

Amy L. Johnson. Implications of American Copyright Law for Policies and Practices Regarding the Digitization of Manuscript Collections. A Master's Paper for the M.S. in L.S degree. April, 2009. 64 pages. Advisor: Deborah Barreau.

With the advent of widely available Internet access and improved digital scanning capabilities, many cultural institutions have begun putting digital copies of their archival holdings online for increased access by scholars and the public. Advances in technology have not been matched by advances in copyright law, which does not address digital issues and remains difficult to interpret for modern technologies. This study investigates how repositories determine the copyright status of manuscript materials, and how they use that information when digitizing collections. Using an Internet survey administered to members of the Society of American Archivists, this study explores the relationship between United States copyright law and the policies and practices of American archival repositories regarding the digitization of manuscript materials.

#### Headings:

Copyright / Computer-stored information

Copyright / Manuscripts

Archives / United States

Surveys -- Archives

IMPLICATIONS OF AMERICAN COPYRIGHT LAW FOR POLICIES AND  
PRACTICES REGARDING THE DIGITIZATION OF MANUSCRIPT  
COLLECTIONS

by  
Amy L. Johnson

A Master's paper submitted to the faculty  
of the School of Information and Library Science  
of the University of North Carolina at Chapel Hill  
in partial fulfillment of the requirements  
for the degree of Master of Science in  
Library Science.

Chapel Hill, North Carolina

April 2009

Approved by

---

Deborah Barreau

## TABLE OF CONTENTS

ABSTRACT .....	i
TITLE PAGE.....	ii
TABLE OF CONTENTS .....	iii
LIST OF FIGURES .....	v
Chapter 1 Introduction .....	1
Chapter 2 Review of the Literature .....	6
2.0 Introduction.....	6
2.1 Overview and Implications of Copyright Law .....	7
2.2 Related Studies .....	11
Chapter 3 Method .....	19
3.0 Introduction .....	19
3.1 Data Collection .....	19
3.2 Questionnaire .....	21
3.3 Data Analysis .....	22
3.4 Ethical Considerations .....	23
3.5 Limitations .....	24
Chapter 4 Results .....	27
4.0 Introduction .....	27
4.1 Context .....	27

4.2 <i>To what extent do copyright issues influence the selection of manuscript materials for digital publication?</i> .....	29
4.3 <i>How do archives allocate resources to obtaining copyright permissions, and how successful are these efforts?</i> .....	33
4.4 <i>What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?</i> .....	39
4.5 <i>What copyright information is provided to users of digitally published materials?</i> .....	41
Chapter 5 Discussion .....	44
5.0 Introduction .....	44
5.1 <i>To what extent do copyright issues influence the selection of manuscript materials for digital publication?</i> .....	44
5.2 <i>How do archives allocate resources to obtaining copyright permissions, and how successful are these efforts?</i> .....	46
5.3 <i>What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?</i> .....	47
5.4 <i>What copyright information is provided to users of digitally published materials?</i> .....	48
Chapter 6 Summary .....	49
REFERENCES .....	51
APPENDIX A: Survey Instrument .....	54

## LIST OF FIGURES

Figure 1: Types of Repositories .....	28
Figure 2: Percentage of time spent on digitization .....	29
Figure 3: Importance of factors for considering publication .....	30
Figure 4: Copyright status of materials chosen for digitization .....	32
Figure 5: Strategies for determining copyright has expired .....	33
Figure 6: Strategies for determining repository ownership.....	33
Figure 7: Factors that limit the identification of copyright owners.....	35
Figure 8: Individual responsible for ending search for copyright holder .....	36
Figure 9: Actions if copyright owner cannot be located.....	37
Figure 10: Copyright holder reaction to digitization .....	38
Figure 11: Copyright holder conditions for digitization.....	38
Figure 12: Parties responsible for developing copyright policies or guidelines.....	40
Figure 13: Parties responsible for approving copyright policies or guidelines .....	40
Figure 14: Strategies for limiting illegal use of materials .....	42
Figure 15: Copyright guidance provided to website users .....	43

## **INTRODUCTION**

Following the introduction of widely available Internet access, many libraries, archives, museums, and other cultural institutions have begun to digitize and make their holdings available online for the education and enlightenment of scholars and the public. While mass-digitization projects of published materials, such as the Google Book Search project, have received a great deal of positive and negative attention (for examples, Baksik, 2006; Thatcher, 2006; Grogg, 2008; and Lackie, 2008), these are not the only types of projects that are currently being undertaken. The digitization of previously unpublished manuscript material is making vast quantities of artworks, letters, diaries, scrapbooks, notes, and other valuable and unique historical documents accessible on a scale that was previously impossible. Formerly, access to these objects was restricted to those who had the financial wherewithal to physically visit a repository and spend the time tracking down relevant materials, primarily scholars.

Now, thanks to advances in digital technologies, not only have manuscript collections become increasingly available to members of the public who have nothing more than an interest and an Internet connection, but search engines have made it easier to find relevant sources using a few keywords. Copyright law has not kept pace with the technology boom, however, leaving repositories unclear as to the legal implications of digitizing historical materials and making them available to the public via the World Wide Web.

United States copyright law is difficult to understand, partly because it leaves many important concepts (such as “original,” “substantial,” and “reasonable”) poorly defined, but the sections that deal with unpublished materials and exceptions for cultural institutions are notably opaque, even in comparison to the indecipherable whole (LeFevre, 1992). The law is also silent or difficult to interpret with regard to digital copying and fair use of digitized materials, making repositories unwilling to take risks without knowing the potential consequences of their actions. The decisions of the courts in these matters have likewise been difficult for lay people to understand. It is therefore no surprise that so many publications have been generated that attempt to explain the implications of copyright law for archives, manuscript materials, and archival digitization projects (for example, see LeFevre, 1992; Spoo, 1999; Spoo, 2000; Minow, 2002; Besek, 2003). What is unexpected is the major shortage of empirical research to date examining how United States copyright law affects the decisions made by archival repositories with regard to their digital collections.

Based on the quantity of published material attempting to guide repositories in their digitization efforts, it is clear that many regard copyright as a major factor for consideration when undertaking a digital publication project. Despite this assumption, no research has yet been carried out to determine the extent to which this is true, or the ways in which copyright concerns impact digitization projects. How and when repositories choose to make their materials more available to the general public on the Internet affects what is available online. For instance, if repositories are only willing to publish digital copies of manuscripts that have lapsed from copyright, vast quantities of important

materials—from most of the 20<sup>th</sup> century and all of the 21<sup>st</sup>—will remain largely unavailable to the public for many years.

The present study will explore the effects of United States copyright law on the policies and practices of American archival repositories in digital publication for public consumption of their previously unpublished manuscript materials. Specifically, the intention is to answer the following research questions:

1. To what extent do copyright issues influence the selection of manuscript materials for digital publication?
2. How do archives allocate resources to obtain copyright permissions, and how successful are these efforts?
3. What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?
4. What copyright information is provided to users of digitally published materials?

First, question one addresses selection criteria for materials to be digitized and the extent to which copyright considerations play a role in the decision-making process. The focus of the second question is to determine when in the project process questions of copyright are addressed, how much time and effort are devoted to copyright issues (including determining copyright status of the materials, identifying and locating rights-holders, and obtaining copyright permissions) and whether the project will proceed if it is difficult to determine copyright status or copyright holders. Considerations of official repository policies are examined in the third question; special consideration will be given to how repositories' policies compare to their actual practices, as determined by the other three questions. Finally, question four focuses on the information provided to the public regarding the copyright status of digitized materials; specific attention will be paid how



repositories inform their users about fair uses of digitized materials and discourage them from using materials in prohibited or illegal ways.

These research questions are similar to the third research question posed by Dryden (2008), who investigated Canadian copyright law and Canadian repositories. Dryden's work is the only empirical study found thus far to explore the connections between copyright law and the ways in which repositories digitally publish their archival materials. Her research also explores the factors that influence decisions to make archival holdings digitally available, the connections between individuals' attitudes to copyright and their knowledge of copyright, and the connections between archivists' understanding of copyright and eventual institutional practices. Using Dryden's work as the methodological foundation, this study focuses on exploring repository policies and practices within the United States.

The purpose of this study is to explore the impact of American copyright law on the policies and practices of repositories making materials available online. This information will inform archivists, historians, and other information professionals of how their peers approach issues of copyright, including official policies and actual practices, as well as document the existence, if any, of divides between what is technologically possible and legally permitted with regard to the digitization of manuscript materials. As a result, a deeper understanding of the implications of copyright law on the policies and practices of digital publication in the United States will be valuable to many archivists and public historians, who will be able to use the best practices of other institutions to improve upon their own digitization projects. Better comprehension of copyright issues may encourage many institutions to begin or continue their digitization efforts, benefiting

countless historians, scholars, and interested members of the public by making historical material more easily accessible for their use.

The concerns inherent in navigating copyright and digital publication are not going to go away. As digital storage becomes less expensive and fast Internet connections become more common, more repositories will look to place archival content online, digitally publishing their materials to share with the world. This study is intended to shed light on current policy and practices in the United States, with the goal of better understanding ways that copyright law affects when and how valuable archival materials can be made public.

## **REVIEW OF THE LITERATURE**

### 2.0 Introduction

In a survey of professional historians, researchers found that U. S. historians use a wide range of primary resources and use an even wider variety of search strategies to locate them (Tibbo, 2003). The study concluded that “the business of the archival enterprise in the digital age” was to make web pages, electronic finding aids, and other materials part of the “daily archival product and perspective” (p. 29). A second survey of Canadian historians concluded that “digital reproductions have an as yet untapped potential,” with the implication that “the World Wide Web offers opportunities to expand and ‘democratize’ access to archival materials and to archival expertise” (Duff et al, 2004, p. 22). These studies indicate a trend toward online description of historical materials and access to the digital reproductions of these materials, which are becoming increasingly important within the ranks of up and coming professional historians.

The following review of the literature will situate the current study in the context of copyright law and the available knowledge about archives’ reactions to the law. The first section briefly outlines sources that explain various aspects of copyright law and their implications for archival policy and practice. These include library rights and exceptions, unpublished works, fair use, and expiration into the public domain. These types of sources form the majority of the professional literature on copyright, digitization, and archival policy. The subsequent section discusses the relevant empirical research and case studies which have implications for, or draw conclusions based on, the observed

policies and practices of archival repositories. The fact that similar studies have been executed in the United Kingdom and Canada is motivation for the current study. Surveying the professional literature leads to the conclusion that further research is necessary before an understanding of the effects of copyright law on the digitization of unpublished historical materials can be reached.

## 2.1 Overview and Implications of Copyright Law

United States copyright law was originally conceived by the framers of the constitution as a way of empowering Congress “to promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries” (U.S. Const. art. I, § 8, cl. 8). The law is described in title 17 of the United States Code (U.S. Copyright Office, 2007), and more than fifty amendments and acts have been appended since the Copyright Act of 1976 was passed.

Copyright protects the creators of original works so that they can benefit from their creations. The creators of manuscript materials own the copyright for those materials, and they have exclusive rights to reproduce their work and to distribute or display the copyright work to the public, unless they have transferred or waived copyright ownership. In case of the original author’s death, the rights pass to the author’s spouse, children, grandchildren, and executor. Transfer of these rights to repositories is becoming much more common, but this was not always the case in the past, which means that current copyright holders must often be located to request permission to digitize. Furthermore, complete provenance information may not be available for every collection,

making it difficult in many cases to either identify the original creator of the manuscripts or the collection donor. Without this information, it can be very difficult to locate the current copyright holders of those works.

