Extending the Reach of Southern Sources
Proceeding to Large-Scale Digitization of Manuscript Collections
Final Grant Report
Prepared by the Southern Historical Collection
University Library
The University of North Carolina at Chapel Hill
for the Andrew W. Mellon Foundation
June 2009
Extending the Reach of Southern Sources
Proceeding to Large-Scale Digitization of Manuscript Collections
Final Grant Report
Prepared by the Southern Historical Collection
University Library
The University of North Carolina at Chapel Hill
for the Andrew W. Mellon Foundation
June 2009

Date: 30 June 2009

Organization: Southern Historical Collection, University Library
The University of North Carolina at Chapel Hill

Address: CB #3926, Wilson Library, Chapel Hill NC 27514-8890

Principal Investigators: Tim West, Director, Southern Historical Collection and Curator of Manuscripts
Kirill Fesenko, Director, Carolina Digital Library and Archives

Project Director: Laura Clark Brown

Contact Person: Laura Clark Brown, ljcb@email.unc.edu

Phone: 919.962.1345 Fax: 919.962.3594

Amount of Grant Awarded: $216,000

Date Grant Received: December 2006

Award Period: April 2007 to June 2009

## Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Executive Summary</td>
<td>4</td>
</tr>
<tr>
<td>Introduction</td>
<td>6</td>
</tr>
<tr>
<td>Summary of the Plan for Large-Scale Digitization</td>
<td>7</td>
</tr>
<tr>
<td>The Challenges of Large-Scale Digitization</td>
<td>9</td>
</tr>
<tr>
<td>Summary of Grant Activities</td>
<td>10</td>
</tr>
<tr>
<td>Key Features of the Digitization Program</td>
<td>21</td>
</tr>
<tr>
<td>Appendices</td>
<td>39</td>
</tr>
</tbody>
</table>
Executive Summary

“Extending the Reach of Southern Sources: Proceeding to Large-Scale Digitization of Manuscript Collections” is a project undertaken by the Southern Historical Collection (SHC), funded through a grant from the Andrew W. Mellon Foundation. The grant enabled the SHC staff to study the feasibility of digitizing the SHC’s collections and to plan a long-term digitization program.

Housed at the Louis Round Wilson Special Collections Library, the SHC encompasses more than 4,600 individual collections comprised of millions of individual documents. Through this project a “decision matrix” was developed for selecting and prioritizing the massive holdings of the SHC for digitization, and the Digital Southern Historical Collection was planned.

The SHC staff investigated the digitization experiences of staff at other archival collections; carefully considered the SHC’s archival mission; and, most critically, conducted extensive conversations with the primary audience for the SHC: the scholars of the history of the American South. Feedback and input were sought through focus groups with graduate students, interviews with individual scholars, attendance at professional conferences, site visits, a scholars’ workshop, and two symposia with archivists and librarians.

The SHC will digitize many of its 4,600+ collections in their entirety. The detailed decision matrix will serve as a guide for the staff as it makes decisions about which collections to digitize and about the priorities for digitization. This matrix has already been used to prioritize 1,030 collections for digitization. Through a digitization-on-demand program, users and collection donors who desire a digital version of a collection that is not slated for digitization, or that was not going to be digitized soon enough to suit their needs, will be able to request that a collection be digitized and will pay a modest fee to help offset the cost of the service.

The SHC and the University Library are committed to developing and sustaining the Digital Southern Historical Collection. This commitment will require the consistent allocation of current staff time and resources as well as the continual expansion of technological capacity and periodic hiring of additional staff to accommodate an ever-growing Digital SHC. Digitization will continue throughout each year, although the pace will be partially determined by the resources allotted through the annual budget and by the availability of external funding.

The digital search room experience will mirror the physical search room experience in which scholars interact with primary sources directly. There are no current plans to frame the digitized manuscripts with contextualizing or interpretive information or to supplement the documents with resources such as scholarly essays, lesson plans, interactive features, or social networking communities. The organization of the digital version of each collection will mirror the organization of its physical counterpart in the SHC.

The digital version of each individual document will be a high-resolution image file. Digitized material will be made available over the Internet free of charge, and users will not be required to enroll in a subscription service, or to be affiliated with a university or other institution.
Although it is difficult to predict the directions in which innovations in technology and historical scholarship will lead, the decisions made at this early stage of the digitization process have left room for the possibility of different choices being made in the future, thus allowing the digital version of the SHC to continually remain flexible and responsive to the needs of its users.
Introduction

For more than seventy-five years scholars have been coming to the University of North Carolina at Chapel Hill (UNC-CH) to explore the manuscripts in the Southern Historical Collection (SHC). There, in the Louis Round Wilson Special Collections Library (hereafter Wilson Library) search room, they can access millions of documents on the American South.

The research process involves going online to the finding aids, identifying collections of interest, then requesting the materials for study. A staff member retrieves the box of materials from the stacks and delivers it to the patron waiting in the search room. Typically, the manuscripts are organized in folders inside each box. After studying the materials the user returns them to the staff member for reshelving.

For those who are able to visit the collection, the SHC’s vast holdings are an invaluable resource—but not everyone can make the trip to Chapel Hill to conduct research in the manner described above. However, what if the SHC was available to anyone, anywhere the Internet can reach? Large-scale digitization of the collection would dramatically change the research process at the SHC—and would likely alter the shape of future scholarship on the American South.

A grant from the Andrew W. Mellon Foundation has made that what if a reality. Through this grant the SHC staff studied the feasibility of digitizing the SHC’s collections and developed an extensible model for a sustainable, long-term large-scale digitization program.

The digitization program is no small undertaking. The collection—built on the work of historian and founding director J. G. de Roulhac Hamilton and his successors, and officially established by UNC-CH in 1930—has grown to more than 4,600 individual collections comprised of millions of individual documents—letters, diaries, maps, ledgers, oral histories, bills, receipts, photographs, literary manuscripts, sound recordings, moving images, pamphlets, printed ephemera, and electronic letters. The collections range in size from single items, such as a soldier’s diary, to thousands of interviews collected by the Southern Oral History Program, to the quarter of a million items in the senatorial records of former senator Sam Ervin Jr. (D-N.C.).

The scope of the collection is no less impressive than its size. The SHC offers strong documentation for all periods of southern history since the late eighteenth century, particularly the early national period, the antebellum era (especially plantation culture and economy), slavery, the American Civil War (especially battlefront and home front experiences), Reconstruction, and the first half of the twentieth century. Politics and activism, religious experience and organization, and family relations are prominent themes throughout all the eras in the collections. The records housed in the SHC tell stories of thousands of individuals from all walks of life—farmers, planters, homemakers, enslaved persons, tradespeople, physicians, lawyers, journalists, educators, politicians, southern writers, and many more. The collection provides fertile ground for professional historians and other scholars, students, and the general public alike. And as users explore and revisit these sources, they will continue to provide others with a unique opportunity to understand and engage with the past.

With large-scale digitization users from all over the world could access this premiere collection easily, conveniently, and without expense.
Summary of the Plan for Large-Scale Digitization

The SHC’s plans for large-scale digitization will incorporate the following key features:

Digitization scope

The SHC staff will digitize many of the 4,600+ collections in their entirety.

Decision matrix

The first steps in the decision matrix take into account the provenance, access and use restrictions, and level of processing for each collection, as well as consider the possible legal and ethical concerns related to publishing the collection digitally. These steps have been applied to all 4,600 collections. To determine which collections to digitize and to set priorities for digitization, the SHC staff will use subsequent steps as a guide. Digitization will begin with collections containing materials that in some way address the African American experience and race relations in the American South.

Digitization on demand

Users who desire a digital version of a collection that was not slated for digitization, or that was not going to be digitized soon enough to suit their needs, will be able to request that the collection be digitized and will pay a modest fee to help offset the cost of the service.

Sustainable program

The SHC and the University Library are committed to a sustainable and continual digitization program based on the priorities set by the decision matrix. This commitment will require the consistent allocation of current staff time and resources as well as continual expansion of technological capacity and periodic hiring of additional staff to accommodate an ever-growing Digital SHC. Digitization will continue throughout each year, although the pace will be partially determined by the resources allotted through the annual budget and by the availability of external funding.

Digital search room

The digital search room experience will mirror the physical search room experience: Users will be able to interact with primary sources directly. And the digitized manuscripts will neither be framed with contextualizing or interpretive information, nor supplemented with resources such as scholarly essays, lesson plans, interactive features, or social networking communities.

Finding aids

The finding aids for the digital collections will be identical to the finding aids found on the SHC’s website. The finding aid for each collection will serve as the descriptive guide for both the digital collection and the physical collection, and the online finding aid for a collection will operate as the anchor for the digitized materials in the collection. Additional description or indexes will not accompany the digital versions of collections.

Organization of collections

The organization of the digital versions will mirror the organization of their physical counterparts. The digital version of each document will be a high-resolution image file. With the exception of oral history interview transcripts, there are no immediate plans
to use optical character recognition (OCR) systems or word-for-word transcription to enable full-text searchability.

Access to digitized material

Digitized material will be made available on the Internet free of charge. No enrollment in a subscription service or affiliation with a university or other institution will be required. Users of the Digital SHC will be able to download copies of the digitized documents to their own computers and will be able to make use of those documents including republishing them to their own websites.

Summary

All decisions about the digitization process were made by the SHC staff after investigating the digitization experiences of staff at other archival collections, after careful consideration of the SHC's archival mission, and, most critically, after extensive conversations with the primary audience for the SHC: the scholars of the history of the American South.

This final report addresses these features in detail, explains the thought process behind each decision, and provides a model that staff at other archives may find useful as they proceed with their own plans for digitization of their collections.

Although it is difficult to predict the directions in which innovations in technology and historical scholarship will lead, the decisions made at this early stage of the digitization process have left room for the possibility of different choices being made in the future, thus allowing the digital version of the SHC to continually remain flexible and responsive to the needs of its users.
The Challenges of Large-Scale Digitization

The SHC already has a strong online presence, in large part due to Documenting the American South (http://docsouth.unc.edu), an award-winning site that makes selected documents available alongside contextualizing documentation from archivists and scholars. Large-scale digitization will not replace established programs like Documenting the American South or other digital projects currently in development but, rather, will exist simultaneously, providing a new model for digital programs in the University Library. Previous digitization efforts of library materials have been project-focused. Large-scale digitization will incorporate entire collections encompassing millions of individual documents. The resulting digitization will be programmatic and sustainable, incorporated into the core mission of the SHC and continuing indefinitely.

