework that results in a system with perfect recall and perfect precision. While the recall is shown to have increased from the first case study to the second, however, the precision remained the same for both persona types between the two case studies. All this means in this scenario is that both frameworks used in the cataloguing metadata resulted in equally pure abilities to recall relevant results, but the significant increase in recall values shown in **Table 6.1** display the real benefit of the more detailed metadata framework. In a real-world application of the detailed metadata framework developed through the work with the Stirling Maxwell collection to the entire 3,000+ item collection would most likely also display a higher recall value than if the metadata framework used in the slip index was carried on from the paper slips into the digitised finding aid.

6.4.1.5 Scenario Five

Looking at the results of the fifth scenario in CS1 compared to those in CS2 shows another clear difference in the effects of the more detailed metadata. The metadata used to catalogue the sample collection in CS1 was comprised solely of the information recorded on the slip index slips, which did not include any description of the physical items themselves, so there was no information about conservation status for the search engine to find. On the contrary, as the metadata

used for CS2 was taken from a spreadsheet that compiled information taken from and about the items themselves that did include a brief indication of conservation status, the accuracy of the searches performed in CS2 increased significantly over those performed in CS1. Concurrently, the description of the physical items that was also included in the bibliographic records for CS2 additionally allowed the librarian personas to compile a brief explanation of the potential conservation needs as requested by the in-house conservation team as part of the scenario. This information was not present in the bibliographic records in CS1, which again meant that the librarians were erroneously led to believe that no items needed urgent conservation despite the presence of seven relevant items in the sample collection for CS1.

The precision in this scenario—out of all the scenarios used—shows the largest increase, going from a 0% precision in CS1 to a value of 100.0% in CS2, again resulting in a demonstration of the rare occasion of an information retrieval system with perfect recall and perfect precision. Just like the previous scenario, a real-world application of the framework used in CS2 to the entire Stirling Maxwell collection would most likely result in a higher recall value, and the perfect precision value demonstrated here backs the theory that more detailed metadata used for cataloguing rare books and special collections is needed to truly meet the needs of the department's users simply through the presence of such a drastic increase from one case study to the next.

6.4.1.6 Scenario Six

Similarly to Scenario Five, Scenario Six again highlights the difference that more detailed metadata creates between the two case studies. Once more, the metadata used in the creation of the CS1 Koha sample collection only included information written on the slips—in this case, there was nothing recorded on these slips that spoke to either the binding used for the item, nor the presence of any form of handwritten annotations. This lack of in-depth, detailed metadata again led both librarian personas to the false conclusion that no items held in the sample collection fit the needs specified in the scenario prompt, when in actuality there were three relevant items held in the CS1 sample collection. The searches performed in CS2 were able to recall three out of three relevant items in the sample collection—again, the items chosen for both case studies were the same, with the only difference being that the slip index would have one slip for an entire work that was published in multiple volumes, while the spreadsheet had individual entries for each volume of a work.

The change in precision from CS1 to CS2 in this scenario most closely mimics that shown in Scenario Two—there is an increase upon application of the new metadata framework, but the precision in CS2 is still a relatively low value overall at 37.5%. This increase in itself adds weight to the argument posed throughout this thesis, and the relatively low value adds weight to the argument that there is still

further work to be done on better and better metadata frameworks that librarians and cataloguers can use in special collections.

6.4.1.7 In summary...

Taking the results of all six scenarios in both case studies, clearly the increase in detail of the metadata used for cataloguing proves a significant positive difference in the ability of both patron and librarian personas to search for items in five out of the six scenarios. The only scenario where the increase in detail failed to aid in more accurate and detailed search results was Scenario Three, where the need for an even higher level of granularity in the detail used while cataloguing was highlighted—even in this scenario, however, the raw results remained equal between the two case studies rather than the more detailed metadata having any sort of negative effect on the personas' search results. This reinforces the overall argument posed in this thesis that rare books and special collections require a significant increase in the level of detail and granularity of the metadata used to catalogue the items under their purview.

Similarly, the overall increases in precision in the majority of the six scenarios also reinforces the argument posed in this thesis. While the increase noted in some of the scenarios is relatively small, the difference between the two case studies is still clear that the higher detail in the bibliographic records used in the newer metadata framework results in a catalogue that more closely fits the stated needs of the two most common users of special collections, as defined in **Chapter**

4. Even the sole case where the precision decreased from the first case study to the second still demonstrates the need for a more detailed metadata framework in rare books and special collections cataloguing.

Now that this positive effect has been examined, the next section discusses limitations that were not only experienced during the process of this research, but also limitations that may appear in future applications of the methodology used herein.

6.5 Limitations, recommendations, and other reflections

There were a significant number of limitations encountered in both cataloguing systems used for this project, some of which could prove as more than simple stumbling blocks in real-world situations. This section will begin with discussing the limitations faced in the process of answering each research question (**Table 3.1**) in order and the issues they posed for the research conducted herein, along with other recommendations and reflections gleaned throughout this process. Some potential solutions accessible to future researchers will also be included for some of these specific limitations. Following this is a discussion of the possible issues these limitations could pose for both this collection as well as other special collections held elsewhere.

6.5.1 Question One limitations and recommendations

This section discusses the limitations experienced in the process of the creation of the initial test collection of 509 items, the selection of items for the sample collections used in CS1 and CS2, and creation of the two Koha catalogues.

6.5.1.1 Cell size

As shown in **Table 4.7** about the mapping of the six core DCRM(B)3 areas onto MARC field entries, the “Note” area described in the standards encompasses all notes from the cataloguer in one area alongside notes about the item. In a scenario where a cataloguer is using a program like Microsoft Excel, rather than a KOS like Koha, the number of notes from the cataloguer (i.e., the source for the title proper) can prove prohibitive for times where the spreadsheet may need to be accessed in a physical manner rather than electronically. For times where the spreadsheet must be printed out, the amount of information entered into each cell of the spreadsheet is limited by the maximum possible size each cell can be. For the version of Excel used throughout the cataloguing of the Stirling Maxwell collection—Version 16.99.2—the maximum possible height for a cell is set at 5.68 inches (409 pixels), while the maximum cell width is set at 21.31 inches (1,534 pixels). This maximum width is not a realistic size for a printed spreadsheet however, as even when printed in a landscape orientation and assuming the ability to print with no margins, a piece of A4 paper only measures 11.69 inches tall, meaning that to print at the maximum possible cell width offered by Excel would

require at least two pieces of A4 paper taped together to display a single cell in its entirety. Additionally, using the same orientation and assuming that each row is set to the maximum height of 5.68 inches, the width of one sheet of A4 paper is 8.27 inches, so each paper would only be able to display 1.46 cells. Expanding this layout to a full catalogue of every item in the test collection of 509 items would require more than 1,000 pieces of A4 paper to display a single cell per item alone—the true number of sheets needed to print out all six cells per item would be *significantly* higher.

Therefore, while the actual maximum cell size in Excel would be sufficient to contain the “Note” area for an item as described in the DCRM(B)3 (Committee, 2011), this is not feasible in a print-ready spreadsheet catalogue. It is also not feasible to maintain two versions of an Excel spreadsheet catalogue for each collection for multiple reasons that will be discussed in detail in a later section. This particular feature was the major driving factor in the decision to split the “Note” area defined by the DCRM(B)3 (Committee, 2011) into the multiple subfields that map onto the six core areas shown in **Table 4.1**.

6.5.1.2 Need for more detailed categories

In addition to the need to have a print-ready spreadsheet catalogue, the other imperative noted by the librarians in the department was the need for categories that were far more specific than the general “Note” area. While systems like Koha that are specifically designed for the express purpose of cataloguing have the

ability to create subfields that are connected to an overarching field to specify or further define the field (i.e., MARC field 245 subfield c separates the statement of responsibility from the remainder of the title and defines it as such), Excel does not have the ability to create a “sub-column” or “sub-cell” that functions in the same manner as a MARC subfield. Therefore, to document the various categories needed to create complete records as requested, the six core DCRM(B)3 areas needed to be split into further subcategories, and those subcategories then had to be treated as categories in their own right. The new categories mapped cleanly onto the six areas as seen in **Table 4.1**, and further mapped cleanly into the MARC coding used by the Koha system, as seen in **Table 4.8**.

6.5.1.3 Need for departure from DCRM(B)3 standards

While best attempts were made at adhering to standards set in the DCRM(B)3 (Committee, 2011) throughout the cataloguing of the test collection and two sample collections, some standards proved to be too time consuming to be applied in full throughout this project. The largest departure from the standards fell in the field of punctuation—0E lays out the first of a set of prescribed punctuation rules to be used throughout a DCRM(B)3 record at any level (Committee, 2011: p.33). These rules were overruled by the needs of the special collections librarians at the Mitchell Library, and also proved to cause errors in both the Excel spreadsheet and Koha records when implementation was

attempted. As the punctuation of a record does not affect its recallability, this omission is not considered a limitation.