Repositories wishing to digitize portions of their manuscript collections must consider the copyright status of their documents. Currently, duration of copyright is dependent upon many factors, the most important of which are the following (U.S. Copyright Office, 2007, § 302, 304):

- In general, copyright for works created on or after January 1, 1978 will extend to seventy years after the death of the author.
- In the case of anonymous, pseudonymous works, and works made for hire (“prepared by an employee within the scope of his or her employment”) created on or after January 1, 1978, copyright will endure for ninety-five years from the date of first publication, or one hundred twenty years from the date of creation, whichever comes first.
- For works in their first term of copyright (that is, copyright has not been renewed) as of January 1, 1978, copyright will endure for ninety-five years from the date the copyright was originally secured.

After copyright expires, works pass into the public domain, at which point they are freely available for use. If repositories were to abide by these rules, they would only be able to digitize and publish extremely old documents, excluding a majority of important, more recent materials. Minow (2002) has created a practical guide to the public domain that further explains when certain types of documents lapse out of copyright based on factors such as the year of the creator’s death and publication date. Repositories can digitize works that have been moved to the public domain without fear of repercussions.

However, sections 107 and 108 of Title 17 offer libraries and archives some justification for the digital copying of unpublished manuscript materials. Though they total fewer than four pages, the complexity involved in navigating these exceptions has

led many scholars and organizations to author explanatory guides to American copyright law with the aim of assisting libraries, archives, and other cultural institutions in navigating these often confusing documents. One of the most important parts of the law with regard to digitization projects is section §108, which deals with the exceptions in the Copyright Act for libraries and archives. The Library of Congress has a great deal of interest in §108, which has led to the publication of some of the best guides on the subject.<sup>1</sup> Besek's (2003) thorough assessment of copyright issues focuses specifically on digital archives, making it particularly pertinent for practitioners. The review of the history of 17 U.S.C. §108 (Rasenberger and Weston, 2005) is useful to digitization teams seeking to understand not only the current state of copyright law as it applies to libraries and archives, but also past applications of the law.

If a work is not in the public domain, it is still possible to use it if "fair use" can be determined. The rules regarding fair use are found in §107, but they are difficult to navigate. The section states that a determination of fair use will be made on a case-by-case basis, determined by four factors.<sup>2</sup> A summary of fair use case law as it applies to unpublished texts follows a hypothetical biographer, giving examples of the difficulties encountered when trying to determine fair use (LeFevre, 1992). Allison-Bunnell (1995)

---

<sup>1</sup> The Section 108 Study Group was convened by the Library of Congress and the U.S. Copyright Office to make recommendations for changes to U.S. Copyright Law, with the intent of balancing the rights of copyright owners and the needs of libraries and archives. They made their final report in May, 2008; their recommendations have not yet been legalized.

<sup>2</sup> (1) the purpose and character of the use, including whether such use is of a commercial nature or is for nonprofit educational purposes; (2) the nature of the copyright work; (3) the amount and substantiality of the portion used in relation to the copyright work as a whole; and (4) the effect of the use upon the potential market for or value of the copyright work. The fact that a work is unpublished shall not itself bar a finding of fair use if such finding is made upon consideration of all the above factors.

examines fair use law from the perspective of an actual archival collection, concluding that “it is impossible for a manuscripts curator to act precisely within copyright law and still serve the needs of the research community.” These papers date to a pre-digitization era, so they must be interpreted with caution using conscious decision-making. Prudence is key for archivists and researchers when weighing the risks and rewards of copying materials.

Terms that are vital to the understanding of copyright law as it applies to digitizing manuscript materials—for example, “unpublished materials”, “orphaned works”, and “reasonable effort”—are often poorly defined in copyright law, and only occasionally touched upon in the professional literature. Copyright law is not always clear about how to manage unpublished materials as compared to published materials. Orphaned works are those whose authors cannot be identified or located. Many libraries and archives have determined that a reasonable effort to find rights holders before determining orphaned status is sufficient to fulfill their obligations under copyright law, but the amount of time, effort, or cost have not been legally defined. Hirtle (2001) examines copyright law, unpublished materials, and orphaned works with regard to how changing technologies have changed the interpretation of the law, concluding that clear legislation on unpublished materials and orphaned works is necessary. He further suggests that a standard for reasonable investigative effort be defined, which could establish a fair use of material.

Aside from published articles, numerous guides and handbooks have been published in the last decade with the goal of informing the digitization efforts of interested cultural institutions. Many of these guides include significant portions that

attempt to explain the nuances of copyright, which they describe as a considerable barrier to digitization projects. The sizeable effort necessary to determine copyright status of materials and locate copyright holders of those materials can be daunting, expensive, time consuming, and fraught with legal hurdles.

## 2.2 Related Studies

In general, past literature suggests a widely held conviction that copyright considerations play a significant role in the digitization of archival materials in American archival repositories. This assumption is supported by Astle and Muir (2002), who address selection criteria for digitization. In a survey of twenty public libraries and archives in the United Kingdom, the study found that respondents ranked copyright as being an important selection criterion, though not the most important one. Preservation issues, access, and public demand were all considered to be more important than a lack of copyright complications. This result was not anticipated by the authors, given “the high profile of copyright issues amongst librarians in general and the emphasis placed on the avoidance of these issues in the literature” (Astle and Muir, 2002, p. 73). However, without good definitions of these selection criteria, which were not provided in the report, it is impossible to know the implications of these findings. The researchers also stated that the range of materials that has been digitized is limited by copyright restrictions, but it was unclear what parts of their data supported this assertion. This study indicates that, while copyright may not be the most important factor when making digitization decisions, it is important enough to justify further exploration.



In a similar survey, Bültmann, Muir, and Wictor (2006) also found that, while copyright does apparently play a role in libraries' and archives' selection criteria, it is secondary to a number of other criteria, including patron demand and rarity of the material. Their survey of digitized content in forty-seven U. K. research libraries and archives—conducted via desk research, Web-based questionnaire, and interviews—found that over a third of survey respondents restricted access to their digitized collections. Copyright or other intellectual property rights restrictions were listed as “the most frequent hindrance to public access” (Bültmann, et al., 2006, 110). In general, this study confirms the conviction, which is a common consensus in existing literature, that copyright is a concern for repositories attempting to begin or continue similar projects.

Liew (2006) examined a range of issues faced by online cultural heritage exhibitions, and found that the management of intellectual property issues was the most important concern among historical associations and foundations, but did not rank in the top five factors for libraries, archives, or museums. However, there were very few representative institutions for libraries (6), archives (3), government agencies (4); the majority was museums (24). This study, together with Astle and Muir (2002) and Bültmann, Muir, and Wictor (2006), motivates the continued analysis of the role and effects of copyright issues and digitization. Copyright clearly plays a role in the digitization decision-making process, but how large a role and what its effects are have yet to be determined.

Some worthwhile research has been conducted that does not relate directly to archival materials and copyright but is nonetheless relevant to those issues. George (2005) examines the difficulty of acquiring copyright permission to digitize published

works. The study took a random sample of 273 titles from the Carnegie Mellon University Library's collection and attempted to obtain permission to digitize the works. This necessitated determining who owned current copyright of those works not in the public domain, and where those parties could be located. Of the 273 requests for permission to digitize that were made, 24% of the total agreed to allow digitization, and 9% of the total requests made placed no restrictions upon that digitization, such as restricted access or a request for payment. This shows the types of problems that libraries might face in acquiring permission to digitize materials, even when it is relatively easy to identify copyright holders. Copyright decisions can be much more difficult when they are based on unpublished manuscript materials, because authors are often anonymous, impossible to identify, or simply difficult to contact.

Beyond the surveys noted above, which focus primarily on the repositories' digital publication policies, a number of studies have undertaken content analyses or case studies of copyright practice in archival collections. In evaluating Web sites featuring U.S. history primary sources, Congleton (2005) verified that copyright information is provided to users who access collections, and is considered an important criterion for evaluating those collections. In a review of twelve Web sites, Congleton determined that copyright information was available on ten. Breaden (2006) analyzed twenty-five online audio exhibits from the United States, Canada, and the United Kingdom for audio quality and accessibility. Audio materials, although not as commonly digitized as print media, are often unpublished and have similar copyright restrictions as other unpublished materials. The author indicates that unclear copyright laws might be one reason that the management of online audio collections can be improved. This study offers evidence that

difficulties surrounding copyright law can cause problems for archival projects regardless of material type.

Two case studies offer in-depth examples of how repositories currently respond to the pressures of copyright law when creating digital projects. The three digitization projects covered in the two articles dealt with a variety of unpublished historical materials, mostly from the late 19<sup>th</sup> and early 20<sup>th</sup> centuries, and represented many different copyright challenges. Despite major investments of time and resources, the project teams concluded that pursuit of copyright holders was a worthwhile element of their respective digitization projects.

Duke University's Ad\*Access project began in 1997 at the very beginning of the digitization movement; its intent was to create an online collection of over 7,000 advertisements published in the United States between 1911 and 1955 (Pritcher, 2000). Many of the ads were not in the public domain, compelling the project team to make important decisions regarding the pursuit of copyright permissions. Although efforts were made to locate copyright owners prior to making digitizing content publicly available, locating copyright holders was a major difficulty due to mergers, acquisitions, buy-outs, and bankruptcies caused company names to change or disappear. Even when current rights holders were identified using historical documentation and Internet sources, the owners were often reluctant to grant digitization permissions due to the lack of information about long-past acquisitions. However, Pritcher notes that the relatively few successes that the Ad\*Access team accomplished indicate that asking for digitization permission was the most expedient technique for acquiring copyrights.

The United Kingdom's Arts and Humanities Data Service, a now-defunct national service to promote the use of digital resources, sponsored the second case study; it examined how two British digitization projects traced copyright holders (Dunning, 2004). Both projects concluded that tracing copyright holders and obtaining permission to digitize materials were the most taxing parts of the projects with respect to digital rights. The Hantsphere project, organized by Hampshire County Council, was an online resource about the history and heritage of the County of Hampshire. The second project, called the Coalfield Web Materials project and sponsored by the University of Swansea, undertook to digitize materials related to South Wales coal-mining. Many of the materials in both collections were difficult to trace due to lack of provenance information. To identify copyright holders, the projects consulted many local information sources, such as historical societies and elderly individuals, with moderate successes. Both projects found that the relationships formed with these sources during the search for copyright holders, many of whom were eager to share their stories, led to gathering more material for digitization and helped inform important stakeholders of the projects' goals.

These case studies indicate that tracing copyright holders of historical materials can be a valuable and worthwhile part of a digitization project, despite the associated costs. All three project teams concluded that the goodwill engendered by making contact with potential rightsholders was worth the time, effort, and resources expended in pursuing permission to digitize. Copyright clearance is a major organizational and administrative issue, but when managed successfully can add a great deal of value to the finished product.

All of these studies inform research undertaken by Dryden (2008), in which the author investigates Canadian copyright law, and the digitization policies and practices of Canadian archivists and repositories using a survey, content analyses of web sites, and interviews. Dryden's dissertation is the basis for the current study, as it is the only cross-repository study thus far to explore the connections between copyright law and the ways in which repositories digitally publish their archival materials. Taking a broad view of copyright implications, Dryden explored the factors that influence decisions about copyright in Canadian repositories, providing data that can provide a benchmark for comparing the findings of the current study on American practices and policies.

The survey, which was mailed in the form of a questionnaire, was completed by 106 respondents. These respondents were generally well-educated and had many years of experience in archives; they worked for a broad range of archival repositories, university collections, historical societies, libraries, museums, and corporate environments. Additionally, Dryden interviewed twenty-two archivists to get a deeper understanding of the survey results. The purpose of this exploratory research was to explore the impact of Canadian copyright law on the practices of repositories making materials available online. The present study draws largely from Dryden's third research question: "What are the copyright practices of Canadian archival repositories in making their archival holdings available on the Internet?" (Dryden, 2008, 5) This question was elaborated in four subquestions.

A total of 80 percent of respondents thought that copyright was an issue for repositories attempting to digitize materials. Data from surveys and interviews suggested that repositories selected items for digitization that were at little risk of infringing upon

copyright or that required less effort and resources in order to obtain copyright permissions or determine copyright status. On the other hand, the existence of copyright policy statements outlining attempts to find copyright owners, and the fact that risk assessment was a part of many repositories' selection strategies suggests that some institutions are willing to take risks when making materials available online.