This shift toward a large-scale model for digitization raised many critical questions for the SHC staff to consider. Would the patrons who use manuscript sources in the SHC want to use digital versions of those sources? If so, how would they use those sources? Which features would they want and need in an online Digital SHC? How comprehensive should the digitization plan be? Should digitization be undertaken on an item-by-item basis, or should collections be digitized in their entirety? Which collections should be prioritized for digitization? How should the digital versions of collections be formatted, organized, and presented online? How would the SHC staff incorporate digitization into the workflows already established for manuscript collections?

The SHC decided to seek grant monies to explore answers to these questions and to develop a plan for meeting the challenges of large-scale digitization, and subsequently received funding from the Andrew W. Mellon Foundation for the project.

The purpose of the grant project was two-fold: 1) to build, in cooperation with the scholarly community, a “decision matrix” for selecting and prioritizing the massive holdings of the SHC for large-scale digitization, and 2) to construct a sustainable model for a large-scale digitization program for manuscript collections that would serve as the beginning of a Digital SHC.
Summary of Grant Activities

Graduate Student Focus Groups, June 2007
Graduate Student Focus Group, August 2007
Interviews with Scholars, August 2007 – Spring 2008
Preparation for the Scholars Workshop
Workshop with Scholars of the American South, April 2008
Symposia with Archivists and Librarians, February and March 2009

To develop the decision matrix and the sustainable model for digitization, the SHC staff worked closely as a group and with the scholarly and archival communities to determine the best courses of action. Through focus groups with graduate students, interviews with individual scholars, attendance at professional conferences, site visits, a scholars' workshop, and two symposia with archivists and librarians, SHC staff gathered the information necessary to develop the digitization plan.

Graduate Student Focus Groups, June 2007

In June 2007 the SHC held two focus group meetings with history graduate students from local universities to gather information on the students' current use of digital history collections in research and teaching; projections for future use of digitized manuscript collections; and preferences and concerns about digitized manuscript collections. Graduate students were chosen for the initial focus groups for several reasons. SHC staff surmised that most graduate students would be comfortable with digital technology; are fluent with the Web; are familiar with many currently available digital history sources; and are likely to benefit, both in the short term and long term, from large-scale digitization of manuscript materials. The fifteen students who participated in the focus groups represented UNC-CH (13), Duke University (1), and the University of North Carolina at Greensboro (1). Most had completed their doctoral exams and were writing their dissertations, but some were at earlier stages in their graduate programs and one was a postdoctoral fellow. All fifteen were studying southern history, and their research interests were within the broad range of the SHC's holdings. The students' research topics spanned the eighteenth through the twentieth centuries, and intersected the subfields of social, cultural, military, and intellectual history. More than half of the students indicated that their work addressed issues of race, class, and gender.

Prior to holding the focus group sessions the SHC staff sent the students a list of discussion questions and links to existing digitized manuscript collections including those on the sites of the Library of Congress's American Memory and the National Library of Australia. Laura Clark Brown, director of the “Extending the Reach of Southern Sources” project, led the groups; and David Silkenat, the project's graduate research assistant (then a PhD candidate in the UNC-CH history department, now an assistant professor of history at North Dakota State University) participated in both sessions, served as secretary, and presented interface models.

Each session was approximately two hours and included an introduction to the holdings of the SHC and to the origins and purpose of the grant project, a presentation
on several models of manuscript digitization, and a structured discussion. Participants had the opportunity to answer each prepared question as well as ask their own questions. (See Questions Posed to Graduate Students in Appendix A.)

All participants had extensive experience with using manuscript materials for research purposes. Many had used digitized manuscript collections in teaching—introducing their students to primary sources using digital history sites such as Documenting the American South, Making of America, and American Memory—but few had used digitized manuscript collections to do their research. (A notable exception was one participant, who had conducted the majority of her dissertation research online by using a digitized collection of court papers and had found online research to be very convenient.) Those who had not used digital collections in their research said they had difficulty finding relevant online material, and generally found “boutique” sites of little use in research because these sites contained a limited selection of documents rather than full collections. Most participants said they would use the Digital SHC in their teaching.

Focus group participants all agreed that:

- Large-scale digitization of the SHC’s holdings would be of enormous value to scholars, students, and nonacademic users, and that virtual access to the SHC from anywhere in the world would expand research and further scholarship on the American South.
- An intuitive user interface with familiar search options would be essential.
- Digitized documents should clearly map back to the original physical manuscripts so that research conducted online could be seamlessly combined with and cross-referenced to research conducted in the physical UNC-CH search room.

Participants disagreed on other issues including:

- There was strong opposition to user fees, licensing, commercial publishing, and restricted or tiered access to the Digital SHC. (Although the students recognized that user fees and commercial publishing might help finance and sustain or hasten digitization efforts, they did not accept restricted access and user fees as reasonable trade-offs.)
- Value of digitizing collections that have already been microfilmed: Although the majority favored digitizing existing microfilm copies of collections, some were concerned that scans of microfilm would not produce the same kind of high-quality images as scans of original documents.
- Merits of incorporating user-generated index terms and Web 2.0 features such as wikis and social networking applications.

Graduate Student Focus Group, August 2007

The second group evaluated a potential model for presenting digitized manuscript collections online: the digital collections on the Smithsonian Institution’s Archives of American Art (SI/AAA) website (http://www.aaa.si.edu/). Eleven graduate students from the UNC-CH history department participated, all of whom were pursuing doctoral degrees and conducting research on southern history topics. Most had used the SHC
extensively in their research; and many had participated in SHC’s earlier series of focus
groups and so had some familiarity with the “Extending the Reach of Southern Sources”
project.

The SI/AAA website contains a growing body of digitized manuscript collections of
important American artists including Winslow Homer, Grant Wood, Romare Bearden,
and Alexander Calder. The SI/AAA’s mass digitization effort began in 2005 with a grant
from the Terra Foundation for American Art, and in the first year the project had
produced 14,769 digital files of manuscript material. The presentation of a digitized
manuscript collection is structured on and linked from the collection's online finding aid. The
site provides images (digital facsimiles) of the collection's documents. It does not offer
transcriptions, text encoding, description or indexing of individual documents, or full-text
search capability.

Participants sat at individual computer terminals, and Laura Clark Brown and
David Silkenat guided them through the collections using a series of structured
questions. Overall, participants had a positive reaction to the website and the
presentation of digitized manuscript collections. Significantly, participants did not
object to the finding aids serving as the digital collections' descriptive architecture;
many said they were comfortable with this model because they have extensive
experience and familiarity with finding aids.

Through this focus group, the SHC staff concluded that the SI/AAA's website could
serve as a good model of large-scale digitization of manuscript collections, and that,
with some design caveats, the staff could emulate and build on the SI/AAA's
innovations and accomplishments.

Interviews with Scholars, August 2007 – Spring 2008

Between August 2007 and spring 2008, SHC staff conducted interviews with scholars
of the American South—at the 2008 Annual Meeting of the American Historical
Association (AHA), by telephone, and at UNC-CH—to gauge interest in a Digital SHC
and to ascertain preferences for selection priorities and Web presentation.

In January 2008 project director Laura Clark Brown and project research assistant
David Silkenat participated in the annual meeting of the AHA held in Washington, D.C.
In addition to attending relevant conference sessions on digital history initiatives, Ms.
Brown and Mr. Silkenat conducted interviews. AHA interviewees were Edward L.
Ayers (University of Richmond), Peter Carmichael (West Virginia University), William
Link (University of Florida), Matthew Mason (Brigham Young University), Adam
Rothman (Georgetown University), and Philip Troutman (George Washington
University). Each interview lasted approximately twenty-five minutes.

Preceding and following the AHA conference, Ms. Brown and Mr. Silkenat conducted
telephone and in-person interviews with other scholars: Bruce Baker (Royal Holloway
University of London), William Blair (Pennsylvania State University), Judkin Browning
(Appalachian State University), Stephanie M. H. Camp (Rice University), James Cobb
(University of Georgia), Barbara Hahn (Texas Tech University), Christine Heyrman
(University of Delaware), James Leloudis (UNC–Chapel Hill), Tracy K'Meyer
(University of Louiville), Michael O'Brien (University of Cambridge), Paul Quigley
(University of Edinburgh), John Michael Vlach (George Washington University), and
Jeffrey Young (Georgia State University).
During the interviews, which were informal and conversational, all scholars were asked the same questions; and approximately the same amount of time was spent with each. (See Key Interview Questions with Scholars in Appendix B.)

Nineteen interviewees were self-selected from an extensive list of scholars whom project staff had asked to participate. Attempts were made to choose those who would represent a diversity of career status, geography, gender, race, and university tiers. Although most who agreed to participate were white males from Tier 1 research universities, the final group of interviewees represented scholars from different parts of the United States and the United Kingdom who were at various stages in their careers. Interviewees also represented a broad variety of chronological, topical, and methodological interests in southern history. (See further identification of all interviewees in the List of Scholars Interviewed in Appendix I).

The interviews yielded useful information about current and projected uses of digitized manuscripts, and provided the basis of a strategy for engaging the scholars who would attend the April 2008 workshop. Interviews were structured around five main avenues of inquiry: 1) the scholars' current uses of digital history collections in teaching and research, and projections of use of the Digital SHC; 2) the SHC's use of finding aids to describe and anchor the digital versions of the manuscript collections without any additional item-level description and indexes; 3) the scholars' preferences for the interface and display of the documents in the Digital SHC; 4) the scholars' interest in Web 2.0 features such as social tagging, bookmarking, wikis, and personal accounts; and 5) the scholars' guidance in prioritizing the 4,600 collections for digitization.