In addition to the punctuation rules, the other rule outright omitted was that of measurement of the physical size of the items. This rule was omitted due to its time-consuming nature, and proved to have no effect on recallability as seen in **Chapters 4 and 5** and summarised in **Table 6.1**, as at no point was size mentioned as a consideration in interviews with the two user groups examined in this research.

Finally, if a field was noted as “mandatory if applicable” (Committee, 2011), the choice as to its inclusion in the Excel spreadsheet fell to the needs of the librarians in the department, due again to the sheer time consumption these inclusions would require. For example, series added entries in the 8XX MARC fields are noted as “Mandatory if applicable” (Committee, 2011: p.172). In the case of the creation of the Excel spreadsheet for this collection, the fact that only one researcher was able to work on this project as well as the needs of the department librarians to have an easily printable spreadsheet meant that omission was needed. This process was followed throughout application of DCRM(B)3. In a real-world scenario, stricter adherence to the standards would likely be required, depending on the policies and procedures prescribed to the library staff, but can be excused in this project due to the previously stated reasons.

6.5.2 Question Two limitations and recommendations

This section discusses the limitations encountered in the process of answering RQ2, specifically in the processes of literature review, constructing interview questions, soliciting and conducting interviews, NVivo coding, and the creation of both personas and scenarios to be tested.

6.5.2.1 Literature gap

When conducting the literature review contained within **Chapter 2**, a significant gap appeared regarding specific information-seeking behaviour (ISB) of special collections librarians that handle local history collections specifically.

As a result of this gap in the literature, while a list of literature-based codes could be utilised for the creation of the patron personas taken from work done by Rhee as an expansion on Meho and Tibbo, who were in turn expounding upon Ellis (Ellis, 1993; Meho and Tibbo, 2003; Rhee, 2012), the same could not be done for the creation of the librarian personas. Instead, as the discussed differences between the librarians at the Mitchell Library and librarians managing other forms of highly specialized collections are significant as explained in **Chapter 2**, the same literature-based codes used for the patron personas needed to be used for the librarian personas, with a stronger reliance on the in vivo codes created in the first examination of the interview transcripts.

6.5.2.2 Differentiation of personas within persona type

Referring to **Chapter 4** when discussing the personas used in this project, the major limitation in their creation and use falls to the topic of differentiation within the types of personas. The two librarian personas³⁶—Donna MacLeod (senior special collections librarian) and Marie Gilligan (junior special collections librarian)—both approached the six scenarios in each case study in strikingly similar ways, with the most obvious difference being Marie often needing to approach other librarians regarding the specific language to use in her searches within the Koha system. The two patron personas³⁷—Joseph Callow (photographic historian) and Dr Sara Childs (Victorian historian)—had an even higher degree of similarity between them.

One potential explanation for this level of agreement between the two personas within each type could come down to the extremely small sample size chosen for the interviews conducted prior to coding analysis. The justification for this sample size discussed in **Chapter 3** still stands, but it is still necessary to recognize the impact of this decision on this research.

In addition to the small sample size for both populations affecting the ability to differentiate between the two personas in each type, the system used to

³⁶ The names of these personas are randomly generated and are in no way related to any information that may identify study participants.

³⁷ The names of these personas are randomly generated and are in no way related to any information that may identify study participants.

demonstrate the significant improvement in recall that detailed metadata affords also impacted the discernment. The Koha system and its search engine only forms a *part* of the searching process noted in both the interviews and the literature review—for example, Marie Gilligan is noted to be more likely to need to approach more senior librarians for help regarding specific language used in the bibliographic records prior to actually performing a search in Koha, while Donna MacLeod is noted to most likely already have this knowledge and therefore not need to perform this action. This step, while an important part of the overall searching process, is one that falls outside of the scope of this project, which is aimed at the demonstration of the concept that rare books and special collections require a much higher level of detail in their metadata in order to be truly accessible to their users through a comparison of recall between the current level of metadata (CS1) and the more detailed metadata (CS2). Therefore, the specific focus of this project justifies the tendency of the two persona types to act as two almost “meta-personas,” rather than two persona type groups of two personas each.

6.5.3 Question Three limitations and recommendations

This section discusses the limitations faced in the process of conducting CS1 and CS2 and the examination, analysis, and comparison of their results.

6.5.3.1 Need for further detail in metadata

As noted multiple times throughout **Chapter 5** and in the earlier section of this chapter comparing CS1 and CS2—specifically Scenario Three—the level of detail in the metadata created for the test collection and the sample collection for CS2 is shown to not be the ultimate level required for collections like the Stirling Maxwell collection.

Scenario Three involved the personas searching for items in the sample collections based on their subject matter. Due to the fact that the metadata recorded in both case studies—that based off the slip index and that based on the Excel spreadsheet—did not include any form of notation regarding the subject matter of the items being catalogued. As discussed in **Chapter 2**, the core-level records as defined by the DCRM(B)3 do include subject headings “if applicable” (Committee, 2011), however this project was limited in time. Returning to **Table 4.4** regarding the language breakdown of the 509 texts chosen for the test collection, only 206 of those items are written primarily in English, which is the only language the sole researcher can understand, let alone be fluent in. When assigning subject headings to a record, the DCRM(B)3 instructs the cataloguer to, “Use judgment in assessing the publication and, if appropriate, assign a complement of headings that provides access to the primary/essential subject of the work (as opposed to secondary or tertiary aspects) at the appropriate level of specificity” (Committee, 2011: p.172). Without a significant skill or fluency in the various languages used in

the items held in the collection, understanding the text to the level needed to assign the subject headings from scratch would not be possible, and where it seems possible, could lead to false identification of appropriate subject headings. Inaccurately assigning subject headings to an item in even the sample collection for CS2 (as the CS1 records were based off a pre-created database that did not include this information) could prove detrimental to both internal and external users of the collection by providing them with false information about the items on record.

There are services available to search a limited number of online catalogues for records, such as the Library Hub Discover database from Jisc, but there are limitations to this approach as well. The first issue with this specific service is that it only includes catalogues of “205 UK and Irish academic, national & specialist librar[ies]” (*Library Hub Discover Advanced Search*, n.d.)³⁸, meaning if an item being searched for is not in those 205 catalogues, there is no information available for the researcher to use in assigning subject headings. Additionally, this approach means the researcher is relying on the judgment of other cataloguers, each with their own perspectives and biases that cannot be understood, accounted for, or addressed, as this researcher would have no information on the cataloguers responsible. The final problem here is that the use of pre-assigned subject

³⁸ This database was used for the purpose of obtaining LCSHs for the items in both sample collections to judge their relevancy for Scenario Three and this alone, as discussed in further detail in **Chapter 5**.

headings in this manner for CS2 would go against the guidelines set in the DCRM(B)3, which require the cataloguer to assess the item themselves (Committee, 2011). In situations where a cataloguer is importing bibliographic records from other institutions, relying on the judgment of others is acceptable—here, however, it is not.

There are additional advanced tools available (e.g., BIBFRAME or Wikidata) that allow for linked data, which could adequately address the issues noted in this area. However, these tools are not functionally available for things such as Excel spreadsheets like the finding aid created throughout this research.

Identification and application of accurate subject headings has the potential for increasing the ability of a user to search for and find relevant items held by a library, as noted by Garrett: “Ideally, subject headings enable users not familiar with the literature of a field to identify and gather together relevant works, regardless of their authors, titles, disparate physical locations, and what other major topics they may treat” (Garrett, 2007). The gathering of these works could extend outside a particular collection, which could link items held within that collection to those held in another, to give the user—internal or external—to gain a thorough understanding of items that may aid in their research needs. Therefore, in a real-world situation—with the presence of more staff with a variety of language skills and subject knowledge to draw from—the application of subject headings

could provide an even greater benefit for a collection than even the increase in detail shown in this project.

6.5.3.2 Need for a more “controlled” vocabulary

While all efforts were made by the researcher to maintain a controlled vocabulary when inputting data gleaned through physical examination of an item into the spreadsheet, by no means was the vocabulary used controlled as strictly as necessary for a proper controlled vocabulary. This can most clearly be seen in Scenario Five in CS2, where the librarian personas were asked to create a list of items held in the sample collection that required urgent conservation. In **Table 5.20**, two variations of the phrase “may need” had to be used in the second search in order to ensure as thorough a search as possible, when both variations were used in the records to mean the exact same conservation status. Simplification of descriptive metadata phrases such as this into a more controlled vocabulary could have made searching for these items faster and easier for librarians, especially as these are users that would be the most familiar with the phrasing and vocabulary used in the process of cataloguing items at the library.