Despite the fact that 79 percent of respondents thought educating the public about copyright was a part of their professional duty, only two-thirds of analyzed websites included a general copyright statement, and only one-third provided any kind of copyright information at the item level. Dryden posits that this dearth of information was due to two factors. First, it is difficult to be certain of the copyright status of any one item, given, for example, a lack of documentation about provenance and a lack of resources to research each item individually. Secondly, a lack of copyright information might impel users to contact the repository for further information, making the terms and conditions of use easier to disseminate.

Dryden concluded that the "best" resources were not being digitized for multiple reasons. Often repositories did not have the resources to dedicate to determining copyright status and ownership, and many archival institutions viewed copyright restrictions very conservatively and were very careful not to digitize risky holdings. The author decided that, in Canada, copyright law is problematic, due to the difficulty of understanding and interpreting it.

Dryden points out several important issues that exist in the process of digitizing published and unpublished materials in Canadian repositories. Further research must be conducted to determine the effect of United States copyright law upon archival policy

and practice. There are numerous factors that cause the relationship between the two to shift over time, and so no single study can describe anything but a snapshot of the effects at particular moment in time. However, the combined impact of these varied studies suggests that considerable further research is necessary to better understand the consequences that observing copyright law might have upon how, when, and why unpublished historical materials are digitized in the United States.

**METHOD:****3.0 Introduction**

The study gathered data about archival repositories' copyright policies and practices using online survey software. Participants were self-selected from subscribers to Society of American Archivists listservs. Both quantitative and qualitative data were collected using an on-line questionnaire and analyzed using SAS statistical analysis software and building on work reported by Dryden.

**3.1 Data Collection**

Determining the correct study population proved problematic. The unit of analysis for this study was the archival repository, and under ideal conditions, the study population would have consisted of all those American archival repositories (which might include libraries, corporations, museums, and historical societies as well as traditional archives) that had digitized or were in the process of digitizing manuscript materials for publication on the Internet. The sampling frame would have been an ideal list of all these organizations. One example of such a repository was the Carolina Digital Library and Archive, located at the University of North Carolina at Chapel Hill, which is responsible for the university's digital materials. Their online collections--Documenting the American South--are an excellent example of the kinds of projects targeted by this study.



In practice, however, the population was made up of employees of archival repositories and the sampling frame consisted of the membership lists of three Society of American Archivists e-mail listservs. While it might have been more effective to sample archival repositories with digital collections or projects from a comprehensive list or database, no such database existed, and compiling such a list was out of the scope of this study. It was therefore necessary to use individual employees of archives that are digitizing materials as a synecdoche to understand the policies and practices of their organizations. This study used a convenience sample of these persons to represent the larger population of repositories interested in digital copyright issues.

The Society of American Archivists is an American professional organization with over 5,000 individual and institutional members. Sections are self-selecting groups of members who have expressed interest in a particular topic or specialty. The Manuscript Repositories Section, with over 900 members, “is composed of SAA members working in archives, libraries, and other repositories that acquire and administer manuscripts created outside the collection institution” (Manuscript Repositories Section). SAA Roundtables are available to members and interested non-members. The Metadata and Digital Object Roundtable claims over 450 individual members and provides space for “collaboration and a source for guidance to archivists at all types of repositories as they engage with the digital environment” (Metadata and Digital Object Roundtable).

Originally, the intent of this study was to survey exclusively members of these two groups, which were chosen for their relevance to the research objectives. Rather than use only one of the groups, both were selected to reach the broadest number of respondents possible. However, two weeks after the survey opened, it was determined

that not enough responses had been made, and the decision was made to open the survey to members of the Society of American Archivists' Archives and Archivists (A&A) List, an open forum for SAA members and interested non-members. This decision resulted in a higher number of responses and a more generalizable sample. There was no way of gauging the number of members actively involved in digitizing manuscript materials, and so it was impossible to determine the total number of possible participants or what might constitute a satisfactory response rate.

Participants were recruited by sending an invitational e-mail to the chosen listservs; the e-mail contained a link to the Qualtrics survey instrument (Appendix A). In the invitation, the researcher requested that only those with current experience working on digitization projects participate in the study, which limited observations to those with some knowledge of how their repository responds to copyright issues. The survey itself consisted of twenty-seven multiple-choice, five fill-in, and six open-ended questions; most participants completed the questionnaire in 10 and 15 minutes. The only cost borne by subjects was the short amount of time necessary to complete an Internet survey, and no inducements for participation were offered. The survey was open from January 23 to February 23, 2009, a period of one month.

### 3.2 Questionnaire

Surveys are particularly appropriate for exploratory research, and while they are best suited for studies that take individuals as their units of analysis, it is also possible to study organizations when individual representatives serve as respondents (Babbie, 2007). This study also fits guidelines for when to use an Internet survey: the sample size is

relatively large, the survey can be conducted with a convenience sample, and there is an e-mail list available for the target population (Schonlau, Fricker, and Elliott, 2002).

A preliminary questionnaire (Appendix A) was reviewed by archivists and digital projects librarians involved with the Carolina Digital Library and Archives at the University of North Carolina-Chapel Hill and their suggestions led to several changes in the language and order of the questions. The final questionnaire consisted of twenty-seven multiple-choice, five fill-in, and six open-ended questions; additionally, many multiple-choice questions included the option to write in a response. An introductory section consisted of demographic questions, followed by four sections informed by the proposed research questions:

1. To what extent do copyright issues influence the selection of manuscript materials for digital publication?
2. How do archives allocate resources to obtain copyright permissions, and how successful are these efforts?
3. What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?
4. What copyright information is provided to users of digitally published materials?

Although the foundation of the survey is based on Dryden (2008), Bültmann, Muir, and Wictor (2006), Mugridge (2006), and Astle and Muir (2002), revisions were made to accommodate the specific purposes of this study.

### 3.3 Data Analysis

Data analysis commenced upon closure of the survey. Since both closed- and open-ended question formats are included, both quantitative and qualitative analyses are

used. Quantitative data was compiled from multiple-choice questions and analyzed for trends. Questions with an “other” option were coded as belonging to one of the other options or as a new category of response. Open-ended questions were analyzed using an open coding method, as described in Babbie (2007, p. 385).

### 3.4 Ethical Considerations

This study and all related promotional materials were reviewed and approved by the University of North Carolina at Chapel Hill Institutional Review Board (IRB) in January, 2009, and given the IRB number 08-2136; this number was attached to all participant communications in accordance with IRB procedures.

The study also adhered to the guidelines set out in the Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research (National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, 1979), which includes respect for persons, beneficence, and justice. The sample selection method and consent form ensured that participants were volunteers and that they were provided with enough information to make an informed decision about their participation.

Anonymity is preserved to encourage accurate responses. Knowledge and practice of intellectual property laws may vary considerably. The purpose of this research is to assess the importance of copyright in decisions to digitize and not to assess compliance with the law.

### 3.5 Limitations

As a survey, there were three major limitations to the study; these stem from the nature of surveys themselves, the sample frame, and the sampling method. The fact that the survey was administered via the Internet was not considered a problem; it was assumed that information professionals working with digital projects use computers frequently and have the technological wherewithal and understanding to participate in an online survey.

While surveys tend to be reliable—results can be replicated consistently—they are of questionable validity. Surveys are excellent at capturing results for large populations, but are not as useful for capturing nuance. Any survey of repository representatives would not be able to capture the full depth and breadth of issues regarding copyright. However, for privacy reasons as well as reasons of scope, it was necessary to administer a survey rather than identify and interview individuals; the privacy inherent in the survey method also provided anonymity, which was important when asking questions about knowledge and use of copyright restrictions. This method also allowed for a larger amount of data to be gathered, which produced more robust results than individual case studies would have done.

The sampling frame was questionable with regard to the population being studied. The population of this study is archival repositories: museums, corporations, libraries, historical societies, and archives. However, there was no extant list of repositories currently involved in digitization projects, and the creation of such a list by the researcher would almost certainly have had bias and taken more time than feasible for a study of this nature. Therefore, it was necessary to use individuals as representatives of the units of

analyses--archival repositories--raising further questions for the validity of the research. The Society of American Archivists listservs acted as the representative population from which the survey sample was drawn. The population of the listservs and the population of archivists working with digital materials were almost certainly not the same, since some members of the e-mail lists have no interest in digital projects. It is difficult to determine an accurate response rate since it was impossible to know how many members of these lists fit the criteria for the study. Also, those who were enrolled members of SAA listservs were probably relatively well-connected and involved in the professional community. Archivists with no interest in networking with their peers would probably not be members of SAA sections or roundtables. Despite these problems, sampling these representative populations was considerably more straightforward than identifying and then sampling the hundreds or thousands of archival repositories in the United States.

Finally, the sampling method raised concerns. The sample was self-selected based on criteria provided by the researcher, so it had to be assumed that respondents had the knowledge and experience that they claimed. Babbie wrote that “all elements must have equal representation in the frame. Typically, each element should appear only once” (2007, p. 201). There may be only one representative of a repository in a listserv, or there may be many, and there is simply no way of knowing how many responses come from the staff of the same repository.

Despite these limitations, most of which raise questions about validity of the study, a survey method was preferable to other potential study methods, primarily for privacy reasons. Field research, content analysis, or interviews would have required the identification of subjects for study, which would endanger the anonymity of participants.

For these reasons, it was decided that a survey was the best method for collecting this potentially sensitive information.

## **RESULTS**

### 4.0 Introduction

The survey questionnaire originally registered seventy-nine individual responses, but some of the responses could not be used in the analysis. Inspecting respondents' Internet protocol (IP) addresses led to the conclusion that some individual respondents had visited the survey site on multiple occasions, but completed the survey once. One individual visited the survey four times before responding, and another visited twice, but did not respond on either occasion. Blank duplicate responses were discarded for the purposes of data analyses. Also discarded were responses that did not contain any information beyond contextual data. A handful of respondents filled out several contextual questions, but did not go on to complete the remainder of the survey. Responses that were incomplete but had answered some relevant questions were used in the data analysis.

After subtracting the invalid responses, there were thirty-eight usable observations. The following results and discussion are based on these data.

### 4.1 Context

As shown in Figure 1, the respondents to this survey represent a wide range of types of repositories.



<b>Types</b>	<b>Freq.</b>	<b>%</b>
University archives	12	31.6
Library	6	15.8
Manuscript repository	6	15.8
National or state archives	4	10.5
Museum	3	7.9
Historical society	2	5.3
Non-profit archives	2	5.3
Municipal, county, or regional archives	1	2.6
Science research archives	1	2.6
Cultural heritage archives	1	2.6

Figure 1: Types of Repositories (n=38)

The majority of repositories engage in digitizing materials (95%), distributing materials within the repository (79%), distributing materials outside of the repository (87%), and distributing materials to the public (95%). The two respondents who wrote that their repositories were not digitizing indicated later that they had undertaken digitization projects in the past but were not actively digitizing additional material.

Survey participants classified themselves as having a wide range of job titles, but were mostly divided into discrete groups. 68% of job titles included the keywords “archives” or “archivists”, and 19% of respondents considered themselves to be librarians of some sort; three respondents (8%) used both terms when describing their job titles. Nine respondents (24%) were in managerial positions, as inferred from keywords such as “senior”, “director”, “chief”, “lead”, and “manager”. Ambiguous job titles like “college archivist” or “university librarian”, which could have indicated head of the archives or library, or simply an archivist or librarian working at a university, were not included in the managerial category. Two respondents described themselves as coordinators and one was a “PRA”, which was taken to probably mean a “public records administrator”. Ten of the respondents’ titles (27%) included the word “digital”.