The majority of scholars indicated that they used digital history collections, although several indicated they used these collections more in their teaching than in their original research. Those scholars who used digital sources indicated that their physical distance from archival sources necessitated their use. The most popular sites were the Documenting the American South site at UNC-CH, the American Memory site at the Library of Congress (http://memory.loc.gov/ammem/index.html), and the Valley of the Shadow site at the University of Virginia (http://valley.vcdh.virginia.edu/).

Nearly all interviewees were enthusiastic about the prospect of a Digital SHC and projected they would use it in both teaching and research. Comments include:

“Having the unequaled collections of the SHC available in digital form would revolutionize the study of southern history. I could also use archival material in my teaching in entirely new ways.”

“I am really excited about the prospect of my students (at both the undergraduate and graduate levels) being able to access the unparalleled collections of the SHC from afar.”

“The great value [of digitizing the SHC] is that it would make collections available to junior scholars [who do not have the resources of senior scholars].”

Regarding the possibility of requiring user fees: As in the graduate student focus groups, the majority indicated that user fees were an unacceptable option. Those interviewees who accepted the possibility of user fees did so with reluctance. Several noted that although they might be willing to pay a modest fee for their own research, they would not ask or require their students to pay for use, thereby effectively negating
the potential for using the Digital SHC in teaching. Any user fee, they said, would dramatically reduce the use of the site.

In considering finding aids and browsing options for the digital collections, the broad consensus among interviewees was that the Digital SHC should replicate the physical arrangement of manuscript collections as closely as possible so that each digital facsimile mapped directly to the physical document it represented. Interviewees agreed that a familiar, easy-to-use search should be a priority, and some hoped that the Digital SHC would have advanced search functions enabling a search for specific types of documents.

As for the interface and document display, interviewees said they wanted the documents to be legible and site navigation to be fast and intuitive, but expressed few other preferences. Several said they would like to be able to simultaneously view multiple documents or multiple pages of a single document. Others noted that on the websites for many digital collections the Save and Print options are often hard to use, and they requested that the Digital SHC provide a logical and user-friendly method for these functions.

The question of incorporating Web 2.0 features (such as personal accounts, social tagging, bookmarking, and wikis) elicited the widest range of opinions. Several scholars, particularly those who rarely used digital sources, indicated they did not see any value in Web 2.0 features. Many others, however, said they would be interested in having some form of personal account on the SHC site that would allow them to track which documents or collections they had examined and that would allow them to save documents for reexamination. Most interviewees were indifferent to the inclusion of user-generated content such as social tagging and wikis. Those in favor said that such content could foster scholarly discourse and promote social interaction among scholars. Those opposed were concerned that such content could be distracting or inaccurate; but if the content was included, they said they hoped that user comments would be moderated by SHC staff.

Interviewed scholars were asked to suggest criteria for prioritizing the collections for digitization. The general consensus was that the highest use and the richest collections—“greatest hits”—should be digitized first. Given the uneven methods of early collecting practices and the SHC's vast holdings of affluent whites, one scholar warned that starting with the “greatest hits” might invite criticism for excluding historically marginalized groups. As with the graduate student focus groups, interviewees were divided on whether the SHC should digitize microfilm copies of collections.

An area of broad agreement was in the idea of digitizing everything in a given collection—even the seemingly insignificant. The majority of interviewees were not averse to partial digitization of collections when issues such as copyright, privacy concerns, or donor restrictions prevented full digitization; but they wanted a clear explanation on the site about the reasons for omission. Ultimately, interviewees wanted to be able to see as much of the collection as possible, and wanted to make their own decisions about which individual documents were valuable or relevant for their research. The majority were excited about the advent of the Digital SHC. Several asked, “When will it be available?”
Preparation for the Scholars Workshop

From January through early April 2008, SHC project staff concentrated on preparations for the “Southern Sources: Focusing the Conversation—A Workshop with Scholars of the American South.” Laura Clark Brown met regularly with the co-principal investigators (Tim West, SHC director / curator of manuscripts, and Kirill Fesenko, director of the Carolina Digital Library and Archive at UNC-CH), and together they conceptualized the workshop and desired outcomes, developed the agenda, and compiled the list of participants. (See the Agenda for the Southern Sources workshop in Appendix C.)

To provide the scholars and other participants with background prior to arrival, Ms. Brown and Mr. Silkenat prepared a report, which was distributed in late March 2008 to all workshop participants. The report outlined the grant project and the SHC’s goals for the Digital SHC, explained the premise of large-scale digitization in contrast to digital publication of selected items, and summarized the findings of the graduate student focus groups conducted in the summer of 2007 and the interviews with scholars conducted in the winter.

Ms. Brown then compiled materials to be used during the workshop including a workbook, which contained background and contextual information; hard copies of slide presentations; and Web-interface features for consideration. She also created an exit survey to be used for gathering scholars’ preferences on selection priorities and interface features. (See the Exit Survey Data in Appendix D.)

Workshop with Scholars of the American South, April 2008: “Southern Sources: Focusing the Conversation”

The Southern Sources workshop opened with a keynote address by Daniel J. Cohen, an associate professor in the department of history and art history at George Mason University and director of the Center for History and New Media. Dr. Cohen’s speech exemplified some of his many technological and historical interests, knowledge, and skills.

The first session was a panel discussion on scholarship on the American South, moderated by Jacquelyn Dowd Hall, the Julia Cherry Spruill Professor and director of the Southern Oral History Program at UNC-CH. Panelists Laura F. Edwards (Duke University), Glenda Gilmore (Yale University), Adriane Lentz-Smith (Duke University), and Mark M. Smith (University of South Carolina) addressed: 1) the current and emerging trends in southern history that they found most exciting, 2) trends that concerned them in some way, 3) future developments that they envisioned in their own branches of southern history, 4) the future role of paper archives in their fields, and 5) the potential of large-scale digitization of manuscripts in the SHC to enhance research in their areas of interest. A Q&A session followed.

In the next session, “Where Do We Start?: Prioritizing Collections for the Digital SHC,” Laura Clark Brown provided possible models for setting digitization priorities for the SHC’s collections. The models suggested approaches for prioritization such as a “greatest hits” strategy that would target the most frequently and heavily used collections and a “hidden treasures” strategy that would focus on little-used but rich collections identified by SHC staff. Ms. Brown also introduced a preliminary decision matrix—a complex outline of questions, considerations, and criteria—for selecting and
prioritizing the digitization of the individual collections. Both the prioritization models and preliminary decision matrix were based on findings from the focus groups with graduate students and the interviews with scholars. In addition, she discussed potential impediments to collection digitization, offered possible solutions for meeting these challenges, and facilitated a discussion on all these aspects of digitization.

In “Manuscripts at Midnight: Finding Aids and Interface Models for Digitized Manuscript Collections,” archivists Jackie Dean, SHC manuscripts processing librarian, and Maggie Dickson, SHC Watson-Brown Project librarian, described the architecture and interface for digitized manuscript collections. Ms. Dean explained plans for a collection's finding aid to serve as the foundation for the digitized collection, with the description contained within the finding aid anchoring the digital facsimiles of materials in that collection. Ms. Dickson showed a variety of interface models for digitized manuscripts, demonstrating how users would access bundled documents, and how they would view and manipulate individual documents. Ms. Dean closed the session by proposing a new search and retrieval system for online finding aids and inviting discussion from workshop participants.

On the second day of the workshop, Christopher (Cal) Lee, assistant professor in the UNC-CH School of Information and Library Science, presented “Beyond Fast Pictures: Exploring the Potential of Interactive Digitized Collections.” He explored ways in which digital access to primary sources can not only transcend geographical boundaries but also offer opportunities for new ways to interact with sources. In an online environment, users can provide comments, annotations, and short descriptive labels (tags) to items; collaboratively author and edit documents associated with collections; customize the interface to best meet their own needs; and form and navigate social networks based on their preferences and research interests. Dr. Lee showed some Internet sites that make use of these features.

Later, project staff administered the Exit Survey inviting the scholars to identify their preferences for the Digital SHC, with emphasis on collection priorities, interface, and special features. The final portion of the workshop included a discussion of the remaining steps to be taken during the grant period.

The scholars who attended the workshop provided invaluable insights and posed questions that were tremendously helpful in the development of the SHC's plan for sustainable large-scale digitization. (See “Key Features of the Digitization Program” for more about the information SHC staff gathered from scholars who attended the workshop.)

**Symposia with Archivists and Librarians, February and March 2009**

Toward the end of 2008, with selection of collections and data gathering under way, the focus of the project shifted toward developing a sustainable digitization program. Ms. Brown recommended that the SHC host two one-day symposia for all University Library staff involved in digitization and colleagues across the country. Richard Szary, director of Wilson Library / associate university librarian for Special Collections at UNC-CH, and co-principal investigators Tim West and Kirill Fesenko worked with Ms. Brown to conceptualize the structure and desired outcomes of the symposia. (See agendas for both symposia in Appendices E and F.)
Legal and Ethical Implications of Large-Scale Digitization of Manuscript Collections

The first symposium was held on February 12, 2009. Ms. Brown presented a brief report on “Extending the Reach of Southern Sources.” Aprille Cooke McKay (University of Michigan), a lawyer and archivist, presented on the legal and ethical obligations associated with large-scale digitization. She focused on the code of ethics of the Society of American Archivists (SAA), which requires that archivists protect the privacy rights of both donors and the individuals/groups who are the subjects of archival records, as well as deal with the ethical obligation to provide access to manuscripts sources for research purposes. Ms. McKay also addressed the legal duties of archivists, pointing out that although archives may be legally liable for breaches of agreements with donors, it would be difficult to find paper archives liable for invasion of privacy because: 1) the archival materials are obscure and not widely distributed; 2) the archive does not typically publish the materials directly; and 3) any legal liability would likely rest with the journalists or scholars making use of the material.

Digital manuscripts, she explained, present very different legal concerns because with digitization the materials become much more widely available, and with broad access any private information that is disclosed can reach a much wider audience and therefore can cause more harm. In digital collections archives also become publishers of the material, making it more likely that they could be held legally liable for the publication of private facts. Ms. McKay also addressed issues of privacy, defamation, libel, and false light—providing examples of how legal issues have been handled by archives that have published on the Internet manuscript materials containing information about living people. She closed by suggesting strategies for managing legal risk including careful selection of materials for digital publication, the possibility of getting input from third parties mentioned in archival collections, creating policies for how to handle potential complaints, and tailoring access to sensitive collections.

