

The Triangle Research Libraries Network's Intellectual Property Rights Strategy for Digitization of Modern Manuscript Collections and Archival Record Groups

January 2011

Laura Clark Brown, Judy Ruttenberg, and
Kevin L. Smith, J.D.

Introduction

The Triangle Research Libraries Network (TRLN) member libraries are Duke University, North Carolina Central University (NCCU), North Carolina State University (NCSU), and the University of North Carolina at Chapel Hill (UNC at Chapel Hill). Each member library holds outstanding special collections including manuscript collections and archival record groups that document the recent American past.

These manuscript collections and archival record groups vary in size, ranging from a few documents to hundreds of thousands of documents (such as in most modern collections—those dating from the mid-twentieth century to the near present). Both the collections and groups contain a variety of personal, professional, and organizational materials including correspondence, memoranda, diaries, journals, ledgers, reports, scrapbooks, newspaper clippings, photographs and photograph albums, audio and video recordings, pamphlets, brochures, and leaflets. The majority of the materials are unique, have never been published, and, until the advent of digital technology, were accessible only in the reading rooms of the libraries where they were housed.

Curated online exhibits have dominated the manuscript and archival digitization projects of academic libraries. These exhibits typically present a small number of selected documents along with a description of each document, transcriptions, and scholarly analysis. An emerging programmatic model—large-scale digitization of manuscript collections and archival record groups—provides online access to entire collections/groups. Given the typical size of modern collections/groups, large-scale digitization cannot support the presentation of the item-level details used in curated online exhibits. This leaner approach to presentation allows for a greater volume of accessible documents. As technology and online applications evolve and the body of materials available online increases, many of these documents will probably be made full-text searchable either by the libraries or by the researchers themselves.¹

¹ Optical Character Recognition (OCR) software makes typescript documents machine-readable. Social web applications and plug-ins could enable scholars and others to contribute transcriptions of handwritten documents.

In 2010, with funding from the State Library of North Carolina, the TRLN member libraries designed a large-scale digitization project that will present free and open online access to a total of forty digitized manuscript collections or archival record groups from the four libraries (ten collections/groups from each library). Called “Content, Context, and Capacity: A Collaborative Digitization Project on the Long Civil Rights Movement in North Carolina” (CCC), the project will present digital images of documents online, accompanied by the broad summary descriptions and contents lists found in the finding aids created when the collections were processed.²

For the first time, these resources will cross the boundaries of the four libraries’ reading rooms—bringing together a vast quantity of research material for the era between the 1930s and 1980s that historians call the Long Civil Rights Movement (LCRM) (Hall, 2005).³ Given the enduring, global interest in this social movement, the free and open online availability of full collections will facilitate new scholarly collaborations across institutions, and even nations, and will support the development of educational tools for students and the use of primary sources in classrooms.

With large-scale digitization of modern manuscripts and archival materials come complex intellectual property challenges because any of the hundreds of thousands of individual documents—rather than the collection as a whole—could potentially remain under copyright protection. In curated online exhibits, libraries have avoided these complexities and ambiguities by digitizing materials old enough to be out of copyright protection or by selecting a small number of documents for which it was reasonable to obtain permissions. As a result, materials dating from the eighteenth and nineteenth centuries dominate the digital landscape for

² A finding aid is a description of a manuscript collection or archival record group. It provides information about a collection/group and its creators along with an outline of a collection/group's contents.

³ In a groundbreaking article, historian Jacquelyn Dowd Hall recast the narrative of the civil rights movement, lengthening the classical chronology—the decade between the *Brown v. Board of Education U.S. Supreme Court* decision of 1954 and The Voting Rights Act of 1965—in both directions to define a “Long Civil Rights Movement” (LCRM). The new narrative also broadens and deepens the classical storyline—a linear progression that resulted in the end of racial segregation—to include the ongoing struggles against economic, social, and environmental injustice. In the LCRM, the cast of well-known heroes and villains swells, and no longer excludes the grassroots players—such as the labor unions and community organizers—or the LCRM’s opponents in the emerging New Right.

southern history; and, because fewer modern documents are widely available online, researchers have had an incomplete view of the recent American past. The CCC project aims to provide an unprecedented level of access to a large aggregation of modern documents, and will employ the Intellectual Property Rights Strategy outlined below.