The percentage of respondent's work time spent on digitization varied widely. Of respondents whose job title included the word "digital", 67% of time was spent on digital projects. Of respondents whose job title did not include words like "digital" or "digitization", 24% of time work time was spent on digital projects. The average for all respondents was 35% of time.

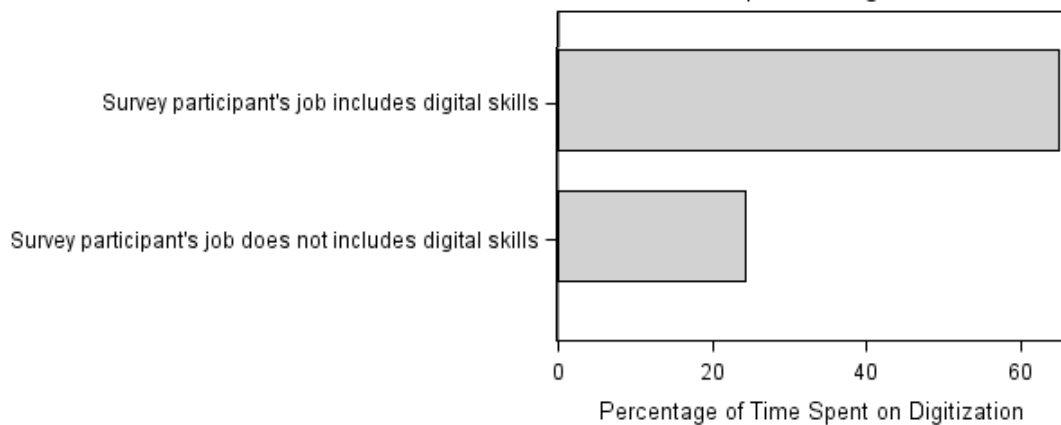


Figure 2: Percentage of time spent on digitization (n=36)

#### 4.2 *To what extent do copyright issues influence the selection of manuscript materials for digital publication?*

The purpose of these six questions was to determine the criteria for selecting materials for digitization in U.S. repositories. Most importantly, this research question addressed the extent to which considerations of copyright issues play a role in the decision-making process, when compared to other factors.

As shown in Figure 3, participants had varying responses to what they believed to be the most important factors for considering Internet publication of materials. While copyright issues were considered "extremely important" by a slim majority—nineteen (52%) out of thirty-six responses—other factors were also considered important. When

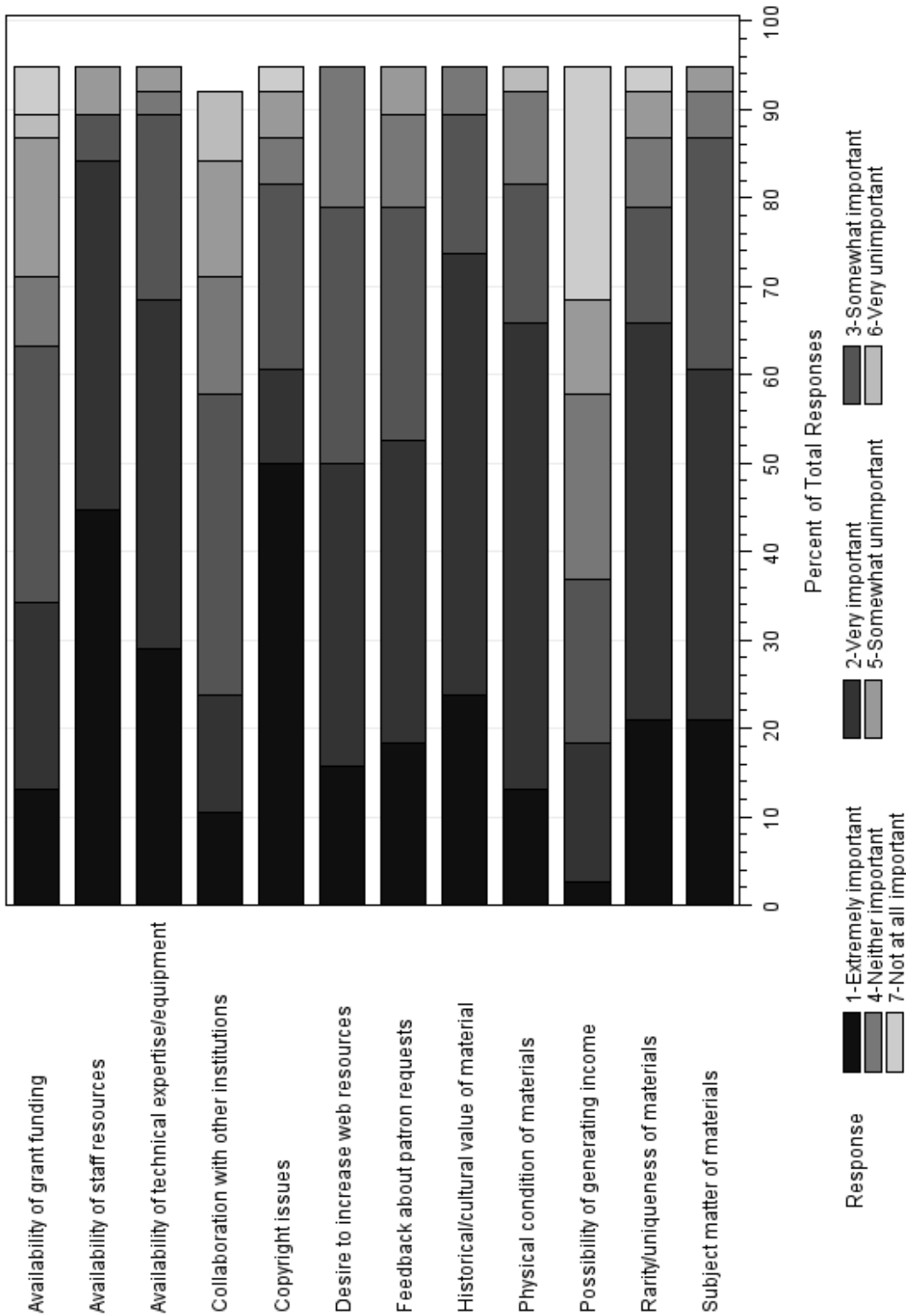


Figure 3 : Importance of factors for considering publication (n=37)

the first two ranking categories (extremely and very important) were added together, staff resources were clearly considered the most important factor, with thirty-two (89%) of thirty-six respondents believing that factor to be extremely or very important.

Respondents also had the option to write in factors and rate their importance. Two respondents ranked institutional mission as extremely important when considering materials to digitize. “User selection and funding” and “anticipated demand” were each written in once and also considered extremely important, while “ease of digitization”, “ownership and/or theft” and “team work” were rated very important by the respondents who wrote in those responses.

Most repositories (34 out of 37 responses, 92%, as one survey participant did not answer this question) considered the copyright status of materials before beginning a digitization project. Two respondents indicated that copyright was not taken into consideration, while one respondent was unsure. These responses are verified by respondents’ answers about what materials are selected for digitization. Repositories appear to select materials for which copyright has expired or is owned by the repository more often than they select materials for which copyright is owned by an outside party. 81% of respondents select materials for which copyright is expired, and 92% select materials owned by the repository, but 65% of respondents will select for digitization some materials for which copyright is owned by a third party. This is still a majority of repositories.

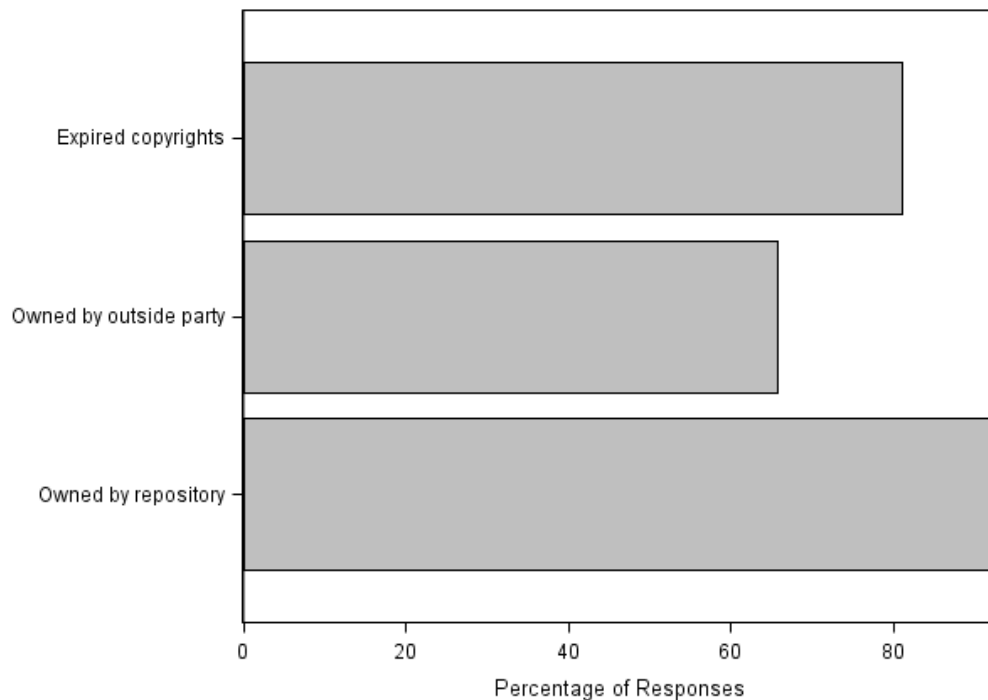


Figure 4: Copyright status of materials chosen for digitization (n=38)

Of the thirteen repositories that do not digitally publish materials when the copyright is owned by a third party, nine (69%) will digitize materials for use within the repository in cases when the repository wishes to make materials available to certain individuals in digitized form or save them to publish at a later date.

92% of repositories consider copyright in some before beginning a digitization project. Repositories use many different strategies to determine whether copyright has expired for any particular document. It was clear from many responses that respondents were considering both published and unpublished materials in describing their practices. For example, when asked how they determined that copyright had expired, thirteen out of thirty-four respondents said that they considered publication date. Respondents' strategies for determining copyright status are illustrated in Figure 5.

<b>Strategy</b>	<b>Freq.</b>	<b>%</b>
Refer to death date of document's creator	7	20.6
Refer to accession file or deed of gift	7	20.6
Refer to published copyright guides	4	11.8
Refer to copyright law	3	8.8
Utilize staff review or research	3	8.8
Consult with experts	2	5.9

Figure 5: Strategies for determining copyright has expired (n=34)

The most popular published copyright guide was Peter Hirtle's copyright and public domain chart, which was referred to four times (Hirtle 2009).

Repositories' strategies for determining that copyright is owned by the repository are illustrated in Figure 6.

<b>Strategy</b>	<b>Freq.</b>	<b>%</b>
Works for hire or generated by repository	16	45.7
Refer to deeds of gift or legal transfers of copyright	19	54.3

Figure 6: Strategies for determining repository ownership (n=35)

Some, but not all, deeds of gift appear to include a clause specifically transferring copyright to the repository.

#### *4.3 How do archives allocate resources to obtaining copyright permissions, and how successful are these efforts?*

While twenty-five respondents (66%) indicated that their repository selects materials to digitally publish when the copyright belongs to a third party, only nineteen of the respondents (50% of all respondents, and 75% of those who digitize third-party-owned documents), make an attempt to obtain authorization to publish materials from the copyright owners.

As indicated in Figure 7, among those attempting to get permission to digitize, repositories tend to end their attempts when they cannot identify or locate the copyright holder or the identified copyright holder does not respond to requests from the repository. Lack of time or financial resources appears not to be as important a factor.