Next Nancy Kaiser, SHC project archivist, and Matthew Turi, SHC manuscripts reference librarian, moderated a discussion of case studies that involved third-party privacy issues. Participants were invited to discuss three specific collections and to consider how the SHC might approach ethical questions regarding collections that contain sensitive materials.

The first case study, the Delta Health Center Records, focused on medical privacy issues. [Although the SHC is not a covered entity under the Health Insurance Portability and Accountability Act (HIPAA), the ethical concerns with providing access to health information continue to be at issue.] The Delta Health Center, originally affiliated with Tufts University and later with the State University of New York at Stonybrook, was located in rural Mound Bayou, Bolivar County, Miss., and served Bolivar, Coahoma, Sunflower, and Washington counties, where poverty remains widespread. Records from the center include correspondence, subject files, and case studies. Among the case studies are some compiled by medical students from northeastern universities who conducted health and living assessments of low-income, African American families living in the Mississippi Delta in 1968. The family case studies include identifying information—e.g., names, birth dates, locations—that can be redacted prior to use in the search room or to digitization. However, rich narrative elements of these studies reveal intimate details of individuals' mental and physical health, diet, living conditions, private lives, familial relationships, and criminal activity. Given the creation date of the records, it is possible that current or former residents of
these small communities could identify the families described in the narratives. Yet, making these family case studies available could be highly valuable, in particular to some scholars and public health professionals.

The second case study focused on personnel issues—specifically, employee evaluations in the Southerners for Economic Justice (SEJ) Records. SEJ was founded in 1976 during a successful campaign to help textile workers at J.P. Stevens and Co. to unionize. Since then, SEJ has focused on empowerment of the unemployed and working poor to develop community-based strategies to solve social problems associated with economic crisis. The records, dating from 1977 to 2001, document the administrative and programmatic work and related interests of the organization. Identified sensitive materials in this collection include personnel files. The personnel files are found chiefly with the administrative materials, but personnel information is also interspersed throughout the programmatic files. For example, the personnel review of a particular employee may be found in that employee’s personnel file, in the supervisor’s personnel file, in loose papers, and in programmatic files. It would be impossible to segregate these materials without an item by item review of the entire extensive collection. These sensitive materials are valuable as documentation of the culture of nonprofit work, especially in a grassroots organization.

The third case study, the diaries contained in the Joseph A. Herzenberg Papers, focused on digitization of intimate information contained within manuscript sources. Joe Herzenberg was a politician, a historian, an advocate for social justice, and the first openly gay elected official in North Carolina. He died in October 2007 in Chapel Hill. The collection contains diaries, correspondence, subject files, photographs, and other materials relating to Joe Herzenberg. The entries in the diaries begin in 1954 during Herzenberg’s early adolescence and continue through 2006, and provide detailed accounts of Herzenberg’s personal and professional life. He recounts his struggles with weight loss and depression, and he discusses his political activities from early work in groups like CORE (Congress of Racial Equality) to his campaigns for town council. In addition, Herzenberg documents his sexual encounters and alludes to some of his friends’ sexual relationships and illegal activities.

While exploring these case studies, workshop participants considered the following questions:

1. Most SHC collections are unrestricted for both research and duplication in the SHC’s search room. In that relatively controlled environment, the SHC transfers the responsibilities for the use of sensitive materials to the researcher. Can the SHC do the same in the Web environment?

2. Does the Web environment change what archivists consider to be sensitive? Should it?

3. Are some material types (e.g., text, images) more sensitive than other types? Is some content (e.g., health, personnel, intimate details) more sensitive than other content?

4. Should it matter who the third party is? Should ethical considerations change when the third party is a public or well-known personality? Should ethical considerations change when the third parties are from historically marginalized or disadvantaged groups?
5. At what archival level—collection, series, item—do archivists need to apply ethics?

6. Are contemporary collections too fraught with ethical dilemmas to digitize?

Most participants agreed that these collections all included materials that would be of immense value to scholars of the twentieth-century American South, but that they are so recent that it would be prudent to wait to digitize and publish them until more time had elapsed. However, making these kinds of materials available in the SHC search room but not in an online collection raises ethical concerns about equality of access. And if such collections were set aside for long periods, would the Digital SHC meet the needs of scholars of the twentieth-century South as well as it does those studying earlier periods?

The next session was a panel discussion on reconciling modern archival practices with large-scale digitization. Panelists included Barbara Aikens (Smithsonian Institution), Max J. Evans (Church of Jesus Christ of Latter-day Saints), Tom Hyry (Yale University), Bill Landis (Yale University) and Dan Santamaria (Princeton University). Merrilee Proffitt, senior program officer at OCLC Programs and Research, was moderator. Topics discussed were: archival practices in the processing of collections in preparation for digitization, digitization on demand, the increased exposure of private data through full-text searchability, the need to work with donors to reduce risk, and ethical obligations in making collections accessible to the general public.

Archivist Maggie Dickson presented “Due Diligence, Futile Effort: Pursuing Copyright Holders,” describing copyright issues in the digitization of the Thomas E. Watson Papers. The Watson-Brown Project at the SHC is a two-year, privately funded grant. One aim of the grant is to digitize and publish online the correspondence series of the Thomas E. Watson Papers. The series consists of approximately 8,500 letters, postcards, telegrams, and notes written between 1873 and 1986 by Watson and his family, friends, and political and business colleagues, with the bulk of the materials dating from the 1880s and 1920s. Unpublished manuscript materials, such as those found in the Thomas E. Watson Papers, are protected by copyright for 70 years following the death of the author.

To investigate the copyright component of digitizing archival materials, the Watson-Brown Project staff conducted intensive copyright research on the correspondence in the Watson Papers by gathering basic metadata—e.g., names, dates, and geographical locations—from the materials. There were 3,304 names included in the correspondence. Using a variety of sources—including Wikipedia, the Social Security Death Index, Ancestry.com, and print reference works such as biographical dictionaries—staff attempted to identify the correspondents to find dates of death for determining copyright status. They 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 in copyright. Life dates could not be found for 1,571 correspondents (47.55%), and no information could be found for 24 correspondents (.73%). The identification process undertaken by the Watson-Brown Project staff was very time-consuming, requiring more than 14 weeks of dedicated time by a full-time employee to evaluate a relatively small body of materials (when compared with the millions of documents in the SHC).

Symposia participants agreed that this would be an untenable methodology of copyright research for the SHC. Furthermore, archivists have no reasonable means to
contact the thousands of descendants of thousands of people. For all practical purposes, it would be impossible to secure copyright permissions for every correspondent whose writing appears in the SHC. Clearly, the SHC staff needed to find a different way to reconcile copyright law with the repository’s mission and the needs of its researchers.

The next panel on legal issues was “Orphan Works, Fair-Use, and Risk Management,” with panelists Heather Briston (University of Oregon), Sharon E. Farb (UCLA), Peter Hirtle (Cornell University), Bill Maher (University of Illinois at Champaign-Urbana), and Mary Minow (librarylaw.com). The panelists, three of whom are both lawyers and information professionals, considered issues of fair use such as whether a digitized item may be considered sufficiently “transformed” by the context in which it is placed within the digital collection to meet the legal requirements of fair use. They also discussed the definition of publication, due diligence, the management of risk in terms of copyright case law, the potential for changes in copyright law, and the importance of documenting efforts to investigate rights to materials as part of risk management. (For more detailed information on participants, see List of Grant Advisors and Symposia Participants in Appendix K.)

Moving from Projects to a Program: The Sustainability of Large-Scale Digitization of Manuscript Collections

The second symposium, held March 12, 2009, focused on the sustainability of large-scale digitization of manuscripts collections. Laura Clark Brown and Maggie Dickinson presented reports from concurrent digitization projects at the SHC. Two panel discussions followed: “Business Models for Large-Scale Digitization Programs,” presided over by Christopher (Cal) Lee, and “Sustaining Large Scale Digitization Programs,” moderated by Will Owen, head of the University Library's Systems Department (UNC-CH). Panelists Liz Bishoff (Bibliographical Center for Research), Oya Rieger (Cornell University), and John Wilkin (University of Michigan) participated in both.

At the “Workflows for Large-Scale Digitization at the Archives of American Art” session, co-presenters (all from Archives of American Art, Smithsonian Institution), were archivists Barbara Aikens and Karen Weiss, and programmer Toby Reiter. Ms. Aikens and Ms. Weiss described the origins of the organization’s digitization program, the shift from microfilming to scanning, and the impact of digitization on archival processing. Mr. Reiter explained the technological infrastructure for the program and demonstrated features of the user interface. (For more detailed information on participants, see List of Grant Advisors and Symposia Participants in Appendix K.)

Symposium Conclusions

The SHC staff gleaned a great deal of valuable information from the panelists and the lively discussions that followed each portion of the program. The three most critical conclusions were: 1) For large-scale digitization to become sustainable, the goals of the digitization program must be integrated with the University Library's broader institutional mission; 2) Responsibilities for meeting the digitization program's goals should be distributed among a wide range of staff members, just as are responsibilities for meeting other parts of the library's core mission; 3) Funding for digitization must become a permanent part of the library's budget.
Key Features of the Digitization Program

Selection Process

The SHC staff will digitize many of the 4,600+ collections in their entirety.

The digitization and online presentation of manuscript collections comprised of millions of documents present staggering challenges of scale. The SHC staff believes these challenges will best be addressed by applying archival theory and practice in the digital environment. In digitizing the collection, the SHC staff will employ the archival principle of provenance: organizing and maintaining the individual collections based on the origins of the materials, rather than piecing together new collections of selected documents based on topics, geography, or chronology, or other characteristics. Under this model, for example, if the topic to be digitized was “Slavery in North Carolina,” the staff would digitize entire collections in the antebellum plantation holdings (such as the Cameron Family Papers and the Pettigrew Family Papers) that may pertain in part to slavery in North Carolina—as opposed to digitizing a selection of individual manuscripts pulled from these different collections documenting slavery in North Carolina. This structure mirrors the physical organization of the SHC's holdings in the Wilson Library where items remain within their collection of origin.