The term “controlled vocabulary” in the traditional sense of the term may be too strict for use in some instances here. Harpring defines a traditional controlled vocabulary as, “an organized arrangement of words and phrases used to index content and/or to retrieve content through browsing or searching. It typically includes preferred and variant terms and has a defined scope or describes a

specific domain” (Harpring, 2010). These lists are highly regulated, “because only under certain specific conditions and review processes may the terms within a controlled vocabulary change or grow” (Hedden, 2010: p.279). While most often these vocabularies are kept to fields such as names, publication places, and subject headings to name a few, the argument can be made that the creation of an internal controlled vocabulary of sorts for the physical description of items in the Mitchell Library’s collections would prove beneficial. For example, a simple *pick list*—one form of a controlled vocabulary “often displayed within drop-down boxes for a field, but could display as button or check-box items” (Hedden, 2010: p.279)—with the terms describing conservation status could streamline the cataloguing process. It could also streamline the searching process for librarians attempting to find this information such as in Scenario Five, as having a clearly defined and controlled vocabulary for conservation status would give the librarians the exact search terms they would need to use when faced with a similar scenario in the real world and thus eliminate the need for multiple searches to ensure all the relevant items have been recalled.

```

000 02201nam a22002534a 4500
003 OST
005 20250512135107.0
008 250415b ||||| |||| 00| u eng d
024 7 0 _2 Stirling Maxwell Slip Index Upper Left Number
    _a 9279a
040
    _c al
    _e dcrmb
110
    _a Académie R. des Inscriptions et Belles-lettres
    _9 40
245
    _a Histoire de l'Academie royale des inscriptions et belles lettres,
    _b depuis son etablisement jusqu'à présent. Avec les memoires de litterature tirez des regitres de cette
    Académie, depuis son renouvellement jusqu'en M. DCCX.
    _c L'Academie royale des inscriptions et belles lettres
    _n Tome Premier
260
    _a Amsterdam
    _c 1719
    _b [s.n.]
300
    _a [12], 488p., 11 leaves of plates: ill.
    _b ill., printed annotations, table of contents
500
    _a Publication imprint reads: "MDCCXIX. A Amsterdam, aux depens de la compagnie."
541 1
    _a Sir William Stirling Maxwell, 9th Baronet's son, Sir John Maxwell Stirling-Maxwell, 10th Baronet
    _c donation
546
    _a French
    _b Roman alphabet
556
    _a Associated slip in slip index has added notes: pencilled number upper left corner; front (pen) "Gros de
    Boze (Claude) comp." crossed out and replaced with institution name as author name, "P.T.O" lower right
    corner; back (pen) "2 vols (2nd)", " ' ' ' (4th)", and "Some are pubd. at La Haye"
562
    _a preserved bookplate inner cover with crest and printed "Maxwell of Pollok, Bart." and handwritten "I.12.1"
    and (pencilled) handwritten "93[circled] 4121"; bookplate with Glasgow coat of arms and printed "Glasgow
    Corporation Public Libraries The Mitchell Library"; carbon paper being used as bookmark; (black ink)
    "C [1/c/e?] [8/=?]" on reverse of title page; Mitchell Library stamp with (pencilled) "18" "[initials?]" and
    "064-36ACA" on reverse of title page; has been rebound; "Histoire de L'Academie Royale" "1" and "1719"
    embossed on spine;
    _b preserved bookplate inner cover with crest and printed "Maxwell of Pollok, Bart." and handwritten "I.12.1"
    and (pencilled) handwritten "93[circled] 4121"
563
    _a has been rebound
583
    _x does not need conservation
700
    _a Gros de Boze, Claude
    _e comp.
    _9 41
942
    _2 ddc
    _c BK
    _n 0
999
    _c 37
    _d 37

```

FIGURE 6.3: INTERNAL MARC DISPLAY FOR A RECORD USED IN CS2

While the conservation needs status (contained in the MARC field 583 in **Figure 6.3** above) would benefit from a traditionally controlled vocabulary in the form of a pick list, with its highly regulated list of terms and their arrangement, the information contained in MARC field 562 (**Figure 6.3**) could also benefit from a controlled vocabulary in the *colloquial* sense of the term. As this field is used to contain information about the physical appearance of an item, if this set of categories is used for other collections held at the Mitchell Library, it cannot be *fully* controlled

as defined by Hedden (2010) and Harpring (2010), as the appearance of items held at the library vary significantly from one another. Attempting to control a vocabulary completely here would result in a vast, unwieldy collection of terms that would most likely prove difficult to train on and utilise. Therefore, a semi-controlled vocabulary, perhaps more strictly prescriptive in the order and arrangement of descriptive terms and lightly prescriptive in the terms themselves (i.e., the use of “bookplate” over “book plate”) could again prove helpful in streamlining the cataloguing and searching processes.

There are some controlled vocabularies already established by the Rare Books and Manuscripts Section for use in special collections cataloguing, but these prescribed vocabularies only cover:

- Binding Terms
- Genre Terms
- Paper Terms
- Printing & Publishing Evidence
- Provenance Evidence
- Type Evidence (Committee, 2014)

Employment of the terms contained within these vocabularies is useful for any collection held in any library, as it will best facilitate linkage of items both within the same collection and within the library as a whole, as well as potentially with

items held in other libraries. In an ideal world, these vocabularies would be universally implemented, but in reality, the application of these lists can fall by the wayside in favour of faster and less detailed cataloguing that allows cataloguers to record more items per day than if they were using the more detailed metadata frameworks and the prescribed controlled vocabularies. Additionally, familiarity with the controlled vocabularies, the contained terms, and the relationship designators used may not be present in external user groups like the historians examined herein, and education as to their presence and application can prove not only a drain on time, but also on money—a resource famously in short supply for public libraries such as the Mitchell Library. Therefore, in addition to the use of more informal finding aids such as the Excel spreadsheet developed for the Stirling Maxwell collection, a happy medium between the potential confusion of the exact application of a prescribed controlled vocabulary and the significant demand for time and money posed by the education needed prior to implementation can be found in the creation of a more colloquial “controlled” vocabulary for use at an institution. Whichever path a library chooses after the first step, it is clear that at a bare minimum, more detailed metadata is necessary in cataloguing rare books and special collections to best serve both staff and patrons alike.

6.5.4 Time consumption and financial cost

Time is money, and money is a finite resource in high demand. For public institutions such as the Mitchell Library, funding is hard to come by for the entire

institution, let alone for a specialised department such as rare books and special collections.

The vast majority of sources regarding public libraries and funding agree on this critical issue: funding for libraries is a constant battle. Clayton and Gorman speak briefly on this matter, noting that, “Another issue of concern to many libraries is budget cuts (or at least, failure to increase materials and other funding in line with increases in costs and usage)” (2006: p.140-1). A 2009 *Library Journal* survey “revealed that the biggest challenge to libraries was rising materials prices, followed by escalating energy costs and the increasing demand for services and programs...[and] that many libraries expected materials budgets to decline” (Hibner and Kelly, 2013: p.127). Even without these budget cuts and restrictions, “core funding is seldom adequate for everything a Special Collections service may wish to do. At best, it maintains basic staffing levels and activities” (Cullingford, 2022: p.333). This level of core funding does not cover the costs of projects such as the examination and cataloguing undertaken in this thesis with the Stirling Maxwell collection, despite the clear evidence shown in this project that such work is necessary to properly manage the collections held in the department as well as provide full access to its users.

It is thus necessary for special collections departments to be considered as a vital department in significant need of a seat at the table of budget discussions for the

whole library because of their unique financial needs, some of which have been demonstrated through this research.

The work on the Excel spreadsheet for the Stirling Maxwell collection alone has taken roughly five years of weekly visits. Granted, this was also completed by a single person, but when funding is tight to the point Cullingford (2022) describes and a department is kept at the bare minimum staffing levels, it is highly unlikely that a special collections department would be able to spare more than a single staff member to perform similar work for the *multiple* collections under their purview. At the time of submission of this thesis, roughly 1,500-2,000 items have been catalogued in the Excel spreadsheet finding aid. This leaves approximately 1,600-2,100+ items left in the collection to catalogue, meaning that should this project proceed at the current rate, the collection may be fully catalogued in another five years. This is five more years of work by a single person with staff support, which is a vast amount of time and staff support, all of which translates to more drain on the already-tight funding available to the Mitchell Library as a public library.