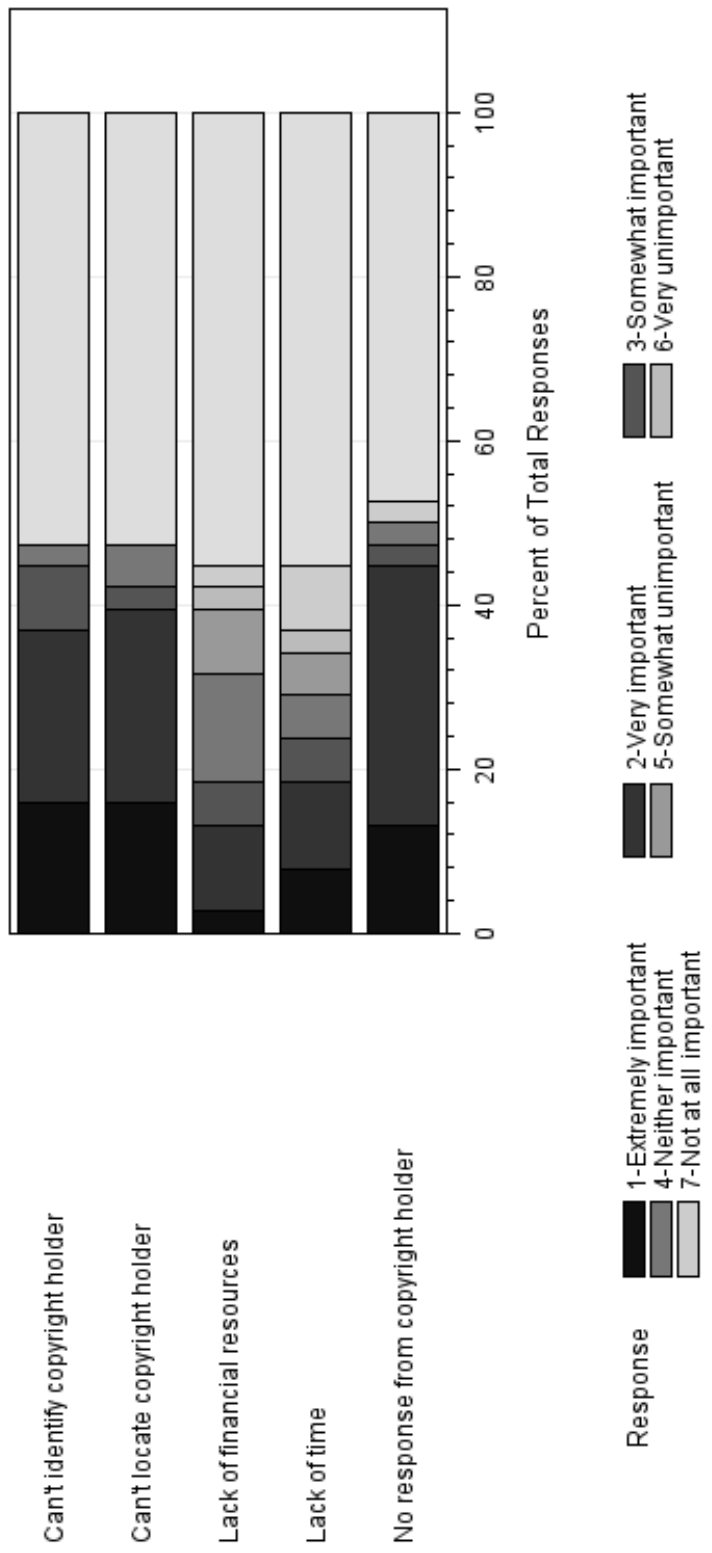


Figure 7: Factors that limit the identification of copyright owners (n=18)



As seen in Figure 8, a variety of individuals are responsible for deciding to end the search for a copyright holder.

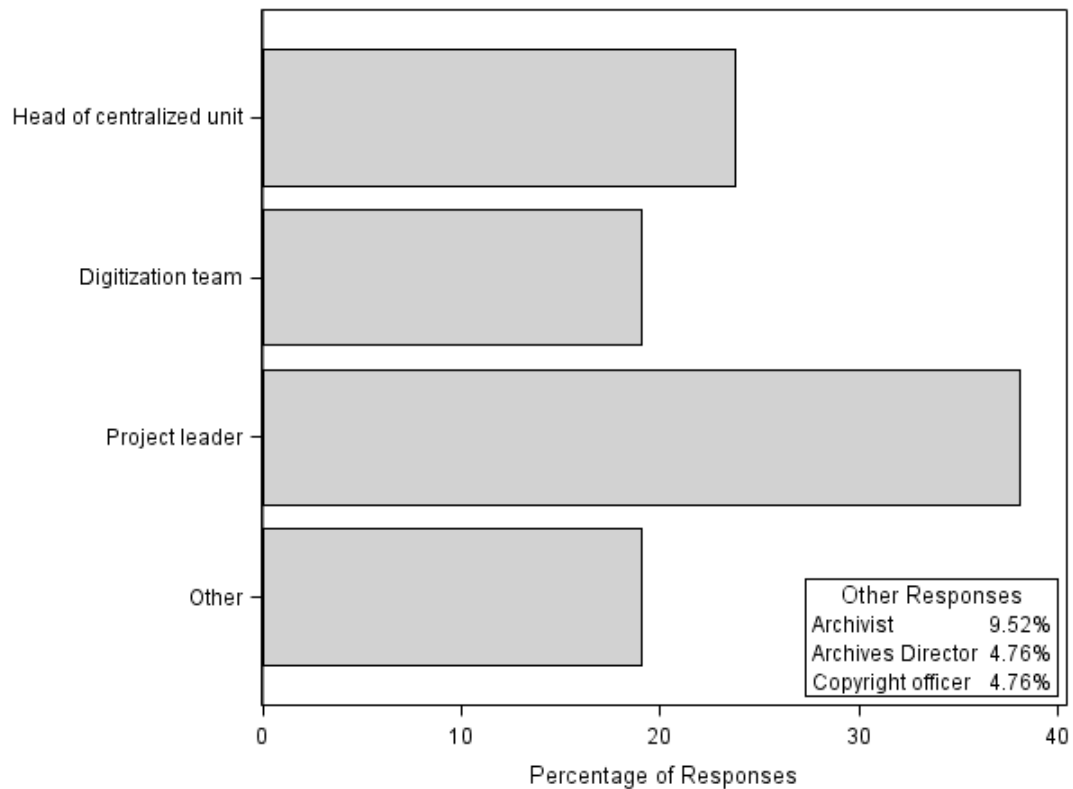


Figure 8: Individual responsible for ending search for copyright holder (n=18)

When the copyright holder cannot be identified, ten of nineteen respondents (52%) whose repositories attempt to identify and locate copyright holders use the document with a disclaimer. Two respondents indicated more than one potential response to this problem: one said that the repository would not use the document or would substitute another document; a second respondent said that the document would not be used, or the document would be used with or without a disclaimer. This makes it unclear what the repository would actually do in this situation, or if the response is dependent on other criteria. Among all respondents, three indicated that their only response would be to

not use the document. Other respondents had other strategies for dealing with non-respondent copyright holders, which are illustrated in Figure 9.

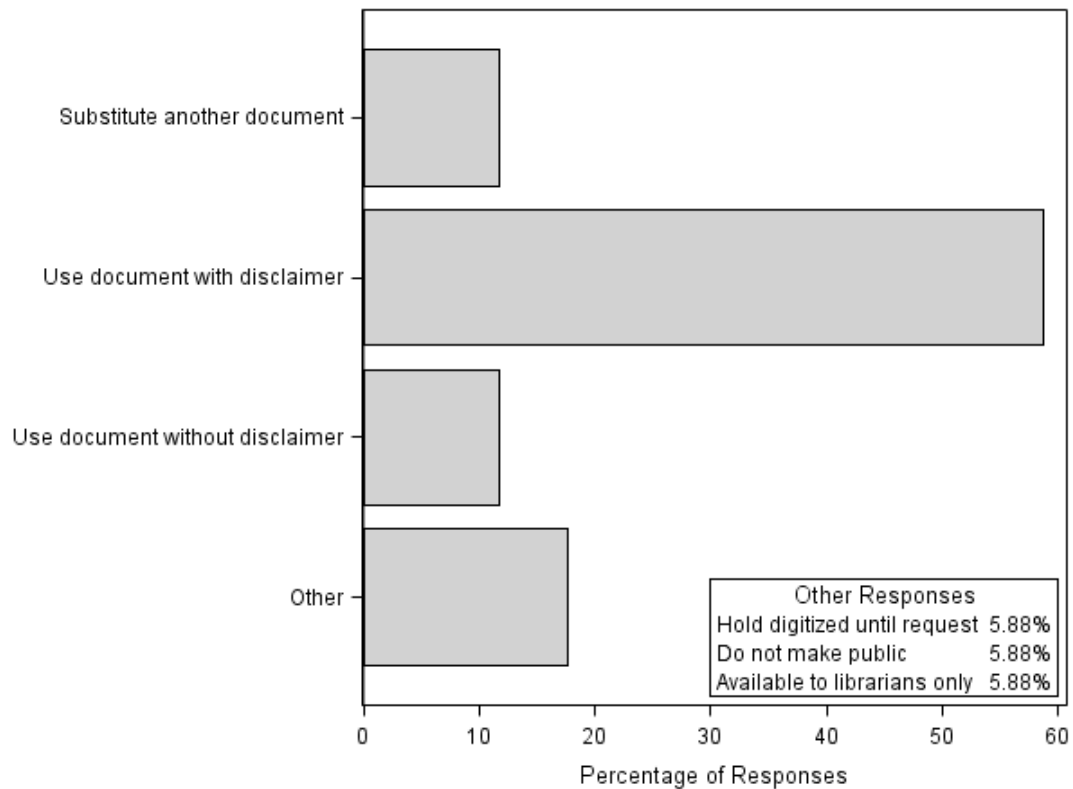


Figure 9: Actions if copyright owner cannot be located (n=19)

Copyright holders appear to be willing to allow their materials to be digitized, usually without attaching any conditions to the digitization, see Figure 10. On one occasion, a copyright owner gave an “ambiguous response” to a request for permission, but in a majority of cases repositories were able to get permission to digitize, even if there was a condition. Few copyright owners declined permission out of hand.

Occasionally, copyright holders will attach conditions to the digitization of their materials; typical conditions are outlined in Figure 11. Usually this only entails the use of a credit line to identify the copyright holder, but other owners would prefer that one-time

payments or royalties be paid to them, and occasionally copyright owners would prefer that access to those materials be limited to certain parties (other category).