Feedback from graduate students and scholars

In exploring the Archives of American Art’s website, graduate students in the August 2007 focus group discovered that portions of the collection had not been digitized—but that didn’t become apparent until they accessed the box or folder level. They thought that an explanation for omissions should have been included upfront in the collection’s finding aid.

Scholars who attended the Southern Sources workshop echoed these sentiments. One said, “I do not like sites that select ‘important’ documents.” Another stated emphatically that even if it’s a box of tiny receipts, “do not sample. Digitize it all.” One scholar observed:

“I don't think you should skip over [collections] simply because you have to omit portions of it. After all, the extent of any collection is determined by arbitrary factors, and in most cases the remaining portions of the collection will be useful on their own. But you should definitely explain why you have had to leave certain sections out. This would be especially important to researchers who might want to come in person to fill in gaps.”

In the Exit Survey following the workshop, only a handful of the scholars (4 of 21) said that the SHC should prioritize digitizing specific genres within collections; the other 17 preferred that complete collections be digitized. Regarding those collections that cannot be fully reproduced due to copyright, privacy, or other restrictions: the majority of scholars (14 of 21) responded that the SHC should not prioritize those
collections that can be digitized in their entirety.

In summary, this feedback indicated that, when possible, the SHC should digitize complete collections; but when some content needs to be omitted from digitization of a collection, that collection should remain in the digitization queue, and the finding aid should include reasons for any omissions.

Order of Digitization

The first steps in the decision matrix take into account the provenance, access and use restrictions, and level of processing for each collection, as well as consider the possible legal and ethical concerns related to publishing the collection digitally. These steps have been applied to all 4,600 collections. To determine which collections to digitize and to set priorities for digitization, the SHC staff will use subsequent steps as a guide. Digitization will begin with collections containing materials that in some way address the African American experience and race relations in the American South.

A critical question facing the SHC staff at the outset of this digitization effort was, “Where to begin?” With a collection as large and as broad as the SHC, full digitization of all the materials would not happen quickly, and so priorities must be set to determine the order. SHC staff considered several options and then presented them to the scholarly community (the SHC’s core audience of users) to help determine what would best meet their needs.

Chronological approach

One option was to digitize collections in the order in which they were received—either beginning with Box 1 of Collection 1 and working methodically toward the most recent collections, or, beginning with Collection \( n \) (the most recently accessioned collections) and working backward toward Collection 1. Both approaches, however, present some inequities: If the Box 1 of Collection 1 approach was used, then some rarely used collections would be digitized before collections that are used frequently. And records from early time periods would be digitized before twentieth-century collections, given that the SHC’s first 2,000 accessions were mostly nineteenth-century records, and scholars exploring the twentieth century may not find much of use in the digital search room. If the Collection \( n \) through Collection 1 approach was used, the bulk of the early digitization would be focused on large twentieth-century collections, and scholars exploring the eighteenth and nineteenth centuries would wait indefinitely for useful digital collections to be put online.
As one scholar commented in the workshop, “Chronological would not work for me because you would inevitably pick the chronological period that I am not working in . . . and that would be frustrating.” Another scholar said that twentieth-century collections tended to be “thin, but impossibly gigantic and so being able to search them in new ways and different ways becomes more and more critical.”

If digitization could enable scholars exploring the twentieth century to navigate those voluminous collections more easily and conveniently, then an approach that placed all such collections near the end of the digitization queue would clearly be inappropriate.

Greatest hits approach

A second option was the “greatest hits” model of digitization— to begin with some of the most-used and most written-about collections (which tend to be the most widely used). There are problems with this approach as well including:

1. The greatest hits model isn’t always preferable or inclusive: As one scholar noted, “We’ve all seen those greatest hits collections. I’ve never seen one that really covers the full breadth that we are talking about here.”

2. Many of these collections have already been microfilmed—and so are available worldwide through interlibrary loan: These resources would certainly be well used if digitized early in the process. However, in prioritizing their digitization, collections that are not available on microfilm—and therefore can be accessed only by users who visit Wilson Library—may not be digitized as quickly. One participant commented, “If I can get it on interlibrary loan, I don’t feel that I need to have it digitized right now. What I would rather see is the stuff that I have no way of accessing except to come here.”

3. Lesser-used and newly added collections could also be of great research value and could very likely become more readily used if digitized: An informal survey of archivists on the SHC staff revealed that usage statistics should play some role in the prioritization for digitization, but that heavy usage should not automatically move a collection ahead in the queue.

Hidden treasures approach

The third approach was to prioritize the digitization of some of the “hidden treasures”—those resources that the archivists may be familiar with, but that few scholars have yet studied extensively. However, what one person considers a hidden treasure, another person may consider only marginally useful or interesting. As one scholar asked, “I want to see them, but how do you pick them?” Moreover, many treasures may still be hidden in boxes little used or unknown to the current SHC staff, and could be revealed as they are brought out of the stacks to be digitized.

Conclusions

As discussions with scholars about these approaches unfolded, no consensus emerged to guide the SHC staff in choosing. Scholars tended to agree that no single research agenda, e.g., the American Civil War, should drive the selection of collections for digitization (as such an approach would necessarily exclude material of great use to scholars with other research interests), and that the materials prioritized for digitization should have broad potential for use by a variety of scholars and other
potential audiences. As one scholar noted, “If you’re doing it because you want to benefit a huge number of people who could never come to the Southern [SHC] and would never pick up a letter, then I think you ought to be biased toward the collections that could usefully be used by that massive number of users” [rather than focusing on small collections, which may be helpful to, say, a scholar writing a monograph, but that few others would find useful.]

Ultimately the SHC staff decided that by prioritizing collections that in some way addressed a very broad and important issue in southern history—one that was historically relevant in all time periods and locations—the initial digitization would represent a broad geographic, chronological, and topical range, covering both the most-used collections and the hidden treasures of the SHC.

Determining the Core Issue

Every archive has particular key issues that are central to its holdings and that transcend the geographic and chronological breadth of its collections. For southern history, perhaps no issue is more critical to an understanding of the region’s past, or is more actively studied, than the issue of race relations between white southerners and African Americans. In the Exit Survey, one scholar who advocated beginning with the digitization of older, handwritten documents recommended prioritizing “collections that include information about or by African American / enslaved people + within the chronological framework.”

Focusing on the African American experience for the initial digitization will tie in with the SHC’s collecting initiatives, and with ongoing projects that involve uncovering documentation of African Americans in existing collections and promoting use of those collections. These initiatives include rewriting the finding aids for ten of the SHC’s collections with substantial African American content and a project to publish an interactive guide to African American holdings in the SHC directed by Holly Smith, the Overholser Fellow in African American Studies.

Concentrating on collections related to race does not mean, however, that only those collections with a primary focus on race will be digitized. Rather, the SHC staff cast a very broad net in determining which collections include some content focused on this issue, and decided that any content based on race within a given collection merited inclusion of the entire collection.

After surveying the SHC’s collections, it was determined that 1,030 collections will be included using race-related content to prioritize collections. The prioritized collections demonstrate the enormous variety of the SHC, containing materials as diverse as plantation and farm journals; physician account books; cotton mill records; letters from federal and Confederate soldiers; diaries of ministers, women, social activists, and children; church records; household accounts; store ledgers; and sociologists’ field notes.

Although a collection may be selected for digitization on the basis of its content related to race, it may also be tremendously valuable to scholars studying unrelated topics, thus extending its value far beyond the considerable number of scholars interested in southern race relations. For example, the Cameron Family Papers—one of the most well-known and widely used collections of plantation manuscripts in the SHC—would clearly merit inclusion in the initial digitization effort using race-related content as a criterion because of its many documents related to slavery. However, the Cameron Family Papers have been used not only by numerous researchers interested in
American slavery, but also by scholars studying many other issues including gender in the Old South, parenting practices among North Carolina planters, folk medicine, the Civil War, the nineteenth-century global economy, the North Carolina railroad, southern architecture, literacy, and Jews in the American South.

**Decision Matrix**

With a strategy in place for priority setting, project director Laura Clark Brown built a detailed decision matrix to analyze the holdings of the SHC. The decision matrix is a complex series of questions, the first of which have already been applied to each of the SHC's 4,600 existing manuscript collections and will be applied to its future holdings. (See a complete copy of the Decision Matrix in Appendix G.) Priority will be given to collections that contain any documents addressing the “great subtext of American history” and of southern history in particular—race and race relations. The core question of the matrix deals with the presence of documentation on race. Subsequent questions explore the collections for size, formats, donor restrictions, copyright concerns, presence or absence of sensitive information on third parties, and status of the finding aids—all factors that might hasten, delay, or prevent digitization. Questions in the matrix determine each collection's place in the lineup of the digitization process.

A collection's priority ranking is based, in part, on the presence of documentation concerning race and race relations, broadly defined. However, the intent of the large-scale digitization model is to digitize the entirety of the prioritized collections rather than to extract those singular items focused on race. For example, the Graves Family Papers, a collection that documents two generations of an affluent white family in Georgia from 1815 to 1901, would be digitized in its entirety even though this large collection contains only a small pocket of documentation of freedpeople. The result of this approach—entire collections rather than selections—is broad coverage of geography, chronology, material types, creators, and content. In the Exit Survey this option was preferred by the scholars who rejected methods that gave priority to such elements as time periods, geographical locations, or document types (e.g., typescripts with OCR possibilities).

**Matrix stages**

The first stage determines whether the collection will proceed further through the matrix and be considered for inclusion in the initial digitization. Collections are first evaluated to determine whether the majority—or at least substantive portions—of the collection is comprised of original materials held in the SHC, as opposed to copies. Collections with original materials are then assessed to determine whether they contain any materials related to race or race relations in the American South. When a collection has met both initial criteria, additional questions are applied and scored to determine the collection's priority for digitization including possible impediments (e.g., imposed restrictions by collection donors, copyright protection, and privacy concerns), the extent to which the collection has been used by scholars or is deemed by archivists to be of great potential use to scholars, and the extent to which the collection has been completely processed by SHC staff.