As demonstrated by the two case studies discussed earlier in this chapter, the presence of more detailed metadata in the bibliographic records for these items clearly provides better access to these materials that are so vital to both internal and external user groups. The length of time spent on this collection alone also demonstrates the massive amount of time a similar process could take up in a

department that handles multiple similar collections, but the argument must be made that such an ask provides far-reaching benefits to the department, the library, and their patrons.

6.6 What does this mean for this collection?

There is still significant work that needs to be performed on the Stirling Maxwell collection. At the time of writing, the chosen metadata framework has been applied to just under 1200 of the approximately 3,600+ items held in the RBC at the Mitchell Library in an Excel spreadsheet of the same format as the spreadsheet used for the test collection within this thesis, but there are still a large number of items that must be found—again, the collection had been split amongst the various other collections at some point with no documentation of these moves—and properly catalogued in the more detailed format. Additionally, it has been shown through both the relative failure of Scenario Three in both case studies and the literature discussed by Garrett (2007) that there is a need for even more detailed metadata to be collected and recorded for these items.

This work shows how beneficial such data is and can be for a collection in the facilitation of access for multiple user groups, and could be used to demonstrate the clear need for further such work to be completed on the Stirling Maxwell collection.

6.7 What could this mean for special collections?

That is not to say that the work done at this point is not significant or meaningful. In fact, the work undertaken in this research clearly demonstrates the chasm between a library's holdings and its users in the special collections department. This chasm is one that may not be well understood by staff in higher positions of authority than the departmental librarians, but this work provides evidence for those librarians to cite in discussions regarding the department needs in the future.

After all, "Libraries exist to link patrons with the information they need...Everyone who works in the library has the same objective: increase access to information. And I mean everyone, from the custodians to the director" (Berger, 2014: p.9-10). The current form of cataloguing in place (i.e., scattered finding aids, a bare-bones online catalogue), though helpful in its own way, has been born out of tight financial and staffing constraints. The staff are able to assist their patrons with the current cataloguing structure, but the application of a more detailed framework clearly shows that the departmental librarians will be able to provide *better* access to their patrons, which will allow them to better fulfil their primary aim as librarians.

It is the hope of this researcher that the work done herein will strengthen the argument of special collections librarians for an increase in funding and staffing

levels with the purpose of strengthening cataloguing for collections like the Stirling Maxwell collection.

6.8 Summary

This chapter began with a reminder of what the research aims and goals were for this thesis. It has then taken the results of the research conducted in the literature review (**Chapter 2**) through the information-seeking behaviours and needs analysis of the chosen internal and external user groups and subsequent creation of personas and scenarios (**Chapter 4**), and culminated with an analysis of the two case studies performed on the two sample collections performed in **Chapter 5**. This work was not without its flaws, a discussion of which (and their potential solutions) then formed the next section of this chapter. Finally, the knowledge gained throughout this research is then applied to a hypothetical future for the Stirling Maxwell collection, the Mitchell Library, and the overall field of special collections.

Throughout this chapter—and indeed, the thesis as a whole—the throughline remains that this is a clear, unequivocal demonstration of the point made by so many researchers: rare books and special collections cataloguing requires not only a highly-specialised knowledge set in the librarians tasked with caring for these collections, but it also requires a unique level of detail in order to adequately meet the unique needs of their users.

Chapter 7: Conclusion and Recommendations

Following the information laid out in this thesis, the argument that rare books and special collections cataloguing requires a highly specialised skill set and depth of knowledge to properly provide access to items for both internal and external users is made clear. Without this level of detail contained in the metadata used in the bibliographic records, users of the department are more likely to obtain a false image of the holdings and their availability, which can lead to frustration and disuse. This chapter will sum up the entirety of the thesis and summarise the argument being made within.

7.1 The initial problem

The beginning of this work came in 2020, as discussed in the introductory chapter of this thesis and in the previous work described in the researcher's master's thesis (Charney, 2020). In brief, the researcher was asked by staff of the Mitchell Library to help in updating and modernising the department's records of one of its collections: the Stirling Maxwell collection (Taylor, 2020b).

This collection is comprised of approximately 3,600+ items as indicated by its initial cataloguing system, which is called a slip index. This collection had been donated to the library by the owner's son upon the owner's death, but very little documentation of the donation is present past the slip index and the information that the items were donated to the Mitchell Library in multiple batches over an unknown period of time. In addition to this, the collection had not been kept as a single unit—rather, it had been split up and stored in a number of other collections within the RBC at the library, such as the Early Foreign Printing collection and Wing Collection—which further confused the staff as to the physical location of the Stirling Maxwell items in the library's holdings and their physical condition. Therefore, work was needed to create an up-to-date digitised finding aid in the form of an Excel spreadsheet that took the information contained in the slip index and combined it with information taken from the items themselves as they were located. This would then clarify for the special collections staff what the library still held in the RBC—as it is believed some items have left the holdings for one reason or another over the years—and their potential conservation needs, with the hope of this finding aid then being a resource for future use by both staff and patrons until the collection can be entered into the full online catalogue.

7.2 Research questions

Returning to the beginning of this thesis in **Chapter 1**, the research questions established by this problem are the following:

1. What qualifies a collection as one that falls under the purview of “rare books and special collections” departments?
 - a. Are there any key features unique to these types of items or this type of collection that sets them apart from conventional, modern-day publications and collections?
 - b. How can these unique features be captured and recorded in the form of metadata?
2. What is the methodology by which librarians may identify who the typical users of rare books and special collections departments are, and what specific information needs they may require from these collections within these departments?
 - a. What features of rare books and special collections items are of particular interest to the identified users?
 - b. Are there any changes to the metadata captured in question 1b required after identification of users and user information needs?
3. How can the unique features of rare books and special collections work with their associated metadata to meet the identified user information needs and improve access to rare books for these users?

7.3 Established knowledge

Collection management is a wide field of knowledge, of which metadata and cataloguing are only a small part. Beyond that, the practice of cataloguing must change along with the type of item being catalogued due to the unique needs of the users who access them. As a subset of book cataloguing, rare books and special collections cataloguing requires a greater and more specialised skillset than that of mainstream books, as the users who interact with these items use them not only as carriers of information, but as valuable information in and of themselves.

The importance of this fact is expounded upon *ad nauseam* in various sources, including Berger (2014) and Cullingford (2022). Both resources dedicate entire chapters to information about the type of vertical metadata required in this field, discussing topics of various paper manufacturing styles, type printing styles, binding styles, etc. This depth of knowledge is not required in cataloguing modern items, as the variation in manufacturing styles has shrunk to an almost negligible amount compared to the variations present in the field of special collections. The multitude of manufacturing differences found in items like those in the Stirling Maxwell collection forms but one form of information that users accessing these items may be looking for and therefore requires specific metadata frameworks and formats to catalogue appropriately.

7.3.1 Gaps in the literature

Through the literature review conducted in **Chapter 2**, multiple gaps in the literature were identified. Using the overall methodology laid out in **Chapter 3** to inform the areas examined in the literature, three main gaps were identified:

1. There is no research around the information-seeking behaviours of special collections librarians, one of the user groups chosen for examination in this thesis.
2. Both Berger (2014) and Cullingford (2022), amongst others, as stated above, thoroughly discuss the need for specifically detailed metadata in cataloguing these items, but real-world, practical examples simulating important aspects like time consumption do not exist.
3. Following the previous gap, there is no known tested procedure for other rare books and special collections librarians to follow for use in their own cataloguing projects, whether at the Mitchell Library or elsewhere.

The work done in this research has clearly addressed all three of these main gaps. While there is still ample opportunity for further research to be done in each area to continue narrowing the gap in knowledge, this project has not only shown the glaring need for such research, but also shown that the tactics used within give clear benefits for both librarians and their patrons.

7.4 Contributions to knowledge

Following the identification of the three main gaps in the literature discussed above, steps were taken to address them throughout this project. This section will explain how these processes and their results form unique contributions to the fields of information science and special collections.

7.4.1 Vertical nature of special collections metadata

Regarding RQ1 about what features separate rare books and special collections from mainstream publishing today, a large portion of these features are found in the items themselves. Variations in the manufacture of mainstream contemporary books have shrunk to a near negligible level compared to the same types of manufacturing variations present in books held in rare books and special collections departments. Researchers rarely need to know how a book published today is bound, while that can be valuable information in the world of historical research.