Intellectual Property Rights in Manuscripts and Archival Documents

In the United States, the statutory term for copyright of unpublished works is seventy years from the death of the author and one hundred and twenty years from the date of creation for anonymous works, which places the majority of modern documents under presumed copyright protection. And potential rights holders are innumerable.

Broadly speaking, manuscripts and archival documents fall in one of three categories of possible copyright status: 1) documents in the public domain, 2) documents that have a copyright holder who has the power to grant permission for their use (although some of those rights holders may be impossible to identify and locate), and 3) documents for which gift agreements specify the transfer of copyright to the institution where they are housed.

These copyright categories apply to the full range of materials including portions of printed publications (e.g., newspaper clippings), works for hire (e.g., organization reports from now defunct organizations), and portions of commercially broadcast media (e.g., radio interviews). A single collection, and potentially a single file folder within a collection, could contain materials in any or each category.

For example:

Frank Porter Graham, president of the consolidated University of North Carolina (1932–1949), U.S. Senator, and United Nations official, corresponded with hundreds of individuals during the five decades (1920–1970) documented in his collection of papers.⁴ Unpublished letters from Populist politician William Jennings Bryan—who

⁴ Frank Porter Graham Papers (01819), Southern Historical Collection, Louis Round Wilson Special Collections Library, The University of North Carolina at Chapel Hill.

died in 1925—are in the public domain because the seventy-year statutory term of copyright has passed. Unpublished letters from American theologian Reinhold Niebuhr—who died in 1971—might remain under copyright protection. An unpublished letter from an individual whose death date is unknown and who is today untraceable might also remain under copyright protection.

It would be impossible to determine all the rights holders who are associated with these multiformat collections, to sort out which rights are held where, and to obtain all possible permissions to create a digitization project completely free of rights concerns (Dickson, 2010).⁵ Therefore, the CCC project must proceed on the basis of an intelligent assessment of the potential risks and reasonable efforts to reduce those risks.

TRLN’s Intellectual Property Rights Strategy

In spring 2010, the national library service organization OCLC Research convened experts in scholarly communications and archival administration to define new practices for managing copyright in the digital age. Many of the models for expanding access that digital technologies engender have often been met with what participants regarded as undue caution by libraries in facilitating that access. Libraries have avoided digitizing modern documents so that their projects would not be encumbered by intellectual property rights concerns.

The goal of the OCLC Research meeting was to encourage libraries to leverage the full range of rights management strategies and the strength of the fair use argument in order to realize their

⁵ In 2007 librarians at UNC at Chapel Hill conducted intensive copyright research on the 3,304 individual letter writers in the correspondence contained in the Thomas E. Watson Papers. Using a variety of online sources—including Wikipedia, the Social Security Death Index, and Ancestry.com, and print reference works such as biographical dictionaries—staff attempted to identify the correspondents and to find dates of death in order to determine copyright status. The librarians found that 608 correspondents (18.4%) had life dates that precluded the materials from copyright protection and thus were in the public domain; and 1,101 correspondents (33.32%) had life dates that placed the materials under copyright protection. They could not find life dates for 1,571 correspondents (47.55%), and they found no information for twenty-four correspondents (.73%). The identification, evaluation, and research processes for a relatively small body of materials required more than fourteen weeks of dedicated time by a full-time, professional librarian and a half-time graduate research assistant. The project librarian, Maggie Dickson, published these findings in the archival profession’s leading journal.

missions as teaching and research institutions. Following the meeting, OCLC Research published a document, titled “Well-intentioned practice for putting digitized collections of unpublished materials online,” which has been formally endorsed by professional organizations including the Rare Book and Manuscript Section (RBMS) of the American Library Association (ALA), and leading academic library professionals and scholarly communications officers (OCLC, 2010).⁶ In keeping with this endorsed statement and in tandem with the ever-growing agreement within the profession, TRLN member libraries propose a three-pronged strategy to promote free and open online access through well-reasoned risk management.