In the second stage, each collection is analyzed in a more detailed way, considering: 1) any donor restrictions on the collection, 2) the existence of sensitive materials within the collection or those that may pose privacy concerns, 3) the provenance of the collection, 4) the inclusion of documents related to subject strengths and collecting
initiatives of the SHC within the collection, 5) the relationship between the collection and other collections within the SHC and other nearby archives, 6) the past and anticipated uses of the collection by scholars, 7) the copyright and intellectual property status of the collection, 8) the existence of microfilmed or other digital copies of the collection’s contents, and 9) the completeness of the collection’s processing and finding aids.

The third stage of the matrix deals with quantitative data about each collection in the priority list. Questions address the size of the collection, the chronological and geographical scope of the collection, the types of manuscripts included (e.g., correspondence, diaries, ledgers, maps, drawings, photographs, recordings), the dates and processing status of the original accession and any additions or expected additions, and whether any part of the collection has been previously reproduced (either on microfilm or in digital form).

Testing the matrix

By applying the three stages of the matrix collections can be prioritized for digitization to best meet the needs of the scholarly community in a balanced and efficient way. SHC staff applied the first stage of the matrix to all 4,600 collections, which enabled prioritization of 1,030 collections for digitization. The staff then tested all stages of the decision matrix by thoroughly evaluating two twentieth-century collections and determining the priority for each: The Delta Health Center Records (#4613) would be a low priority due to privacy concerns (presence of health information). The Delta and Providence Farm Papers (#3474) would be a high priority, as risk would be low for privacy concerns because, despite the fact that the Farm Papers is a twentieth-century collection, it does not contain sensitive information about identified third parties likely to be living today.

Previously microfilmed collections

Microfilm can be digitized quickly and cheaply without exposing original manuscript sources to additional handling. Thus digitizing collections from microfilm could result in a large volume of materials being digitized safely and made available online very quickly. However, scans of microfilm are only as readable as the original microfilm itself and may not be of the same high quality as scans created directly from the original documents. The graduate students and scholars were divided on the issue of digitizing materials for which microfilm copies already exist.

Many of the graduate students said that the SHC should digitize its existing microfilmed collections to get a large quantity of material online quickly; others were more wary of digitized microfilm—stating concerns about image quality and readability, and questioning whether the cost savings and speed truly mitigated the loss of quality. Some said that the degree of image-quality loss was substantial enough to forego microfilm digitization altogether, and they advocated for original manuscript digitization only. Others argued that any initially digitized microfilm could be replaced by digitized manuscripts at a later date as resources permitted.

When presented with the probability that the digital surrogates of the documents would be legible, most interviewed scholars said that large-scale digitization should emphasize quantity of materials digitized rather than perfection of image quality. Several argued that the SHC staff should examine microfilm editions for quality—and advocated for digitizing only those microfilmed collections with superior legibility and
for using the original archival material when the microfilm proved illegible or otherwise poor. The majority said that microfilm digitization should be considered on a case-by-case basis. More than half (12 of 21) of the scholars who completed the Exit Survey responded that the SHC should not digitize microfilm holdings.

Given the current availability of SHC staff time and resources, it was decided that when collections that are available on microfilm reach the top of the priority list for digitization, working from the original documents instead of the microfilmed versions will likely be best for the following reasons:

1. Many of the microfilmed copies of SHC collections were created by commercial publishing companies and therefore could not be digitized by the SHC.6

2. Microfilmed copies created by the SHC would have to be compared to original sources to determine image quality and may require a page-by-page assessment of each collection (as microfilm quality varies with different kinds of documents). And after scanning the film, individual pages or folders within a linear section of microfilm may have to be delineated in order to fit correctly into the digital presentation of the collection.

3. Many collections were microfilmed decades ago, and the original collections may have grown with additions, been reorganized and reprocessed, or changed in other ways. The microfilmed documents would therefore have to be aligned with the finding aid and the original documents as they are currently organized.

Legal and ethical concerns

The SHC must also find a way to balance the critical goal of providing greater access to manuscripts materials through digitization with the legal and ethical concerns regarding privacy and copyright. As mentioned previously, some collections may contain sensitive materials. Examples include: student records, correspondence related to refereed journals or grants, financial information, medical or health information, legal records, business records with trade secrets, materials conveying personal information about identified third parties (e.g., extramarital affairs, drug use, and juvenile crime), and sexually explicit and graphic content.

Federal law provides some guidelines in handling privacy concerns. For example, the Census Bureau releases aggregate data about the census as soon afterward as is practical; it does not, by law, release individual census responses for 72 years following the census due to privacy concerns. Following this guideline, collections whose materials were created 72 or more years ago should be free of legal concerns about third-party privacy; however, ethical obligations need to be taken into consideration as well. As a result, in the decision matrix higher priority is assigned to those collections that do not include materials that may infringe on the privacy of living third parties, as well as those in which such materials could be easily segregated from the bulk of the collection and could be excluded without significantly decreasing the collection’s research value.

The University Library and the SHC respect the intellectual property rights of others and do not claim any copyright interest in most SHC collections. The SHC will make the digital reproductions of archival materials available under an assertion of fair use (17 USC 107). However, staff will adhere to a take-down policy that will guide decision making in the wake of copyright infringement claims. The take-down policy
will be published on the SHC’s website. Higher priority will be given to those collections unlikely to contain materials that remain under copyright protection and that contain copyright-protected materials that could be removed without significantly decreasing the research value of the collection as a whole.

Under copyright law, unpublished materials created by authors who have been dead 70 years or more (120 years or more for anonymous authors) are not protected by copyright; most works published after 1923 are protected. The SHC will assess all collections for risk, particularly noting the presence of works published after 1923, commercially produced sound recordings and moving images, unpublished works by identified literary authors, photographs with credit lines of photographers who are either living or have been deceased fewer than 70 years, and materials with later creation dates that may be under copyright protection.

As noted in the “Legal and Ethical Implications of Large-Scale Digitization of Manuscript Collections” section, attempts to identify copyright holders revealed that it would be prohibitively time-consuming to attempt to find the copyright holders for every piece of correspondence in a given collection. The SHC staff is developing a strong and visible policy for addressing any copyright complaints and will be prepared to remove material when and if any claimants notify the SHC that they hold copyright to a particular digitized document.

Extensibility

The SHC staff hopes that this decision matrix will help other repositories in designing a plan for large-scale digitization of their own collections. Specifics would necessarily be divergent, but the basic elements may be transferable: The model focuses on a core question of interest (SHC staff identified collections related to race) to the primary research audience (for the SHC it is scholars of the American South) of the archive. In shaping the core question, the keys are to determine the intended audience and to engage that community in setting priorities for digitization. Then a series of questions is developed to enable archivists to prioritize collections according to their relevance to that question.

Conclusion

After the initial assessment of all 4,600 collections, the identification of 1,030 high-priority collections, and the detailed assessment of two key collections, the SHC determined that the decision matrix will be a useful long-term tool for evaluating and prioritizing collections in a large-scale digitization effort.
Digitization Elements

Additional streams of digitization

The SHC staff worked closely with scholars of the American South (who form the primary audience for its holdings) to develop a decision matrix that prioritizes the materials that are likely to be of greatest value to scholars, and these materials will be first in the queue of materials to be digitized.

However, the stream of digitization defined by the decision matrix will not be the only stream resulting in digitized materials at any given time. Grant-funded digitization on specific topics and the digitization of materials by and for other entities within the University have already enabled digitization of some materials sooner than the decision matrix alone has dictated. For example, the University Library is a partner in “Publishing the Long Civil Rights Movement,” a grant project funded by the Andrew W. Mellon Foundation, and the SHC digitizes manuscripts to support scholarly works published online through that program. The SHC also digitizes materials for use in the North Carolina Digital History Textbook Project at LEARN NC, a program of the School of Education at UNC-CH that creates and distributes Internet-based resources for K-12 teachers.7

Users who desire a digital version of a collection that either was not slated for digitization, or that was not going to be digitized soon enough to suit their needs, will be able to request that the collection be digitized and will pay a modest fee to help offset the cost of the service.

Digitization on demand

Currently, the SHC offers a fee-based, digitization-on-demand service for users who request duplication of large quantities of collections. In the Exit Survey, two thirds (15 of 21) of scholars indicated that they would be willing to pay (either through grants, research funding from their home institutions, or other sources) to digitize a collection or a portion of a collection that they needed for their research, which otherwise might not be digitized for months or even years.

The digitization-on-demand services provided for remote users and UNC-CH instructors take place at the folder level: thus, even if only a few letters from a file of correspondence are needed, the entire folder is digitized. As items are digitized, the digital copies will eventually be connected to the collection’s finding aid and made available, even if the entire collection has not yet been fully digitized.

SHC staff has begun to digitize materials on demand for remote individual researchers, other institutions, and UNC-CH faculty members needing digitized documents for instructional use. Between October 2008 and June 2009, approximately 18 linear feet of material were digitized on demand.
The SHC and the University Library are committed to a sustainable and continual digitization program based on the priorities set by the decision matrix. This commitment will require the consistent allocation of current staff time and resources as well as continual expansion of technological capacity and periodic hiring of additional staff to accommodate an ever-growing Digital SHC. Digitization will continue throughout each year, although the pace will be partially determined by the resources allotted through the annual budget and by the availability of external funding.

Sustainability

For digitization of the SHC collections to become truly sustainable, it must be a part of the core mission of the SHC that is institutionally supported. Participation in the digitization process must be built into the job descriptions and work plans of nearly all SHC employees, and consideration for digitization must be a part of the workflow associated with every new collection. In this model, the stream of digitization may need to slow to a trickle at times and it may speed up dramatically at others, depending on available resources, but it will not ever cease.

The immediate long-range goal of large-scale digitization is to digitize the more than 1,030 collections (totaling more than 5 million items) that have been identified as priorities using the decision matrix. These digitized collections will be presented online in conjunction with the collection's finding aid through an interface that provides free and open access to the digitized collections. In the short term, the SHC has determined that existing resources will allow, over the next three to five years, for the digitization of 500 collections (each of which is comprised of one linear foot or less of material and together total approximately 200 linear feet), which will be presented in conjunction with their finding aids through a Web interface that provides free and open access. The SHC plans to seek funding to enable a more rapid schedule of digitization of these initial 500 collections and to establish a defined timeline for digitizing the prioritized 1,030 collections.