The importance of this information for the patrons is emphasised by both Berger and Cullingford. Berger dedicates an entire chapter of his book to information about the various types of manufacturing of paper, binding methods, printing type, and even ink (2014), as does Cullingford (2022). Information like this is invaluable to researchers like those studying book history and bibliography, so having that information available as captured metadata present in the catalogue is a necessity

for libraries to truly meet the needs of their users. This necessity is clearly demonstrated in a real-world simulation in this research, emphasising the importance of the significant level of detail needed in cataloguing this field.

7.4.2 Information-seeking behaviours

The second research question answered in this project regards how both internal and external user groups search for information in rare books and special collections departments, as well as the identification of a potential methodology for others to follow in future research. This methodology can also be used by librarians attempting to improve their own libraries' abilities to serve their communities.

As discussed throughout this thesis, there is no literature around the information-seeking behaviours of special collections librarians handling historical materials, only around these behaviours in the historians they serve. This project tested the viability of the historian ISBs as ISBs for these librarians, with some success, as discussed in **Chapter 6**. In the future, this could prove a potential route for further research in the area.

Likewise, while this project has only examined the information-seeking behaviours of a very select sample population of special collections librarians, the knowledge gleaned from this process already shows its usefulness. Given the unique information needs of the patron group selected for this thesis when accessing

items held in rare books and special collections—notably, the in-depth information contained in the very structure of the items themselves—special collections librarians have had to adapt the standard forms of ISBs to accommodate them. Further research in this area would certainly expand the sample population examined, which would then increase the field’s understanding of the general population of special collections librarians overall.

Returning to the point made in **Chapter 6** regarding the general financial issues facing specialised library departments such as this field and the ever-looming presence of budget cuts facing public libraries, it is clear that such slashes to funding are causing both immediate and lasting harm to every person involved in the care and study of such items.

7.4.3 Creating a procedure for others to follow

As mentioned numerous times in this thesis and in the general literature, funding and staffing are significant issues facing libraries globally, not just the Mitchell Library. Cataloguing projects like the one undertaken with the Stirling Maxwell collection require large amounts of time and staff to complete, both of which cost money. This cost can make some organisations balk at the request; however necessary the projects are.

A large part of such cataloguing projects can be taken up by the design process. Having an example of a process to follow can streamline the design

process—minimising this stage of the project can result in more of the allotted funds being spent on more vital parts of the project like the cataloguing itself. Being that this research takes the theory given by other researchers and applies it to a real-world situation, the process herein can thus serve as a loose roadmap of sorts for other librarians to follow when cataloguing their own collections elsewhere.

7.5 Future research opportunities

Through the limitations and stumbling blocks experienced in the process of this study, the researcher has concluded that there are some significant opportunities for future research in this area. This section divides these opportunities between those present for the field of special collections and those present for the wider field of information science and discusses them in detail.

7.5.1 For special collections

There are multiple opportunities for further research in this field unveiled by the work completed herein, two of which are discussed below.

7.5.1.1 Wider use of detailed metadata

Firstly, research into the possible wider use of the chosen detailed metadata framework both for other collections held at the Mitchell Library as well as those held in other libraries is highly recommended. The goal of this project was

primarily to examine the exact usefulness of a higher level of detail of metadata applied to a collection with an existing, primitive set of metadata, which meant that this work was highly specific. Future work could expand this work to other collections to examine its wider, generalised use, as it could perhaps lead to updates to the current metadata standard of DCRM(B)3. Additionally, given the long-standing and well-known financial constraints placed on the field of special collections discussed previously, examination into the use of informal finding aids could adequately bridge the gap between a complete lack of cataloguing and full online cataloguing while perhaps reducing some of the ask such work requires of a library's budget could be completed here as well.

6.8.1.2 Subject headings

The need for an even higher level of detail in the metadata used to record and catalogue rare books and special collections—most notably around subject headings—has been discussed *ad nauseam*. However, further research in this area is still required, as noted not only within this thesis, but by other researchers as well. Work in this area is indeed ongoing, as Garrett mentions a peculiar gap in the subject indexing in the field, stating, “While the English Short Title Catalogue (ESTC) contains Library of Congress Subject Headings (hereafter LCSH) for books from the English-speaking world printed before 1701, entries for material dating from 1701 to 1800 are not subject indexed” (2007: p.70), though it is vital to note that this information is outdated at the time of this doctoral research. According to

the ESTC home page, it “is the most comprehensive resource available for the output of the printing press in the English-speaking world before 1801” (*English Short Title Catalogue (ESTC)*, 2025), meaning that the gap noted by Garrett has closed sometime in the years between 2007 and 2025.

The question remains if use of LCSH in their current iterations is truly appropriate for historical items, as:

...title keyword searches miss an enormous amount of relevant material, since words we might use to describe a historical topic today are often entirely different from words populating actual eighteenth-century book titles. For a number of reasons, some having to do with changes in the lexicon, some with a century-specific perceived need for circumlocution, words such as “hygiene” and “prostitution” occurred far less frequently in the eighteenth century than they do today—not to mention the often disastrous effects of pre-1800 orthography on modern day keyword searches (Garrett, 2007: p.70).

LCSH in their status as “the most widely used system for providing natural language subject access in libraries in the English speaking world” (MacEwan, 2000: p.199-200) are, by nature, difficult—if not impossible—to apply directly in cataloguing systems in the non-English speaking world. Creation of a similarly thorough indexing language for historical items like those most commonly held in

rare books and special collections departments would be of massive benefit to the field, and poses plenty of interesting questions for researchers to investigate.

7.5.2 For information science

If there had been a set of literature regarding the ISBs of rare books and special collections librarians, the section of this research dealing with the interviews and transcript coding of the librarians could have proceeded in a much more directed manner than it did in this thesis. It is also of note that this is an under-researched population once the fact of the highly specialised knowledge required in such a position is taken into consideration. While this research has made the first step towards expanding the existing knowledge on this population's ISBs, as discussed in **Chapter 6, Section 6.3.2**, it serves more as a call to researchers regarding the desperate need for a greater understanding on the matter.

Such research could be conducted in a similar manner as that laid out in **Chapters 3 and 4**, expanded to a larger sample population that could more easily be generalised as a representative of the entire population. It could also be conducted in the manner of Meho and Tibbo (2003), Ellis (1993), or Rhee (2012). The information gleaned from such examination would not only prove an expansion on the current understanding of ISBs, but could also result in new frameworks or discoveries that might echo down the line far into the future.

7.6 The final word

There is no question as to the necessity of a rare books and special collections department in libraries, both public and otherwise. Beyond the invaluable local historical information contained within the items held in such collections, there is also a wealth of information to be found in the very manufacture of these items. This information is useful for numerous users of the department—from staff conservators charged with preserving the aged, fragile items to historians of any type who look to these items as resources, and everyone in between. As such, accurate and detailed cataloguing is required to provide full and proper access to these items.

Terry Walker again explains the extreme importance of these departments in a most succinct and poignant way, stating:

If history is a story we tell ourselves about the past, these artifacts are the fabric of that past, linking elements that may allow researchers to build a factual and balanced understanding of not only what happened, but how it happened and perhaps even why. The opportunity for local researchers to have hands-on interaction with primary source documents in a local setting is not something that can be fully replicated on line. This is the reason these documents have been preserved and so carefully cared for—and this

treasure trove is available for present and future generations to explore at the library (2014)

Current cataloguing methods such as the Stirling Maxwell slip index created sometime at the beginning of this collection's life cycle at the Mitchell Library—while they can contain the bare bones of metadata necessary to catalogue these items—are not sufficiently detailed enough to provide the *true* access patrons need from such special and unique items. This is because the users of these items have unique and specific information needs when accessing items held in the special collections department, most often due to the incredible depth of information both carried within the text as well as contained in the manners of manufacture of the item itself. These needs have been examined and identified through interviews of the two prime user groups involved with the Special Collections department at the Mitchell Library (librarians and historians), and the importance of those needs being met by a metadata framework with more detail than the bare minimum has been shown through two case studies using persona testing. By comparing the two case studies in **Chapter 6, Section 6.4.1**, the massive benefit to both primary user groups is clear: more detailed metadata results in higher recall, which then results in a more positive and productive experience for the user.

Metadata frameworks like the widely used DCRM(B)3 are useful for capturing the immense amount of information contained by a special collections item, but they

are most easily translated to purpose-built cataloguing systems. When a special collections department is faced with tight budgets and minimal staffing, justifying the expenditure of staff, time, and money on thoroughly cataloguing the department's holdings in a system such as Koha is difficult. One option that can alleviate some of these concerns is a more informal solution such as the Excel spreadsheet requested by the Mitchell Library staff—there is no excess training required in the use of the software itself, and this research has clearly demonstrated the ability of the software to properly store the full level of detailed metadata required to meet the department's users' needs.