1. Recognize that some materials may already be in the public domain

Most modern archival documents under consideration are not in the public domain. But if the materials meet any of the following criteria, they are considered to be in the public domain: 1) If the material was published prior to 1923; 2) If unpublished material was created by an author who has been dead seventy or more years; 3) If the material was “work made for hire” in the course of employment (e.g., official university records); or 4) If the work was published either without copyright notice, or with notice but without renewal, during the periods of time when notice and renewal was required.

2. Obtain permissions and rights when and where practicable

In consultation with the scholarly communications officers in the TRLN member libraries, the CCC project group will develop a decision matrix listing the various formats, along with the rights concerns associated with each of them, including the level of care to be exercised in the initial selection of collections/record groups. During the selection process, the CCC project group will avoid including collections/record groups that contain substantive works by authors with literary estates, medical records, or student work.⁷

Identifying all potential rights holders for all the documents in the selected collections is impossible. However, TRLN member libraries will request permissions from known and likely

⁷ In addition to intellectual property considerations, student records and medical records raise privacy concerns. Student records are protected by the Family Educational Rights and Privacy Act (FERPA) of 1974; medical records are protected by the Health Insurance Portability and Accountability Act (HIPAA) of 1996.

rights holders including prominent public figures, collection donors, and newspaper publishers. Any problematic materials discovered after the initial selection will be removed from the digitization queue by the libraries.

Although this intellectual property rights strategy establishes a community standard for the TRLN member libraries, specific practices in the individual libraries will differ. Each library will determine the necessity and practicality of obtaining permissions for its own materials, and will respond to permission denials as appropriate for its own circumstances. At the sole discretion of the library that owns the materials, digitization might still proceed even after a permission that was sought was denied. Denied permission does not irrevocably harm a fair use defense (see #3 below).

All efforts to obtain permissions will be documented, and the documentation will be retained in the permanent files of each library.

3. Acknowledge the strength of a fair use argument

As educational institutions with missions to provide access to the materials housed in the libraries, the strongest argument the TRLN member libraries can offer for the right to present manuscript collections and archival record groups online is fair use.

In the unlikely event that a TRLN member library is challenged on the presentation of the digitized collections/groups, and in the even more unlikely event that the library is unable to resolve those challenges (see section below on responding to challenges), the library will rely on a fair use argument.

Four main factors are considered in an assessment of fair use, as defined in the U.S. Copyright Act, Title 17 of the *United States Code*, Section 107, excerpted below:

“In determining whether the use made of a work in any particular case is a fair use the factors to be considered shall include —

- (1) the purpose and character of the use, including whether such use is of commercial nature or is for nonprofit educational purposes;
- (2) the nature of the copyrighted work;
- (3) the amount and substantiality of the portion used in relation to the copyrighted work as a whole;
- and
- (4) the effect of the use upon the potential market for or value of the copyrighted 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.”

Fair use is a balancing test based on these four factors. The factors are not a list of requirements, and all four factors need not be met to have a successful fair use argument. Each factor as it might pertain to the CCC project’s selected manuscript collections and archival record groups is addressed below.

The purpose and character of the use

The CCC project is not for profit; the project’s purpose is to promote historical scholarship and support educational uses of primary sources by providing free and open online access to a large corpus of research materials: the digitized manuscript collections and archival record groups from the four libraries. Individual documents contained in the collections and groups may remain under copyright protection, but they are used in this project for research and educational purposes.