In keeping with the goal of creating a sustainable, long-term program of digitization that integrates digitization with the core mission of the SHC, this digitization initiative will not be undertaken by a set of specialized archivists working on this project apart from the rest of the SHC staff. Rather, roles in the digitization process will be divided between staff members who will work primarily with digitization and those who work with the existing physical archive. For example, the SHC proposes hiring a digital SHC coordinator, a collection preparation assistant, and a scanning technician to work primarily on the digitization process. However, ten other University Library staff (the director of the SHC, the Senior Research and Instruction Librarian for the SHC, the head of Technical Services, the Manuscripts Processing librarian, the conservator, the assistant conservator, the conservation assistant, the manager of CDLA Digital Production Center, the CDLA project management librarian, and the CDLA digitization
technician) will have important roles in the proposed digitization program. (For job descriptions of staff members and proposed workflows associated with digitizing manuscript collections, see Proposal for a Sustainable Program of Digitization in Appendix H.)

**Technological issues**

Large-scale digitization also requires planning for the physical technologies and storage capacities that will enable continual digitization and user access. Technological issues were not among the core questions of "Extending the Reach of Southern Sources: Proceeding to Large-Scale Digitization of Manuscript Collections." The simultaneously occurring Watson-Brown Project, however, explored the technological aspects of large-scale digitization such as file-naming conventions, file management, scanning specifications, and metadata.

Building on the work of the Watson-Brown Project, SHC staff considered technology in all phases of the digitization planning process. Digitizing fragile manuscripts requires the use of an overhead scanner to minimize the handling of and potential damage to the original sources. The Carolina Digital Library and Archive’s Digital Production Center currently has one Zeutschel overhead scanner among its imaging equipment, and this scanner is used for the majority of manuscript material. Images of SHC collections that are scanned using the Zeutschel scanner are stored as lossless (compression with retention of original data) JPEG 2000 files (instead of TIFF files, which are larger), which retain extremely high image quality and create more manageable file sizes. Even with the smaller JPEG 2000 files, digitization of huge quantities of manuscripts materials will require a monumental amount of storage capacity. The SHC staff anticipates that the average file size will be 12 MB and that the Digital SHC will require more than a terabyte of additional storage space in the Digital Archive (server space designated for long-term storage of electronic files) as well as a second terabyte on the Web server (server space for the online, publicly available collections) annually.

Any increase in the speed at which digitization occurs will necessitate increasing storage capacity apace, so plans will be put in place for backing up data and for regularly increasing the Digital Archive’s capacity and upgrading servers as needed. The expansion of storage needs and the continual need for upgrades will require increased staffing in the University Library’s Systems Department to ensure the integrity of both the Digital Archive and the Web servers. The Digital SHC interface will be built using commercial information architecture applications that will not require extensive on-site programming, reducing the need for on-staff programmers.

Throughout the ongoing digitization effort, the SHC staff will work closely with the staff of the University Library’s Systems Department and Digital Production Center to continually revise scanning, storage, and interface solutions to meet evolving needs.
The digital search room experience will mirror the physical search room experience: Users will be able to interact with primary sources directly. And the digitized manuscripts will neither be framed with contextualizing or interpretive information, nor supplemented with resources such as scholarly essays, lesson plans, interactive features, or social networking communities.

Digital search room

As previously described, the collections themselves will be digitized in their entirety (with the exception of some sensitive or copyright-protected material that may not be digitized due to legal or ethical considerations); so users will, with few exceptions, have access to the same resources that they could view in the library, and will interact with digitized documents directly.

The initial version of the Digital SHC will not provide users with tools to permit the addition of user-created content, or with other Web 2.0 or social networking features. (In theory, such tools could allow for social tagging, user-generated commentary, wikis, and other forms of communication among users. However, there was no consensus among the graduate students and scholars who shared their views with SHC staff for enabling these types of tools.)

Although many of the graduate students were intrigued with the potential of social tagging, others raised objections and expressed skepticism about the value and reliability of user-generated content. Among the questions asked were: Would scholars really want to share their research work through social networking applications? How would users be identified when adding content? Could users contribute content anonymously (and, if so, would uncertainty about the identity of the writer make the content less valuable to scholars)? How would the SHC prevent or manage vandalism and offensive user-created content?

Many of the scholars also expressed skepticism about the value of user-generated content, concerned that it could unnecessarily “clog” the site with inaccurate and inconsequential information that would distract from the value of the documents. One interviewee said she wanted a “clean research experience . . . one that at least gave the illusion of materials being untouched by others.” Another commented, “I just want to see the documents.” Others hoped that such content would be moderated or vetted by the staff of the SHC.

During the workshop, the scholars seemed focused far more on the potential problems with social networking components than on the benefits of them. In these conversations, they cited discomfort with any kind of registration system that would allow content to be generated only by approved scholars, but also expressed concern about potential problems with anonymously authored content (such as false entries in Wikipedia) and the difficulties in developing an efficient and economical way to
monitor, vet, or edit user-generated text or tags. In the Exit Survey, however, many scholars said that the Digital SHC should incorporate social networking features and that if those features were available, they would possibly use them.

Given the feedback from the graduate students and scholars—strong support for a digital search room experience that mirrors the SHC physical search room, and concern about problems that could arise with user-generated content and other Web 2.0 or social networking features—and considering the costs associated with setting up such tools and then moderating, vetting, or editing them, there are no plans at this time to build in such features.

The finding aids for the digital collections will be identical to the finding aids found on the SHC's website.

Finding aids

The SHC’s archivists arrange and describe manuscript material in aggregate, rather than cataloging or indexing at the item level. Whether using the physical or digital collections, the same finding aids will be available for the primary descriptions of collections and the materials within them and for identifying collections of interest. Digital facsimiles of documents will be organized in the same way as their physical counterparts.

Programmers working at UNC-CH with the University Library's Systems Department and the Carolina Digital Library and Archive have created an infrastructure to allow newly digitized material to be automatically associated with the container structure described in the finding aid. So, as users explore a finding aid online, they will immediately see folders with newly digitized content as hyperlinks that will allow them direct access to those digitized materials. If a letter in the physical collection, for example, is contained in Folder 12, then in the digitized collection the user will be able to access that letter by clicking on the name of the collection, then on Folder 12, then through the individual files to find the letter in question.

Graduate students in the 2007 focus groups expressed a strong preference that digitized documents should clearly map back to the original physical manuscripts. In other words, users should be able to identify the box or folder in which the physical item resided so that the document could be accurately cited and easily double-checked with the original. The graduate students had substantial experience working with finding aids in the SHC and other repositories, and they felt comfortable with finding aids and confident in their ability to use them effectively.

At the August 2007 focus group, graduate students responded positively to the organization of digital collections on the Archives of American Art’s website by finding aids—suggesting that this method seems intuitive and useful to historical scholars. When asked what they would suggest if full-text searching of documents was not implemented in the SHC digitization, participants named subject searching (which can
be easily achieved by searching the finding aids) as a priority.

During interviews with scholars, several interviewees indicated that they liked how the SHC arranged and described its collections and wanted to preserve as much of that as possible in the Digital SHC. All but one scholar (who needs item-level description of photographs for his research) indicated that the SHC finding aids would provide a solid architecture and ample descriptive information for the Digital SHC. Most said they preferred the modern finding aid to item-level calendars or lists, which they agreed provide too much unnecessary detail, and make assessment and contextualization difficult. One interviewee said that the “SHC finding aids [are] great [and] should be [the] primary means of organizing digital collections. The finding aid is a sufficient delivery mechanism for the digital collection.” Another said that the remainder of the work—selecting, reading, and interpreting—was the job of the researcher and that the SHC “should not take as its goal to change or simplify [archival] research.”

The digital version of each document will be a high-resolution image file. With the exception of oral history interview transcripts, there are no immediate plans to use optical character recognition (OCR) systems or word-for-word transcription to enable full-text searchability.

Image resolution

Most paper documents will be represented as digital facsimiles (images of original documents) rather than as full-text searchable transcriptions. The majority of the original documents in the SHC are handwritten documents, and transcribing documents into typed versions one by one as they are digitized to facilitate full-text searching would require huge amounts of staff time, making this option for full-text search capabilities impractical. Just as when exploring documents in the physical search room, users of the digital search room will need to examine each document carefully, decipher the handwriting of the author, and analyze the text for themselves.

Not surprisingly, many of the graduate students initially expressed a desire for either full-text search capability or in-depth indexing at the item level, which will not be possible with image-based digitization. With one exception, participants quickly acknowledged the impracticality of transcribing or indexing the enormous mass of mostly handwritten documents and agreed that online access to facsimiles of millions of documents was a reasonable trade-off for full-text searching. The participant who disagreed advocated for intensive indexing for fewer collections.

In the Exit Survey with scholars, most comments about accessing digitized documents through the user interface focused on the value of a user-friendly, easy-to-navigate way to open individual documents; to move between documents easily; and, if possible, to view multiple documents simultaneously; and to zoom in, pan, or rotate images for closer examination. The scholarly community that uses the SHC for research
purposes is familiar with the types of documents housed in the SHC and understands that OCR technology cannot yet provide for inexpensive full-text searching for handwritten collections. If image files of digitized documents are clear and are easy to find and manipulate within the digital interface, they should meet the needs of the scholars who comprise the SHC’s primary audience.

Digitized material will be made available on the Internet free of charge. No enrollment in a subscription service or affiliation with a university or other institution will be required.

Free accessibility

Interviewed scholars expressed great enthusiasm for democratizing historical research by making manuscript sources freely available to those outside their own professional ranks and to those residing far from or unable to travel to Chapel Hill. As one scholar noted in an interview, “This digitization project has the potential to make a huge difference in the ability of students far away from North Carolina being able to undertake research using SHC holdings—in many cases, it will make the difference between a research topic being viable or not—and it would be a shame if fees presented an obstacle.”

The SHC and the University Library are committed to making the Digital SHC available and freely accessible to as many existing and new users around the world as possible. Therefore, access will not be limited to individuals affiliated with universities or other institutions, but rather will be open to everyone—professional historians, college students, K-12 teachers, genealogists, and the general public alike.