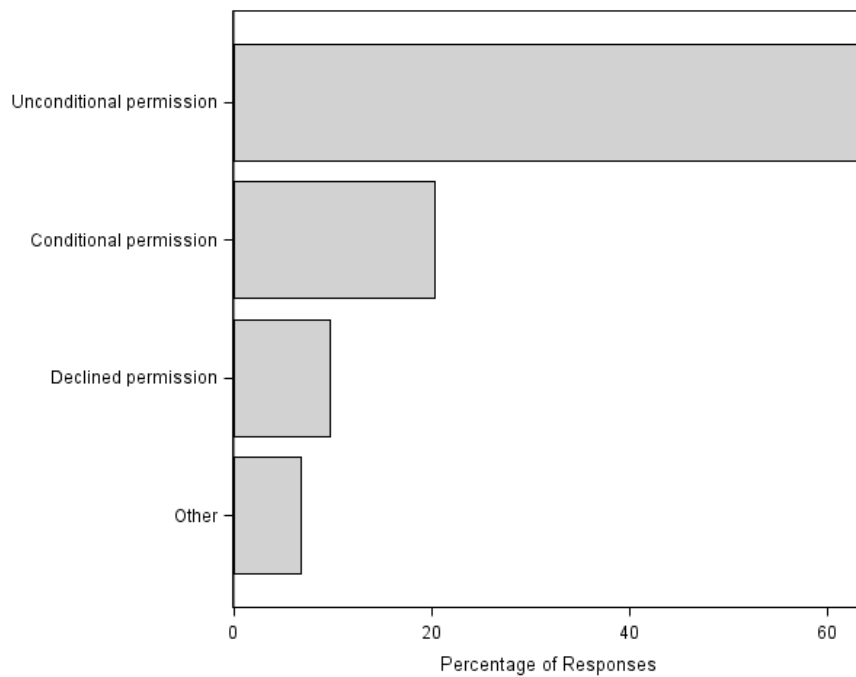


Figure 10: Copyright holder reaction to digitization (n=16)

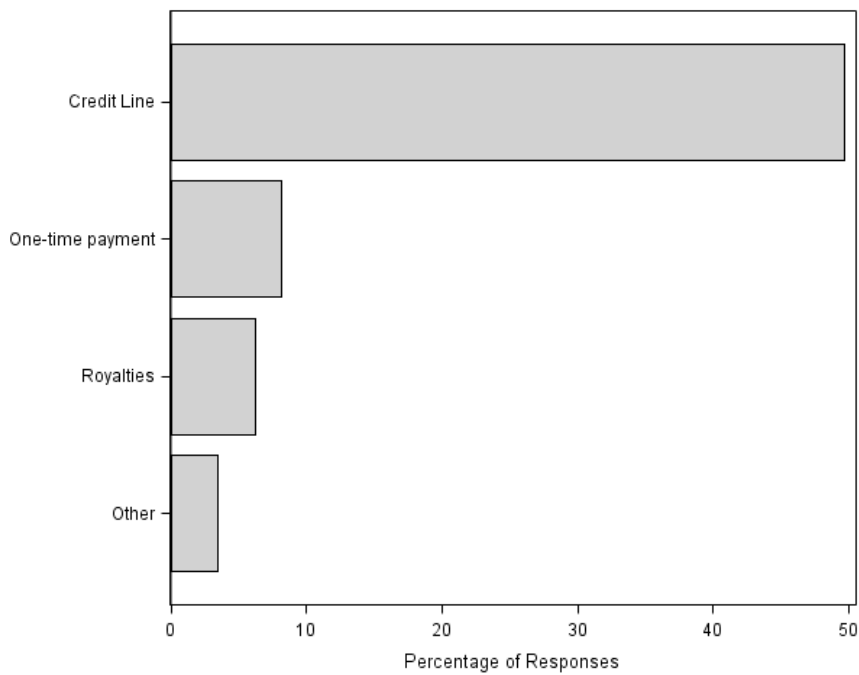


Figure 11: Copyright holder conditions for digitization (n=16)

Five respondents (21% of those who answered the question) had been challenged by a copyright holder because of digitized materials published without permission; three of these made no attempt to identify copyright holders before digitizing and subsequently publishing these materials. Occasions upon which repositories were challenged ranged from a secret fraternal organization that objected to publication of a fraternity symbol in a digitized yearbook, to copyright owners of these electronically published without permission, to complaints made by the Visual Artists and Galleries Association (VAGA). Only one of the respondents indicated in open-ended responses that they were frequently challenged by VAGA; their response was “to analyze the image and either ignore the complaint or remove the image, or only make the thumbnail copy available [sic].” Other repositories removed images of disputed materials and in some cases replaced the image with one not in dispute.

#### *4.4 What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?*

43% of respondents indicated that their repository has official policies or guidelines with regard to copyright. As shown in Figure 12, these policies, when they exist, have been developed by a wide variety of stakeholders. Policies have been approved at a variety of different levels as well, as shown in Figure 13.

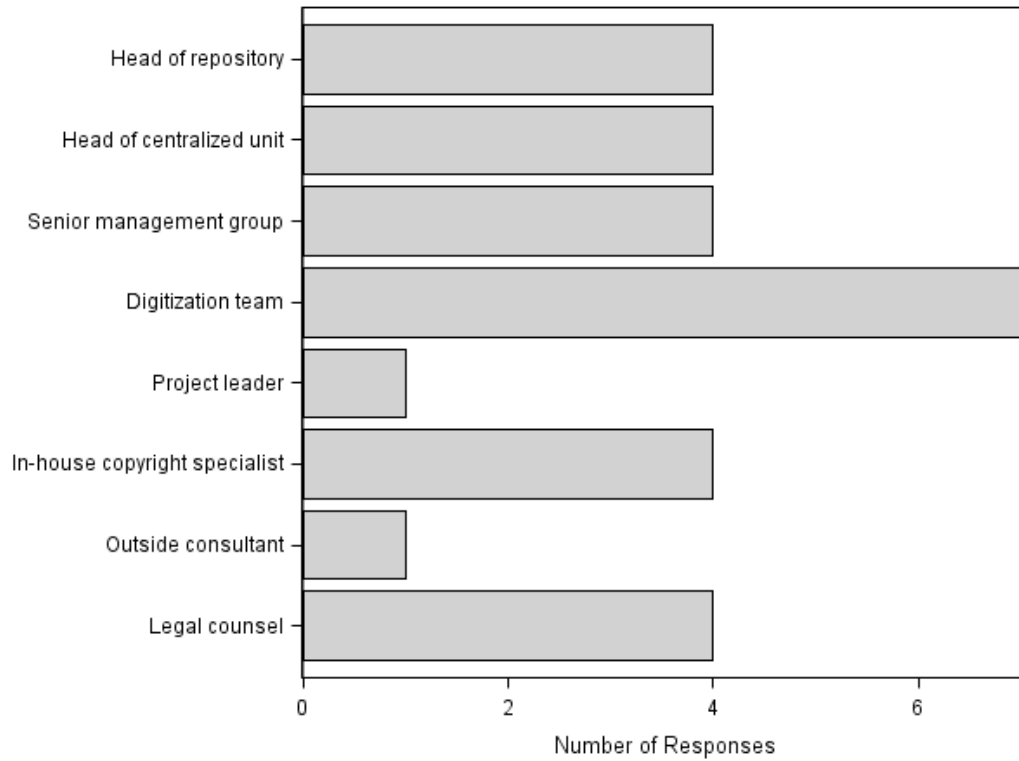


Figure 12: Parties responsible for developing copyright policies or guidelines (n=16)

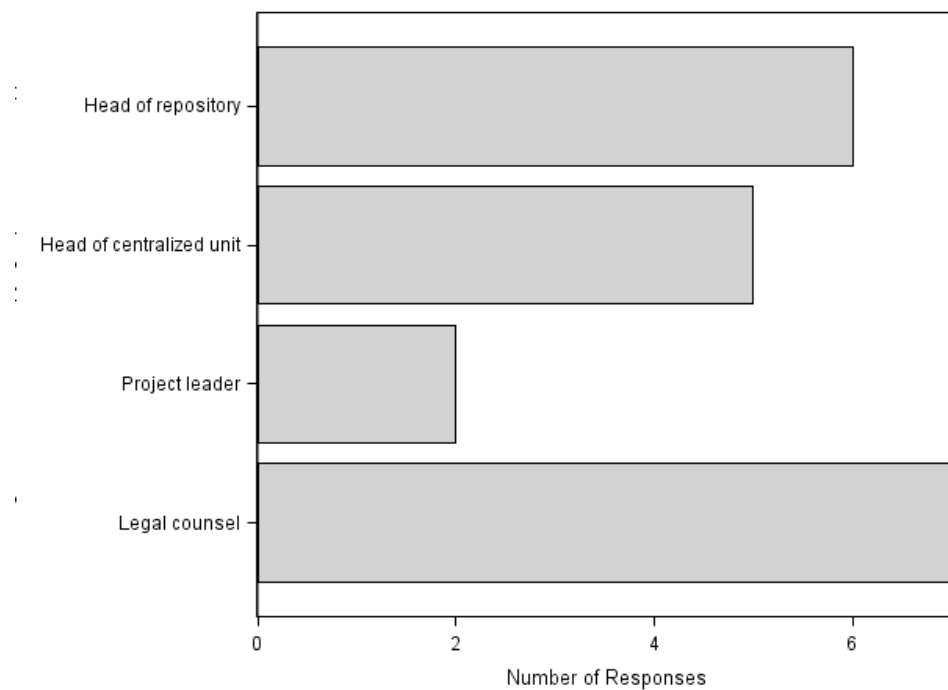


Figure 13: Parties responsible for approving copyright policies or guidelines (n=16)

66% of respondents did not know when the policies were written or adopted, and 40% did not know how often the policies are reviewed or revised. 53% of repositories that currently have policies have updated those policies within the last two years; however, respondents from the other 47% did not know the last time those policies had been revised.

Fifteen of thirty-four respondents (44%) have a staff member who is designated as responsible for copyright concerns; nine of those fifteen (60%, 24% of the total respondents) were the designated copyright individual for their institution.

#### *4.5 What copyright information is provided to users of digitally published materials?*

58% of respondents claimed that their repositories are concerned that website visitors may violate the copyright of available materials available, but all but three repositories use a variety of strategies to limit illegal use of their digitally published materials. Seventeen repositories make a copy of their copyright policy available to users of their digital materials; this is particularly interesting as fifteen respondents indicated that their repository had written copyright policies. Eleven of the fifteen made those policies available online, while a further six repositories that had earlier indicated that they did not have official policies said that they had posted them on the Internet for their users. This raises questions about whether these are official policies and whether respondents understood what was meant when asked about official policies.

Figure 14 illustrates other strategies utilized by repositories to limit illegal use, including hosting low-resolution images and including different kinds of statements about acceptable use.

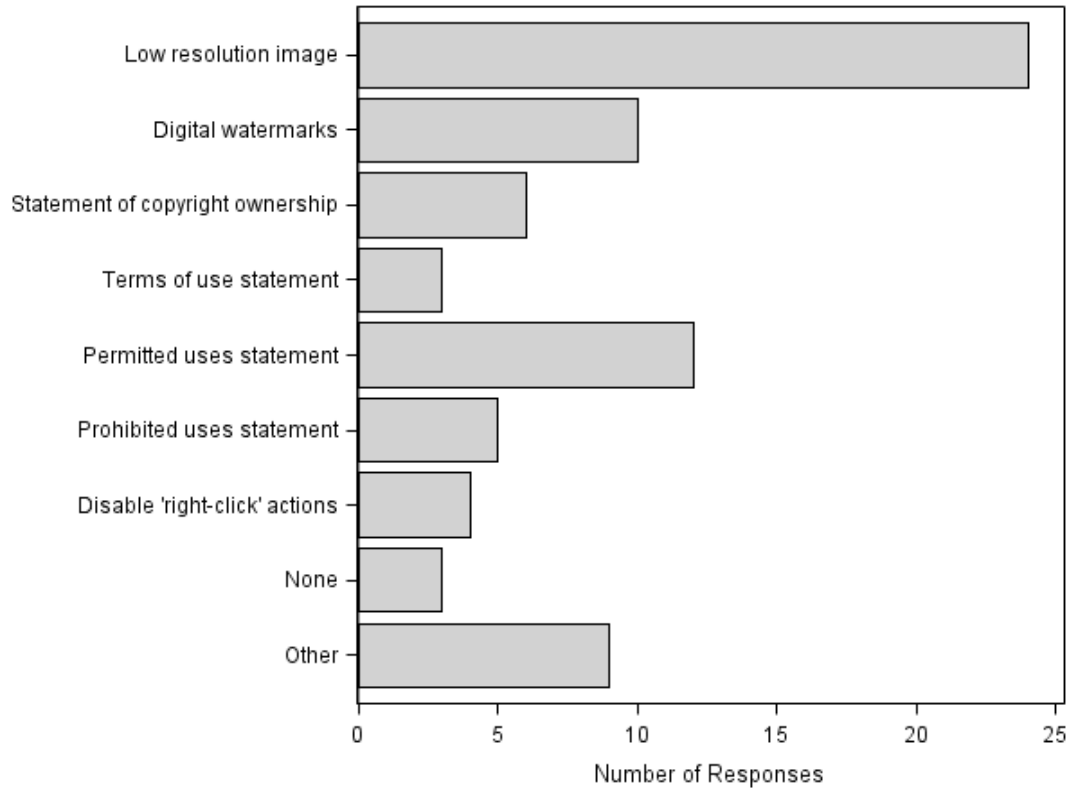


Figure 14: Strategies for limiting illegal use of materials (n=33)

Other strategies for limiting use include statements about the public domain, links to terms of use, copyright statements on each page and in image metadata.