This all returns to the core tenet of collection management given by Matassa:

“collections management can be summed up as: know what you have and where to find it” (2011). The detailed metadata required by the type of items examined herein tick *both* boxes, allowing rare books and special collections librarians to better manage their collections to meet—and potentially exceed—the unique needs of their users.

References

AACR, T.J.S.C.f.R.o. (2005) '*Anglo-American Cataloguing Rules*'. 2nd edition, 2002 revision. Chicago, Illinois; Ottawa, Ontario; London, England: American Library Association; Canadian Library Association; Facet Publishing.

Adu, P. (2019) *A step-by-step guide to qualitative data coding*. London: Routledge.

Allen, B. and Emerald (1996) *Information tasks [internet resource] toward a user-centered approach to information systems*. San Diego, Calif.: Academic Press. Available at: [https://doi-org.proxy.lib.strath.ac.uk/10.1108/S1876-0562\(1996\)96](https://doi-org.proxy.lib.strath.ac.uk/10.1108/S1876-0562(1996)96)

Annan, T. (c. 1870) '*Sir William Stirling-Maxwell*'.

Arksey, H. and O'Malley, L. (2005) 'Scoping studies: towards a methodological framework'. *International Journal of Social Research Methodology*, 8 (1), pp. 19-32. <https://doi.org/10.1080/1364557032000119616>

Attar, K. (2021) 'Why Cataloguing Is Cool'. *Journal of the Australian Library and Information Association*, 70 (4), pp. 424-430. <https://doi.org/10.1080/24750158.2021.1922128>

baseimage (n.d.) '*Business Portrait of Young Woman*'. Canva.

Berger, S.E. (2014) *Rare Books and Special Collections*. London: Facet Publishing.

Bird, R. (2015) 'Moving or Relocating a Library'. *Legal Information Management*, 15 (4), pp. 260-263. <https://doi.org/10.1017/s1472669615000626>

'Blinken OSA Archivum'. (n.d.). <https://archivum.org/>

Buckland, M. (2017) 'Evaluation of Selection Methods'. *Information and Society*. Cambridge, Massachusetts; London, England: MIT Press, pp. 153-165.

Campbell, M. (1996) *Random Name Generator*. Available at: <https://www.behindthename.com/random/>

'Castle Objects of the Month: Anne Boleyn's Books of Hours'. (2021) Hever Castle & Gardens. Available at: <https://www.hevercastle.co.uk/news/castle-object-of-the-month-books-of-hours/>

Charney, A.B. (2020) *Sir William Stirling-Maxwell: A Case Study of Collection Management Policy Concerning a Victorian Era Collection*. MSc Information and Library Studies Case Study. University of Strathclyde.

Clark, T. (2021) *Bryman's social research methods*. 6th edn. Oxford: Oxford University Press.

Clayton, P. and Gorman, G.E. (2006) *Managing Information Resources in Libraries: Collection Management in Theory and Practice*. London: Facet Publishing.

Committee, R.B.S. (2007) '*Descriptive Cataloging of Rare Materials (Books)*'. Washington, D.C.: Library of Congress Cataloging Distribution Service.

Committee, R.B.S. (2011) '*Descriptive Cataloging of Rare Materials (Books)*'. 3rd printing with corrections. Washington, D.C.: Library of Congress Cataloging Distribution Service.

Committee, R.B.S. (2014) *RBMS Controlled Vocabularies for Use in Rare Book and Special Collections Cataloging*. Available at: <https://rbms.info/vocabularies/index.shtml>

Committee, R.S. (2010) '*RDA: Resource Description and Access*'. American Library Association; Canadian Federation of Library Associations; CILIP: Chartered Institute of Library and Information Professionals. Available at: <https://www.rdatoolkit.org/about>

corelens (n.d.) '*Mindfulness Practices Portrait of Older Woman*'. Canva.

Cornelison, B.R. et al. (2021) 'Accuracy of Google Translate in translating the directions and counseling points for top-selling drugs from English to Arabic, Chinese, and Spanish'. *American Journal of Health-System Pharmacy*, 78 (22), pp. 2053-2058. <https://doi.org/10.1093/ajhp/zxab224>

Crockett, M. (2016) *The no-nonsense guide to archives and recordkeeping*. London: Facet Publishing.

Cullingford, A. (2022) *The Special Collections Handbook*. 3rd edition. London: Facet Publishing.

diversiflylens (n.d.) 'Portrait of a Middle Aged Woman in Brown Coat'. Canva.

Duff, W.M. and Johnson, C.A. (2002) 'Accidentally Found on Purpose: Information-Seeking Behavior of Historians in Archives'. *The Library Quarterly: Information, Community, Policy*, 72 (4), pp. 472-496. <http://www.jstor.org/stable/40039793>

English Short Title Catalogue (ESTC) (2025) Available at: <https://datb.cerl.org/estc>

Ellis, D. (1993) 'Modeling the Information-Seeking Patterns of Academic Researchers: A Grounded Theory Approach'. *The Library Quarterly: Information, Community, Policy*, 63 (4), pp. 469-486. <http://www.jstor.org/stable/4308867>

Flyvbjerg, B. and Sampson, S. (2001) *Making Social Science Matter*. Oxford, UK: Cambridge University Press.

Garrett, J. (2007) 'Subject Headings in Full-Text Environments: The ECCO Experiment'. *College & Research Libraries*, 68 (1), pp. 69-81. <https://doi.org/10.5860/crl.68.1.69>

Glasgow City Archives. (n.d.) Available at: <https://www.glasgowlife.org.uk/libraries/family-history/family-history-at-the-mitchell/glasgow-city-archives>

Harpring, P. (2010) *Introduction to Controlled Vocabularies: Terminology for Art, Architecture, and Other Cultural Works*. 1st edn. Los Angeles, CA: Getty Research Institute.

Hedden, H. (2010) 'Taxonomies and controlled vocabularies best practices for metadata'. *Journal of Digital Asset Management*, 6 (5), pp. 279-284. <https://doi.org/10.1057/dam.2010.29>

Henka, A. and Zimmermann, G. (2014) 'Persona Based Accessibility Testing'. In: Stephanidis, C., ed. *HCI International 2014 - Posters' Extended Abstracts, 2014//2014 Cham*. Springer International Publishing, pp.226-231.

Hibner, H. and Kelly, M. (2013) *Making a Collection Count: A Holistic Approach to Library Collection Management*. 2nd edn. Chandos Publishing.

Huggard, M. (2019) 'Good Intentions: Distinguishing Deaccessioning from Weeding'. In: Jackson, L.U. (ed.) *Reappraisal and Deaccessioning in Archives and Special Collections*. Lanham: Rowman & Littlefield Publishers, pp. 1-11.

Hunter, G.S. (2020) *Developing and Maintaining Practical Archives*. London: Facet Publishing.

Jankowski, A., Schultz, A. and Soito, L. (2018) 'Motley Crew: Collaboration across an Academic Library to Revive an Orphaned Collection'. *Library Resources & Technical Services*, 62 (3), pp. 114-126.

<https://link.gale.com/apps/doc/A548441542/AONE?u=ustrath&sid=bookmark-AONE&xid=d3941bef>

Jesson, J. (2011) *Doing your literature review : traditional and systematic techniques*. Los Angeles, Calif.; London: SAGE.

Johnson, P. (2025) *Fundamentals of collection development and management*. 5th edn. London: Facet Publishing.

Kelly, M. (2015) 'The Materials-centred Approach to Public Library Collection Development: A Defense'. *Library Philosophy and Practice*.

<https://link.gale.com/apps/doc/A529046014/AONE?u=ustrath&sid=bookmark-AONE&xid=5af772e2>

Koha Library Software. (2025) Available at: <https://koha-community.org/>

Laubheimer, P. (2020) *3 Persona Types: Lightweight, Qualitative, and Statistical*. Available at: <https://www.nngroup.com/articles/persona-types/>

Library Hub Discover Advanced Search. (n.d.) Available at: <https://discover.libraryhub.jisc.ac.uk/advanced-search/>

Library of Congress, O.f.D.C.P. and Association of College and Research Libraries, B.S.C.o.t.R.B.a.M.S. (1991) '*Descriptive Cataloging of Rare Books*'. 2nd edn. Washington D.C.: Library of Congress. Available at: <https://babel.hathitrust.org/cgi/pt?id=mdp.39015025368443&seq=5>

Lynch, P., Tomas, P. and Hattenstone, A. (2024) 'Public libraries in 'crisis' as councils cut services'. *BBC.co.uk*, <https://www.bbc.co.uk/news/articles/cn9lexplel5o>.