The character of the use is transformative. An individual document’s original use was temporally bound, its value practical. For example, at the time that any individual letter in the Frank Porter Graham papers was written, it served only to share information with Graham. But now that letter is part of a manuscript collection that contains more than one hundred thousand documents. And as with any individual item in a manuscript collection/archival record group,

the document serves as a small part of a larger resource, one used in scholarly inquiry and education. The aggregation and organization of individual documents to create manuscript collections/archival record groups transform the purpose and function of the individual documents, as do the finding aids for these collections/groups, which also add to the research value. The digitization and online presentation of the documents in the collection/group further remove the individual document from its original purpose, and deepen its transformed purpose and use as a historical resource that contributes to our understanding of the past.

The nature of the copyrighted work

Most of the documents in the manuscript collections/archival record groups were created in the course of the daily life of an individual or in the routine business of an organization. Created without commercial motivation or artistic intent, these works were not meant for publication at the time of creation, and today are not publishable in isolation. The research value of manuscripts lies not with the individual document, but rather with the collection of documents that together provide context and insight into the past.

The amount and substantiality of the portion used

Researchers expect and need access to entire manuscript collections and archival record groups, and to the full content of the individual documents. Access to partial content impedes scholarly inquiry and hampers creative educational uses. The presentation of entire documents and entire collections/record groups is therefore appropriate for the intended use by students, educators, and scholars. Individual documents in the collections/groups are the copyright-protected works; but the law does not specify a particular amount of a work that can be used without permission. The educational and transformative purposes of the use require the presentation of the works in their entirety and so satisfy this factor.

The effect of the use upon the potential market

The scholarly research value and educational significance of these collections are incalculable, but their aggregated online presentation will have little to no effect on the market value of individual documents. In virtually all cases, no such market exists. In the rare instance in which an individual document has a commercial market, the downloadable digital images will not be of commercial quality and therefore will pose no threat to that market.

Response to complaints

The TRLN member libraries will respond to all complaints arising from the digitization and online presentation of the manuscript collections/archival record groups. Each library's website will include a statement about intellectual property along with contact information. Each library will review complaints for materials on its site, and, if warranted, will remove the digital objects associated with the complaints.

Any remaining risk of a dispute over intellectual property related to these materials is small. Efforts to obtain permissions when it is practical to do so will reduce the number of potential disputants and will also demonstrate the TRLN member libraries' good faith efforts to pursue all due diligence in avoiding infringement.

Lastly, the strong rationale for fair use outlined above will discourage complaints because it significantly reduces the damages for which TRLN member libraries could be held liable in the unlikely event that a fair use defense ultimately failed and infringement was found. Copyright law explicitly provides for this reduction of liability when employees of nonprofit educational institutions make use of protected material within the scope of their employment based on a good faith belief that the use was a fair use (17 U.S. Code § 504(c)(2)).

Conclusion

The risk of conflict over intellectual property rights is small because the challenges will be few, and can be addressed and rectified without litigation. The benefits to education and research are enormous and outweigh the minimal risks. Therefore, the CCC project should proceed as planned.

Endorsement

The set of analyses and policies laid out in this document were endorsed by the directors of the TRLN member libraries and the provosts of the four institutions (Duke University, NCCU, NCSU, and UNC at Chapel Hill) in fall 2010.

References

1. Dickson, M. (2010). "Due Diligence, Futile Effort: Copyright and the Digitization of the Thomas E. Watson Papers." *American Archivist*, 73(2), 626–636.
2. Hall, J. D. (2005). "The Long Civil Rights Movement and the Political Uses of the Past." *Journal of American History*, 91(4), 1233–1263.
3. OCLC Research (2010, May 28). "Well-intentioned practice for putting digitized collections of unpublished materials online." Retrieved December 26, 2010, from <http://www.oclc.org/research/activities/rights/practice.pdf>.