Copyright guidance is provided to users in multiple forms, as well. Copyright statements, information about how to order copies, and owner information are most common, but there are other strategies, too. Few respondents fail to provide any guidance for their users.

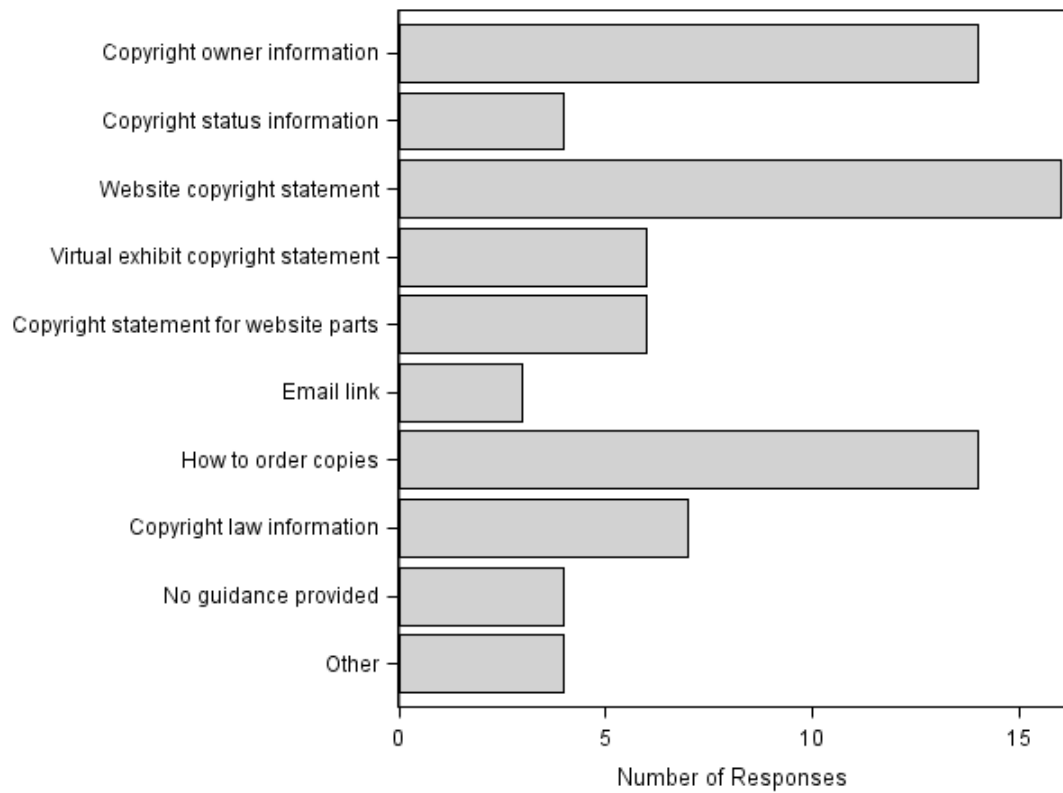


Figure 15: Copyright guidance provided to website users (n=33)



## **DISCUSSION**

### 5.0 Context

While thirty-eight responses is not empirically high, the wide range of repository types implies that the data are somewhat representative of the copyright behavior of American archival repositories. Responses to contextual questions about digitization activities undertaken by repositories as well as the respondents' functional job descriptions indicate that as a group, respondents were knowledgeable about digitization. Based on job titles (26% of which included the words digital or digitization) and descriptions, many could be described as digitization experts, and many appeared to have personal knowledge of digitization practices and policies at their institutions. This can be taken as a measure of the respondents' expertise, and therefore the trustworthiness of the data.

#### *5.1 To what extent do copyright issues influence the selection of manuscript materials for digital publication?*

The first research question had much in common with Astle and Muir's study (2002), and the similar study undertaken by Bültmann, Muir, and Wictor (2006), both of which addressed selection criteria for digitization. Both studies found that, while copyright issues were considered important criterion for selecting materials for digitization, copyright was not as important as other issues, including "access", patron demand, and condition and rarity of materials.

This research indicates that copyright plays a more significant role in selection than in previous studies. Copyright issues were considered to be “extremely important”—the highest rating—by a majority of respondents; no other factor was rated by more respondents as being “extremely important”, though it is important to note that since this study focused on copyright issues, this would be at the forefront of respondents’ minds during the completion of the survey.

Astle and Muir also stated in their study that the range of materials that has been digitized is limited by copyright restrictions, but it was unclear what parts of their data supported this assertion. The current study showed that a majority—92%—of all respondents’ repositories consider copyright before they begin a digitization project, and 65% of respondents’ repositories are willing to publish digital objects on the Internet when the copyright is owned by a third party. Repositories appear to select materials for which copyright has expired or is owned by the repository more often than they select materials for which copyright is owned by an outside party. Of the 35% of repositories that do not publish copyrighted digital objects, 69% digitize documents for their own use, without publishing them. These data support the claim made by Astle and Muir that repositories feel limited in what they can digitize based upon copyright restrictions.

Repositories put a varying amount of effort into determining the copyright status of their materials. It would appear that many repositories’ legal contracts or deeds of gift include clauses specifically transferring copyright to the repository. This is a helpful strategy that can save institutions a great deal of time and frustration when they later wish to digitize objects; being secure in the knowledge that they own copyright precludes a lengthy search to identify and locate copyright holders.

It can be concluded from these responses that copyright issues play a significant role in what repositories are willing to publish publicly on the Internet. Further research should be undertaken to determine the extent to which these issues are dependent upon the type of repository and the types of materials being digitized.

*5.2 How do archives allocate resources to obtaining copyright permissions, and how successful are these efforts?*

While twenty-five respondents (66%) indicated that their repository selects materials to digitally publish when the copyright belongs to a third party, nineteen of the respondents, 50% of the total number and 76% of those who digitize third-party-owned documents, make an attempt to obtain authorization to publish materials from the copyright owners.

Research question two expanded on research completed by George (2005). George's study of attempts to identify and locate copyright holders concluded that of 273 requests for permission to digitize published materials that were made, twenty-four percent of the total agreed to allow digitization, and nine percent of the total requests made placed no restrictions upon that digitization, such as restricted access or a request for payment. The current study data indicate that obtaining permission to digitize unpublished materials is significantly easier than obtaining permission to digitize published materials, as found in George's study.

An important avenue for future work is gaining knowledge about the success or failure of attempts to contact copyright holders. This type of research can be used to indicate the effectiveness and efficiency of the techniques used to contact copyright

holders. Combined with the results of the current study, it may be possible to better understand the value of these techniques and the behavior of repositories and copyright holders during the process of digitization.

An important avenue for future work is gaining knowledge about the success or failure of attempts to contact copyright holders. This type of research can be used to indicate the effectiveness of the techniques used to contact copyright holders, as well as the efficiency of these techniques.

### *5.3 What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?*

Surprisingly, a minority of respondents stated that their repositories had official copyright policies or guidelines and there was no consensus among respondents regarding who wrote or approved their policies, or how often their policies had been updated. Repositories would be well advised to develop copyright policies that take risk assessment into consideration that can help guide their employees when dealing with copyright matters.

Generally, the survey responses suggest that employees need more guidance and that repositories should create policies in consultation with copyright experts that help them manage their risks. This is important as a number of respondents' repositories have been challenged after publishing a digital object on the Internet.

#### 5.4 *What copyright information is provided to users of digitally published materials?*

Congleton's review of twelve websites revealed that ten of them made copyright information available to their users (2005). In this study, 91% of repositories seek to limit illegal use of digital materials through watermarks or other digital means, and 88% offer some form of copyright information to users and. Three respondents (9%) admitted that their repositories provide no deterrents to illegal use on their Internet pages, but all three of these repositories make their official copyright policy available, which may limit their liability, apparently trusting users to do the right thing.

More worrying are the 12% of respondents whose repositories provide no copyright guidance for users on their web pages. None of these repositories have official copyright policy, but they have not made any other copyright information available to users in other forms. Repositories should put some time and effort into developing copyright policies, not only for employees, but also to assist their users.

It is encouraging that such a small proportion of repositories offer no guidance to their users, but it is unclear what the significance of this fact is or how valuable the guidance is. Copyright law can be very confusing; it is possible that copyright statements may be confusing as well, and there is potential that by providing statements about copyright to users, repositories are transferring responsibility for using images to their users. More studies need to be undertaken in order to draw conclusions about how well repositories are interpreting copyright law to their patrons.

## **SUMMARY**

In spite of the breakneck pace of technological innovation, copyright law has not kept up, which makes interpreting the law for new technologies difficult and imprecise. The copyright status of some manuscript materials can make it risky to put these materials on the World Wide Web, where they would benefit many more people. There has been considerable discussion in the professional literature about how repositories can best protect themselves from copyright litigation while still maintaining a high standard of accessibility for their materials. The current study explored the ways in which American repositories navigate copyright law while serving their institutional goals and the best interests of their patrons.

As technology continues to improve, repositories will have more opportunities to place digital copies of their archival holdings online for a larger audience. In addition, technological advances will allow more options for the representation of those materials in a digital environment. Those most interested in the results of this study are likely archivists, historians, and other information professionals with a stake in the digitization of manuscript collections. This study will inform these professionals about how their peers at other repositories approach these issues. Because the legality of digitizing archival materials is still poorly described and understood, repositories may be more willing to take risks if they know how others have justified their copyright decisions.

Further research should be undertaken, including a broad, in-depth study such as Dryden's should be carried out with regard to American repositories. Dryden's study was able to survey 106 repositories and conduct follow-up interviews with 22 staff members. This survey of 38 cannot match Dryden's depth. Furthermore, while this survey only asked questions about copyright policies and practices in American repositories. Dryden's study asked many more questions about copyright in Canadian repositories, including the role of Canadian archivists in shaping repository policy. The present study has created a base-line from which subsequent studies may continue to explore questions about copyright and digitization in American archival institutions.