- MacDougall, I. (2017) *Voices of Scottish librarians : the evolution of a profession and its response to changing times*. Edinburgh: John Donald.
- MacEwan, A. (2000) 'Crossing Language Barriers in Europe: Linking LCSH to Other Subject Heading Languages'. *Cataloging & Classification Quarterly*, 29 (1-2), pp. 199-207. https://doi.org/10.1300/J104v29n01_14
- Mak, S. and Thomas, A. (2022) 'Steps for Conducting a Scoping Review'. *Journal of Graduate Medical Education*, 14 (5), pp. 565-567. <https://doi.org/10.4300/JGME-D-22-00621.1>
- MARC21 Format for Bibliographic Data*. (2000) Available at: <https://www.loc.gov/marc/bibliographic/>
- Matassa, F. (2011) *Museum Collections Management: A Handbook*. London: Facet Publishing.
- Meho, L.I. and Tibbo, H.R. (2003) 'Modeling the Information-Seeking Behavior of Social Scientists: Ellis's Study Revisited'. *Journal of the American Society for Information Science & Technology*, 54 (6), pp. 570-587. <https://doi.org/10.1002/asi.10244>
- Monterde, J. (n.d.) 'An older man wearing a hat and a blue sweater'. Canva.
- Oder, N. (2009) 'Adjustment time: libraries are either tightening budgets or preparing to do so'. *Library Journal*, 134 (1), pp. 32.
- Ovenden, R. (2008) 'Special Collections Librarianship'. In: Earnshaw, R. and Vince, J. (eds.) *Digital Convergence - Libraries of the Future*. London: Springer, pp. 295-302.
- Pearson, D. (2019) *Provenance Research in Book History: A Handbook*. New and revised edn. Oxford: Bodleian Library Publishing.
- Peters, M.D.J. et al. (2020) 'Updated methodological guidance for the conduct of scoping reviews'. *JBI Evidence Synthesis*, 18 (10), pp. 2119-2126. <https://doi.org/10.11124/JBIES-20-00167>
- 'Rare'. (n.d.) *Merriam-Webster.com Dictionary*. Merriam-Webster.

- Rhee, H.L. (2012) 'Modelling historians' information-seeking behaviour with an interdisciplinary and comparative approach'. *Information Research*, 17 (4). pp., <http://InformationR.net/ir/17-4/paper544.html>
- Rhee, H.L. (2010) 'Unique qualities of historians' information-seeking behavior in historical research'. *Proceedings of the American Society for Information Science and Technology*, 47 (1), pp. 1-2. <https://doi.org/https://doi.org/10.1002/meet.14504701362>
- Saldaña, J. (2021) *The coding manual for qualitative researchers*. 4th edn. London : Thousand Oaks, California: SAGE Publications.
- Schleicher, M. (2010) 'Assembling Selection Criteria and Writing a Collection Development Policy for a Variety of Older Medical Books'. *Journal of Hospital Librarianship*, 10 251-264. <https://doi.org/10.1080/15323269.2010.491424>
- Special Collections*. (n.d.) Available at: <https://www.glasgowlife.org.uk/libraries/special-collections>
- Stalker, L. and Dooley, J.M. (1992) 'Descriptive Cataloging and Rare Books'. *Rare Books & Manuscript Librarianship*, 7 (1), pp. 16. <https://doi.org/10.5860/rbml.7.1.75>
- Stenstrom, C. and Haycock, K. (2015) Public Library Advocacy: An Evidence-Based Perspective on Sustainable Funding. *Public Libraries Online*. Available at: <https://publiclibrariesonline.org/2015/09/public-library-advocacy-an-evidence-based-perspective-on-sustainable-funding/>
- 'Stephen Mitchell (1789-1874) - The 150th Anniversary of his death'. (2024) Available at: <https://www.glasgowlife.org.uk/libraries/family-history/stories-and-blogs-from-the-mitchell/special-collections-blogs/stephen-mitchell-1789-1874-the-150th-anniversary-of-his-death>.
- Subject and Genre/Form Headings*. (n.d.) Available at: <https://www.loc.gov/aba/cataloging/subject/>
- Taylor, S. (2023). *RE: Email to Arthur Charney*. Email to Arthur Charney, 30 October.
- Taylor, S. (2020a) *Selection of Stirling Maxwell Slips*.
- Taylor, S. (2020b) 'Student Placement Project Proposal'.

The Mitchell Library. (2026) Available at:

<https://www.glasgowlife.org.uk/libraries/venues/the-mitchell-library>

Thomas, G. (2021) *How to do your case study*. 3rd edn. Los Angeles: SAGE Publications, Ltd.

Unwin, G., Unwin, P.S. and Tucker, D.H. (2024) 'history of publishing'.

<https://www.britannica.com/topic/publishing/The-age-of-early-printing-1450-1550>

Walker, T. (2014) 'Local Treasures: The Value of Special Collections in the Public Library Setting'. *Alki*, 30 (2), pp. 18-23.

Yin, R.K. (2018) *Case Study Research and Applications: Design and Methods*. 6th edn. United States of America: SAGE Publications, Inc.

Zeng, M.L. (2016) *Metadata [internet resource]*. 2nd edition. Chicago: Neal-

Schuman. Available at: [https://research-ebsco-](https://research-ebsco-com.proxy.lib.strath.ac.uk/c/j46aul/search/details/3vn5td7c4b?db=nlebk&db=nla)

[com.proxy.lib.strath.ac.uk/c/j46aul/search/details/3vn5td7c4b?db=nlebk&db=nla](https://research-ebsco-com.proxy.lib.strath.ac.uk/c/j46aul/search/details/3vn5td7c4b?db=nlebk&db=nla)
[bk](https://research-ebsco-com.proxy.lib.strath.ac.uk/c/j46aul/search/details/3vn5td7c4b?db=nlebk&db=nla)

Appendix A

Participant solicitation flyer



PARTICIPANTS NEEDED!

I am a doctoral student at the University of Strathclyde conducting a study on the information behaviour and needs of people who use the special collections department at the Mitchell Library.

IF YOU:

- Are 18 years of age or older,
- Have used the resources within the special collections department at the Mitchell Library at least twice, *and*
- Are able to do a 30 minute to 1 hour interview over Zoom,

PLEASE EMAIL ME TO SCHEDULE AN INTERVIEW!

These interviews will be conducted over Zoom before 31 March 2024.

THANK YOU!

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Arthur Charney
arthur.charney@strath.ac.uk

Appendix B

Example Interview Questions

- How often do they access Mitchell Library special collections?
 - Online? In person?
- Are they looking for something specific or just browsing?
 - If specific → does language factor into it? (i.e. if they know a resource exists in Latin that may be useful, will they then work to find a translation or use the original resource?)
- Do they place importance on handwritten annotations? (not necessarily ownership inscriptions alone but also notes w/in text or possibly notes not about the text) What if the person who wrote them is/isn't identifiable?
- What is their usual process in either specific or general searches?
 - Can they give me a couple scenarios in which they might need to access items/information from this collection? How would they go about doing that? What happens when they find that information?
- Is there any information they look for that maybe isn't already available?
 - i.e. conservation information, physical extent/attributes (like printed annotations, indices, etc)

Appendix C

Participant Information Sheet and Consent Form

Participant Information Sheet

Name of department: Computer and Information Sciences

Title of the study: "Know what you have and where to find it": Exploring information needs of multiple user groups and how they shape better finding aids

Introduction

My name is Arthur Charney, and I am a Computer and Information Sciences doctoral student at the University of Strathclyde. I can be reached directly via email at arthur.charney@strath.ac.uk.

What is the purpose of this research?

You are being invited to take part in a research project which will investigate information behaviours and needs of multiple user groups accessing finding aids in the special collections department of the Mitchell Library and how those behaviours and needs can be used to adjust the structure/format of currently used finding aids. Finding aids are tools meant to help locate specific resources within a collection and can include (but are not limited to) inventories and catalogues. These tools are accessed by multiple user groups, which have been narrowed down to the special collections department librarians and the historians that access the special collections department resources at the Mitchell Library. I will be identifying the major information behaviours of these groups via interview and using them to adjust the structure of a finding aid for use with a Victorian era collection currently held at the library. This research will inform my work towards earning a Doctor of Philosophy degree in Computer and Information Sciences at the University of Strathclyde.

Do you have to take part?

Your participation is entirely voluntary, and neither refusing to participate nor withdrawing participation will affect any other aspect of the way you are treated. You entirely retain the right to withdraw from the research without detriment to yourself for whatever reason.

Upon completion of the interview, you will be provided with a period of 30 days, commencing from the day of interview completion, to deliberate your continued participation. It is essential to note that any withdrawals occurring beyond this designated timeframe would have implications on the final research product. Therefore, you are urged to carefully consider your decision within the allocated timeframe to ensure the integrity and reliability of the study.

What will you do in the project?

This study will consist of one interview conducted virtually via Zoom. This interview will take approximately 30 minutes to 1 hour, and will be conducted by the end of the day on 31 March, 2024.

I may contact you again sometime in 2024 to evaluate the new finding aid.

Why have you been invited to take part?

You have been invited to take part because you either work as a librarian in the special collections department at the Mitchell Library in Glasgow or are an adult (18 years of age and over) member of the public who access the resources held by the department.

What information is being collected in the project?

Your responses in the interview will be recorded via an audio recorder and transcribed, but no personally identifying information will be collected or retained at any point. My contact with you will be the only time any information of any sort will be collected, and you will always be made aware of my recording of any information I collect, when I collect it.

Who will have access to the information?

All data will be handled in full confidence and anonymity will be preserved. All data will be kept in a password protected computer file and only viewed by myself and my research supervisors. No names or direct quotes will be used in any unpublished or published writing about this research project.

Where will the information be stored and how long will it be kept for?

All data will be kept in a password protected computer file and only viewed by myself and my research supervisors. All data will be kept until 31 December, 2035. Following this date all data will be wiped and deleted.

Thank you for reading this information – please ask any questions if you are unsure about what is written here.

What happens next?

If you would like to find out more about this project, feel free to contact me by email with any questions.

Researcher contact details:

You can contact me, Mr. Arthur Charney, at arthur.charney@strath.ac.uk .

Chief Investigator details:

You can contact my primary supervisor, Professor Gobinda Chowdhury, at gobinda.chowdhury@strath.ac.uk .

You can contact my secondary supervisor, Professor Ian Ruthven, at ian.ruthven@strath.ac.uk .

This research was granted ethical approval by the Department of Computer and Information Sciences Ethics Committee.

If you have any questions/concerns, during or after the research, or wish to contact an independent person to whom any questions may be directed or further information may be sought from, please contact:

Secretary to the Department Ethics Committee
Department of Computer and Information Sciences
University of Strathclyde
Livingston Tower
26 Richmond Street
Glasgow
G1 1XH
Email: ethics@cis.strath.ac.uk

Consent Form

Name of department: Computer and Information Sciences

Title of the study: "Know what you have and where to find it": Exploring information needs of multiple user groups and how they shape better finding aids

- I confirm that I have read and understood the Participant Information Sheet for the above project and the researcher has answered any queries to my satisfaction.
- I confirm that I understand how my personal information will be used and what will happen to it (i.e. how it will be stored and for how long).
- I understand that my participation is voluntary and that I am free to withdraw from the project at any time, up to the point of completion, without having to give a reason and without any consequences.
- I understand that I can request the withdrawal from the study of some personal information and that whenever possible researchers will comply with my request. This includes the following personal data: audio recordings of interviews that identify me; my personal information from transcripts.
- I understand that anonymised data (i.e. data that do not identify me personally) cannot be withdrawn once they have been included in the study.
- I understand that any information recorded in the research will remain confidential and no information that identifies me will be made publicly available.
- I consent to being a participant in the project.

I consent to being audio and/or video recorded as part of the project

(PRINT NAME)	
Signature of Participant:	Date:

Appendix D

Template for persona creation

	Name Occupational Title
Age:	Education:
Sex:	Occupation:
Length of Employment:	Location of Employment:

BIOGRAPHY

[Brief description of person]

WORK DUTIES

- [Work duty 1]
- [Work duty 2]
- [Work duty 3]
- [Work duty 4]

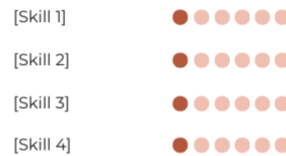
SEARCHING BEHAVIOURS

- [Searching behaviour 1]
- [Searching behaviour 2]
- [Searching behaviour 3]
- [Searching behaviour 4]

PERSONALITY



SKILLS



TASTES AND HOBBIES

			
[HOBBY 1]	[HOBBY 2]	[HOBBY 3]	[HOBBY 4]

Appendix E

Senior librarian persona



Donna MacLeod
Special Collections Librarian

Age: 64 years	Education: BA Hons, MSc
Sex: Female	Occupation: Special Collections Librarian
Length of Employment: 23 years	Location of Employment: The Mitchell Library

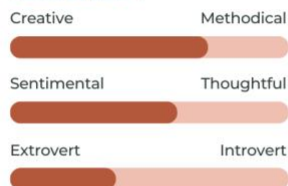
BIOGRAPHY

Graduated from a CILIP-accredited MSc programme. Has worked in two other libraries prior to current job, both in special collections. She particularly enjoys working with uncommon bindings, and often travels to view exhibitions on the subject.

SEARCHING BEHAVIOURS

- Most likely to have prior specialist knowledge about select subjects
- Still likely to use other staff for help, especially if looking for subjects outside their knowledge
- More likely to use narrower terms for initial searches
- More likely to take less time when searching

PERSONALITY



WORK DUTIES

- Patron interaction and handling patron resource inquiries
- Curation of resources based off of patron requests and needs
- Significant collection management duties (i.e., keeping up with preservation/conservation needs of the items, deaccession/weeding, loans, etc.)
- Public outreach and collection promotion

SKILLS



TASTES AND HOBBIES



Appendix F

Junior librarian persona



Marie Gilligan
Special Collections Librarian

Age: 35 years	Education: BA Hons, MSc
Sex: Female	Occupation: Special Collections Librarian
Length of Employment: 6 months	Location of Employment: The Mitchell Library

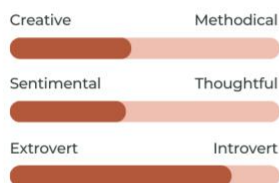
BIOGRAPHY

A recent graduate of a CILIP-accredited MSc programme, she was hired right out of school. This is her first job in the field, having worked previously in retail and food service prior to obtaining her degree.

SEARCHING BEHAVIOURS

- More likely to rely on help from other, senior librarians/staff as she has very little, if any, knowledge/understanding of the library's holdings
- Might cast a broader net when searching
- More likely to take a longer time while searching

PERSONALITY



WORK DUTIES

- Patron interaction and handling patron resource inquiries
- Curation of resources based off of patron requests and needs
- Minor collection management duties (i.e., keeping up with preservation/conservation needs of the items)
- Public outreach and collection promotion

SKILLS




TASTES AND HOBBIES



Appendix G

Photographic historian persona



Joseph Callow

Freelance Photographic Historian

Age: 59 years	Education: BA
Sex: Male	Occupation: Freelance Photographic Historian
Length of Employment: 11 years	Location of Employment: Self-employed

BIOGRAPHY

Graduated from university and began working in hospitality while developing his photography skills. Worked for a few years as a photographer, then got into historical photography. Began to work as a freelance photographic historian, helping non-profit groups and city councils on various projects.

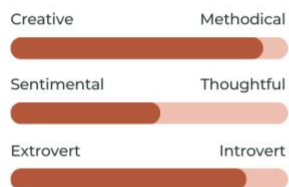
SEARCHING BEHAVIOURS

- Most likely to have already searched through online catalogues prior to walking in
- Has already identified sources of interest and set up appointment with librarians
- Will often spot small details in the pictures he looks at and goes off of those for his next search ("breadcrumbing")

WORK DUTIES

- Searching for photographs relevant to a prompt given by the specifications of the contract he's taken on
- Identifying minor details that would be of interest
- Writing short (250-500 words) pieces that will be used as online blog posts or captions based around a given prompt or theme

PERSONALITY



SKILLS



TASTES AND HOBBIES



Appendix H

Victorian historian persona



Dr Sara Childs
Historian

Age: 47 years	Education: BA Hons, MA, PhD
Sex: Female	Occupation: Historian
Length of Employment: 8 years	Location of Employment: Local Museum

BIOGRAPHY

Interested in Victorian history from a very young age and proceeded to obtain multiple academic degrees in the subject. Has worked at other museums prior to current job.

SEARCHING BEHAVIOURS

- Works very closely with historians at other institutions
- More likely to reach out to library staff prior to conducting a database/catalogue search
- Enjoys browsing in the same area as a given source
- Diligently follows citations in both primary and secondary sources

PERSONALITY



WORK DUTIES

- Researching the provenance of artifacts held by the museum
- Collaborating with other institutions on exhibitions to highlight particular artifacts and collections
- Managing public outreach and education programs
- Managing artifact curation, preservation, conservation, storage, and display

SKILLS



TASTES AND HOBBIES