## References

- Allison-Bunnell, J. L. (1995). Access in the time of Salinger: Fair use and the papers of Katherine Anne Porter. *American Archivist* 58(3), 270-82.
- Astle, P. J., & Muir, A. (2002). Digitisation and preservation in public libraries and archives. *Journal of Librarianship and Information Science* 34(2), 67-78.
- Babbie, E. (2007). *The practice of social research* (11th ed.). Belmont, CA: Thomson Wadsworth.
- Baksik, C. (2006). Fair use or exploitation? The Google Book Search controversy. *Portal* 6(4), 399-415.
- Besek, J. M. (2003). *Copyright issues relevant to the creation of a digital archive: A preliminary assessment*. Washington, DC: Council on Library and Information Resources.
- Breaden, I.C. (2006). Sound practices: On-line audio exhibits and the cultural heritage archive. *American Archivist* 69(1), 33-59.
- Bültmann, B., Hardy, R., Muir, A. & Wictor, C. (2006). Digitized content in the UK research library and archives sector. *Journal of Librarianship and Information Science* 38(2), 105-22.
- Congleton, R. J. (2005). Evaluating web sites featuring primary sources on United States history. *Behavioral & Social Sciences Librarian* 24(1), 61-82.
- Crawford, M. J. (1983). Copyright, unpublished manuscript records, and the archivist. *American Archivist* 46(2), 135-47.
- Dryden, J. E. (2008). Copyright in the real world: Making archival material available on the Internet. (Doctoral dissertation, University of Toronto, 2008). Retrieved August 28, 2008 from [https://tspace.library.utoronto.ca/handle/1807/11198/1/Dryden\\_Jean\\_E\\_200806\\_PhD\\_thesis.pdf](https://tspace.library.utoronto.ca/handle/1807/11198/1/Dryden_Jean_E_200806_PhD_thesis.pdf)
- Duff, W., Craig, B., & Cherry, J. (2004). Historians' use of archival sources: Promises and pitfalls of the digital age. *The Public Historian*, 26(2), 7-22.
- Dunning, A. (2004). *Tracing Copyright Holders: How two digitisation projects coped with copyright for historical material*. London: Arts and Humanities data service (ADHS). Retrieved September 7, 2008 from <http://ahds.ac.uk/creating/case-studies/tracing-copyright/index.html>



- George, C. A. (2005). Testing the barriers to digital libraries: A study seeking copyright permission to digitize published works. *New Library World* 106(7/8), 332-42.
- Grogg, J. E. & Ashmore, B. (2008). Google Book Search libraries and their digital copies. *Journal of Library Administration* 47(1/2), 125-40.
- Hirtle, P. B. (2001). Unpublished materials, new technologies and copyright: Facilitating scholarly use. *Journal of the Copyright Society of the USA* 49(1), 258-75.
- Hirtle, P. B. (2009). Copyright term and the Public Domain in the United States. Retrieved April 1, 2009 from [http://www.copyright.cornell.edu/public\\_domain/](http://www.copyright.cornell.edu/public_domain/)
- Lackie, R. J. (2008). From “Google Print” to “Google Book Search”: The controversial initiative and its impact on other remarkable digitization projects. *The Reference Librarian* 49(1), 35-53.
- LeFevre, K. B. (1992). The tell-tale ‘heart’: Determining ‘fair’ use of unpublished texts. *Law and Contemporary Problems* 55(2), 153-83.
- Liew, C. L. (2006). Online cultural heritage exhibitions: A survey of strategic issues. *Program: Electronic Library and Information Systems* 40(4), 372-88.
- Society of American Archivists. (2008). Manuscript Repositories Section. Retrieved November 17, 2008 from <http://www.archivists.org/saagroups/mss/>
- Society of American Archivists. (2008). Metadata and Digital Object Roundtable. Retrieved November 17, 2008 from <http://www.archivists.org/saagroups/metadata/>
- Minow, M. (2002). *Library digitization projects and copyright*. Retrieved September 5, 2008 from <http://www.llrx.com/features/digitization.htm>
- Mugridge, Rebecca. (2006). *Managing digitization activities*. Washington, DC : Association of Research Libraries.
- National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research. (1979). *The Belmont report: Ethical principles and guidelines for the protection of human subjects of research*. Retrieved November 30, 2008, from <http://ohsr.od.nih.gov/guidelines/belmont.html>
- Pearce-Moses, Richard. (2005). *A glossary of archival & records terminology*. Chicago: Society of American Archivists.
- Pritcher, L. (2000). Ad\* Access: Seeking copyright permissions for a digital age. *D-Lib Magazine* 6(2). doi: 10.1045/february2000-pritcher
- Rastenberger, M. & Weston, C. (2005). *Overview of the Libraries and Archive Exception in the Copyright Act: Background, History, and Meaning*. Retrieved October 3, 2008 from [http://www.section108.gov/docs/108BACKGROUND.PAPER\(final\).pdf](http://www.section108.gov/docs/108BACKGROUND.PAPER(final).pdf)

- Schonlau, M., Fricker, R. D. & Elliott, M. N. (2002). *Conducting research surveys via e-mail and the web*. S.I.: Rand.
- Spoo, R. (1999). Fair use of unpublished works: Scholarly research and copyright case law since 1992. *Tulsa Law Journal* 34(1), 183-200.
- Spoo, R. (2001). Copyright law and archival research. *Journal of Modern Literature* 24(2), 205-12.
- Thatcher, S. G. (2006). Fair use in theory and practice: Reflections on its history and the Google case. *Journal of Scholarly Publishing* 37(3), 215-29.
- Tibbo, H. R. (2003). Primarily history in America: How U.S. historians search for primary materials at the dawn of the digital age. *American Archivist*, 66(1), 9-50.
- United States Constitution. Article I, § 8, clause 8.
- United States Copyright Office. (2007, October). Copyright law of the United States of America and related laws contained in title 17 of the United States Code (circular 92). Washington, DC: U.S. Government Printing Office. Retrieved October 19, 2008 from <http://www.copyright.gov/title17/>.

## Appendix A: Survey Instrument

In completing this survey, please consider current or recent (within the past five years) digitization projects undertaken by your repository that will be published on the Internet and available to the general public. If you have already participated in this survey through another group, please do not respond a second time.

### A. Background Information

1. In what type of repository do you currently work?
  - National or state archive
  - Municipal, county, or regional archive
  - Manuscript repository
  - Religious archive
  - University archive
  - Corporate archive
  - Historical society
  - Museum
  - Library
  - Other (please specify): \_\_\_\_\_
  
2. Which of the following activities is your repository currently engaged in?
 

*Check all that apply*

  - Digitizing materials in our collection
  - Making digitized materials available to individuals within the repository
  - Making digitized materials available to individuals outside the repository
  - Making digitized materials available to the general public
  
3. In what year did your repository begin digitizing manuscript materials?
  
4. In what year did your repository begin publishing digitized manuscript materials on the Internet?
  
5. What is your current job title?
  
6. In your current position, approximately what percentage of your time is spent working on digitization projects?
  
7. Briefly describe your duties that relate specifically to making your repository's holdings available on the Internet.

B. *To what extent do copyright issues influence the selection of manuscript materials for digital publication?*

8. How important are the following criteria when considering what manuscripts to digitize and publish on the Internet?

*Question was coded using a 7-point Likert-style scale, which ranged from extremely important to not at all important*

Availability of technical expertise and equipment

Availability of staff resources

Availability of grant funding

Desire to increase web resources so that researchers can serve themselves

Historical or cultural value of materials

Feedback from patrons about what they want

Copyright issues

Possibility of generating income

Subject matter of documents

Physical condition of documents

Rarity or uniqueness of documents

Collaboration with other institutions

Other (please specify): \_\_\_\_\_

Other (please specify): \_\_\_\_\_

Other (please specify): \_\_\_\_\_

9. Is the copyright status of documents considered before a digitization project is begun?

Yes

No

Don't know/Not sure

10. In selecting documents from your repository's holdings to make available in digital form to the public on your repository's website, does the selection include documents for which the copyright has expired?

Yes

No

Don't know/Not sure

11. Briefly describe how it is determined that copyright has expired for any particular document or set of materials.

12. In selecting documents from your holdings to be digitized and made available to the public on your repository's website, does the selection include documents in which the repository owns the copyright?

Yes

No

13. Briefly describe how it is determined that your repository owns the copyright for any particular document or set of materials.

14. In selecting documents from your holdings to be included on your repository's website, does the selection include documents for which the copyright is not owned by your repository?

Yes (If yes, then skip to Section C, Q16)

No (If no, continue to Q15)

15. Does your repository digitize documents without publishing them on your website (for example, for internal purposes or to wait until copyright has expired)?

Yes (If yes, then skip to Section D, Q24)

No (If no, then skip to Section D, Q24)

C. *How do archives allocate resources to obtaining copyright permissions, and how successful are these efforts?*

16. Does your repository try to obtain the authorization of copyright owners to allow you to put digital copies of their documents on your repository's website?

Yes (If yes, continue to Q17)

No (If no, then skip to Q22)

17. How important are the following factors in your repository's decision to end attempts to locate copyright holders?

*Question was coded using a 7-point Likert-style scale, which ranged from extremely important to not at all important*

Lack of financial resources

Lack of time

Can't identify copyright holder(s)

Can't locate copyright holder(s)

Have received no response from apparent copyright holder(s)

Other (please specify): \_\_\_\_\_

Other (please specify): \_\_\_\_\_

Other (please specify): \_\_\_\_\_

18. Who makes decisions about when to end the search for a copyright holder?

*Check all that apply*

Head of centralized unit

Digitization team/committee/working group

Project leader or officer

Other (please specify): \_\_\_\_\_

19. What does your repository do if the copyright owner cannot be located or does not respond to your request?

*Check all that apply*

Do not use the document

Substitute another document, if possible

Use the document with a disclaimer (for example, a statement that informs users that efforts to locate the copyright owners were unsuccessful but that copyright owners who come forward will be appropriately acknowledged)

Use the document with no disclaimer

Other (please specify): \_\_\_\_\_

20. If your repository is able to successfully locate a copyright owner to obtain permission for the use of a document, what has been the copyright owner's reaction to your request?

*Please indicate a percentage by typing in a number between 0 and 100*

Declined permission

Gave unconditional permission

Gave conditional permission (for example, wanted royalties or other payment)

Other (please specify): \_\_\_\_\_

21. If you answered "gave conditional permission" to Question 18, what were the conditions?

*Please indicate a percentage by typing in a number between 0 and 100*

Wanted one-time payment

Wanted royalties or licensing fees

Wanted credit line

Other (please specify): \_\_\_\_\_

22. Has your repository ever been challenged by a copyright owner over the use of a document on your repository's website?

Yes

No

23. Please describe the most recent or most notable situation when this occurred and how your repository dealt with it.

*D. What are the official policies of archival repositories regarding copyright issues, and who is responsible for maintenance and compliance?*

24. Does your repository have written policies, procedures, or guidelines relating to the administration of copyright in your institution?

Yes (If yes, continue to Q25)

No (If no, then skip to Q30)

25. Who developed your repository's copyright policies/procedures/guidelines? *Check all that apply*

Head of repository

Head of centralized unit

Senior management group

Digitization team/committee/working group

Project leader or officer

In-house copyright specialist

Outside consultant

Other (please specify): \_\_\_\_\_

26. At what level of your organizational structure have your repository's copyright policies/procedures/guidelines been approved?

*Check all that apply*

Head of repository

Head of centralized unit

Project leader or officer

Other (please specify): \_\_\_\_\_

27. In what year were your repository's copyright policies/procedures/guidelines originally written?

Year: \_\_\_\_\_

Not sure/don't know

28. How often are your repository's copyright policies/procedures/guidelines reviewed or revised?

More than once a year

About once a year

About every five years

Don't know/Nor sure

Our policies have not been updated since they were written

Other (please specify): \_\_\_\_\_

29. In what year were your repository's copyright policies/procedures/guidelines last reviewed or revised?

Year: \_\_\_\_\_

Not sure/don't know

30. Who in your organization makes decisions in situations where the copyright policies/procedures/guidelines are difficult to apply? (Job title only please, not given name)

31. Does your repository have a designated staff member who is responsible for copyright matters?

Yes

No

32. Are you that person?

Yes

No

33. Briefly describe the duties of the person responsible for copyright matters.

E. *What copyright information is provided to users of digitally published materials?*

34. Is your repository concerned that visitors to your repository's website may violate the copyright of the original materials?

Yes

No

35. What measures, if any, does your repository take to limit illegal use of digital copies of materials?

*Check all that apply*

Low resolution images

Digital watermarks across the image

Statement about copyright ownership across the image

Statement about terms of use across the image

Statement about permitted uses on each web page

Statement about prohibited uses on each web page

Disabling the 'right-click' function

None

Other (please specify): \_\_\_\_\_

36. Does your repository make a copy of its copyright policy available to users on its website?

Yes

No

My repository does not have an official copyright policy

37. What copyright guidance, if any, does your repository provide to visitors to your website?

*Check all that apply*

Information about the owner of the copyright in individual documents

Information about the copyright status of individual documents (e.g., whether the copyright has expired)

Copyright statement for the entire website

Copyright statement for each virtual exhibit

Copyright statement for each part of the website

Email link specifically for copyright inquiries

How to order copies of documents

Information about copyright law

My repository does not provide copyright guidance to visitors to our website

Other (please specify): \_\_\_\_\_

38. If you have any further comments to make about copyright and its role in digitizing manuscript materials at your institution, please feel free to make them below.