Users of the Digital SHC will be able to download copies of the digitized documents to their own computers and will be able to make use of those documents including republishing them to their own websites.

User interface

The SHC will continue to explore many possible features for the user interface of the Digital SHC, and is particularly interested in finding ways of allowing users to print, download, and repurpose digital images of manuscript sources in useful and creative ways. In the Exit Survey the majority of scholars (18 of 21) noted that they would like to be able to download documents so they could store them for examining again. All but one said they would like to be able to print documents, and all agreed that it would be
helpful if citation information automatically appeared in any downloaded or printed version of a document.

For a variety of other reasons, the option of saving a digital version could appeal to scholars working with digital sources. Some may want to create a library of research files that can be viewed even when they do not have access to the Internet. Others may have specialized image viewing software on their personal computers that they could use to enlarge, enhance, or otherwise manipulate image files. Scholars may wish to use image files as part of a course assignment or educational website.

Given this feedback, the SHC plans to continue to explore the technical issues associated with allowing users to print, download, and repurpose digitized content, and is committed to finding ways to facilitate these kinds of uses to the extent possible.

Visions for the Future

Although it is difficult to predict the directions in which innovations in technology and historical scholarship will lead, the decisions made at this early stage of the digitization process have left room for the possibility of different choices being made in the future, thus allowing the digital version of the SHC to continually remain flexible and responsive to the needs of its users.

Reevaluation

The plans for large-scale digitization that have been created as a result of this grant project form the foundation of an exciting new direction for the SHC. After a significant amount of material has been digitized in accordance with the SHC's initial large-scale digitization plan, it will be important to: 1) reconvene representatives of the scholarly community, 2) consider the successes and drawbacks of this model, 3) revisit digitization procedures with an eye toward improving efficiency, 4) learn how scholars are using this new means of accessing the rich materials of the SHC, and 5) consider possible new directions for the digital search room.

To enable the most efficient and accurate digitization of the collections, SHC staff will, as digitization proceeds, evaluate both the process of digitization and the ways in which patrons are using the Digital SHC. And as more and more collections are digitized, staff will learn the best ways to digitize the myriad materials, how to provide the best training, how to accurately anticipate the amount of staff time needed, and how to determine the optimal workflow. Usage statistics will be kept including the number of unique users who access the collection in a given period of time, which collections are most used or requested for digitization, the number of digital documents downloaded, and which browsers and operating systems are used to access the Digital SHC.
Keeping in touch with community of scholars

SHC staff will continue to be in close contact with the scholarly community—inviting their feedback as the digitization process unfolds, and learning more about how the Digital SHC is enhancing their scholarship and teaching and how the digital search room could be modified to best meet their needs. As one scholar expressed in the workshop, the initial stage of digitization may place a large number of digitized collections online “in a way that's very familiar for scholars to use but would also create the basis for future stages” that could incorporate additional features and functionalities, some of which have not yet been envisioned.

Impact of further expansion of access

Many other communities could greatly benefit from the large-scale digitization as well. For example, as historians and graduate history students are able to access digital versions of SHC collections from anywhere in the world, these resources may become much more widely interpreted and cited in scholarly articles and monographs. In turn, examples from these collections may find their way into undergraduate lectures, and professors and teaching assistants may invite students to explore digital sources through original research projects—an option that would be impossible at any university outside the immediate Chapel Hill area without large-scale digitization.

The Digital SHC may also find its way into elementary, junior high, and high school classrooms as history teachers discover new possibilities for bringing U.S. history to life through the SHC’s diaries, letters, maps, account books, photographs, and oral history interviews. Local libraries and historical organizations may be able to use the Digital SHC to reconnect local communities to the materials that originated within them through links or special Web pages on their own websites, local historical society meetings, or other community events. Genealogists and amateur historians with an interest in the American South will be able to peruse manuscripts sources at any time without having to travel to do so.

As all these different types of users—and perhaps some not yet anticipated—begin to discover the rich resources of the SHC in digital form, it will be important for SHC staff to learn more about how they are using the collections and about the kinds of digital resources, formats, and tools that will help them make the most of their work with the SHC’s holdings.

Partnerships

The SHC may also wish to consider the possibility of partnerships that could facilitate the digitization of other archives. At the Southern Sources workshop, some scholars noted that some smaller archival collections have smaller search rooms, fewer staff members, and more limited hours—making it difficult for them to work extensively within those collections. After digitization has successfully begun at the SHC, the knowledge gained from that process could be profitably applied at other institutions. Moreover, partnerships between the SHC and smaller archives could allow for the digitization of materials from other institutions and for meaningful collaborations among repositories housing similar or related materials.

This process of supporting the digitization needs of other repositories may be begun through the North Carolina Digital Heritage Center, a project of the State Library of North Carolina and the University Library at UNC-CH. The project will provide
digitization and hosting services for cultural heritage materials held by libraries, archives, historical societies, museums, and other institutions throughout the state, and will provide broad access to digitized materials to teachers, students, researchers, and citizens both within and beyond North Carolina. And even more opportunities could arise from the large-scale digitization—opportunities in keeping with the SHC’s mission to support research and scholarship on the American South and, by extension, the world.
## Appendices

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A.</td>
<td>Questions Posed to Graduate Students in Focus Group, June 2007</td>
<td>40</td>
</tr>
<tr>
<td>Appendix B.</td>
<td>Key Interview Questions with Scholars</td>
<td>41</td>
</tr>
<tr>
<td>Appendix C.</td>
<td>Agenda for “Southern Sources: Focusing the Conversation—A Workshop with Scholars of the American South”</td>
<td>42</td>
</tr>
<tr>
<td>Appendix D.</td>
<td>Exit Survey Data</td>
<td>45</td>
</tr>
<tr>
<td>Appendix E.</td>
<td>Agenda for “Legal and Ethical Implications of Large-Scale Digitization of Manuscript Collections Symposium”</td>
<td>55</td>
</tr>
<tr>
<td>Appendix F.</td>
<td>Agenda for “Moving From Projects to a Program: The Sustainability of Large-Scale Digitization of Manuscript Collections Symposium”</td>
<td>56</td>
</tr>
<tr>
<td>Appendix G.</td>
<td>Decision Matrix</td>
<td>57</td>
</tr>
<tr>
<td>Appendix H.</td>
<td>Proposal for a Sustainable Program of Digitization</td>
<td>67</td>
</tr>
<tr>
<td>Appendix I.</td>
<td>List of Scholars Interviewed</td>
<td>75</td>
</tr>
<tr>
<td>Appendix J.</td>
<td>List of Workshop Participants</td>
<td>75</td>
</tr>
<tr>
<td>Appendix K.</td>
<td>List of Advisers and Survey Participants</td>
<td>81</td>
</tr>
</tbody>
</table>
Appendix A. Questions Posed to Graduate Students in Focus Group, June 2007

1. Do you or would you use digitized manuscript collections? Why or why not? If you do use digitized collections, how would you describe and quantify your usage?

2. In your opinion, are there any examples of digital collections that have a particularly effective or ineffective means of presentation and navigation? If so, what are those sites and what specifically do you like or dislike about them?

3. Accepting that there are always trade-offs, is it better to have digital accessibility to a greater volume of materials while potentially sacrificing image quality or to have higher image quality and less volume? For example, the SHC has a good number of collections that have been microfilmed. Does it make sense to digitize that film—a relatively quick and inexpensive way to digitize dozens of collections—and thereby sacrifice image quality and color?

4. It is difficult to imagine having the resources necessary to transcribe the millions of nineteenth-century documents held in the SHC, and therefore, full-text searching of all the documents is an unlikely prospect. In lieu of full-text searching, what are the most desirable modes of searching digital manuscript collections?

5. Assuming that a digitized collection would have some degree of systematic and controlled indexing supplied by archivists, would you also want to have the ability to add and to search social tags? (Social tags are user-generated, natural language descriptors used on sites like Flickr and Del.icio.us.) Why or why not?
Appendix B. Key Interview Questions with Scholars

1. Do you or would you use digitized manuscript collections? Why or why not? If you do use digitized collections, how would you describe and quantify your usage?

2. Should the SHC digitize its microfilmed collections from the microfilm or should it spend the time and money returning to the original documents in order to create the images? Why?

3. How do you envision the online presentation of digitized documents? To what degree, if any, should the online presentation of digitized archival material attempt to mimic the analog archival research experience?

4. With scholarly citation in mind, is it important for the digitized item to map to the physical container (e.g., folder, box) where the original document is housed?

5. Would user fees be acceptable to scholars, educators, and students?
Appendix C. Agenda for “Southern Sources: Focusing the Conversation—A Workshop with Scholars of the American South”

All workshop sessions will be plenary and will be conducted in the Pleasants Family Assembly Room on the main floor of the Wilson Library at the University of North Carolina at Chapel Hill.

April 10, 2008

5 – 5:45 PM
Opening Reception

5:45 – 6:00 PM
Welcome and Introduction of Daniel J. Cohen by Richard V. Szary

6:00 – 6:45 PM
Keynote Address
Speaker: Daniel J. Cohen

7:15 PM
Dinner (Carolina Inn)

April 11, 2008

8:30 – 9:00 AM
Continental Breakfast

9:00 – 9:30 AM
Welcome and Introductions by Tim West, Deborah Barreau, and Laura Clark Brown

9:30 – 10:15 AM
“Current and Emerging Trends in Southern History”
Moderator: Jacquelyn Dowd Hall
Panelists: Laura F. Edwards, Glenda Gilmore, Adriane Lentz-Smith, and Mark M. Smith

10:15 – 11:00 AM
Discussion

11:00 – 11:15 AM
Break
11:15 – 11:30 AM
“Where Do We Start?: Prioritizing SHC Collections for Large-Scale Digitization”
Presenter: Laura Clark Brown

11:30 AM – 12:30 PM
Discussion

12:30 – 1:30 PM
Box Lunch

1:00 – 1:30 PM
Tour of the Carolina Digital Library and Archive (optional)

1:30 – 2:30 PM
“Manuscripts at Midnight: Finding Aids and Interface Models for Digitized Manuscript Collections”
Presenters: Jackie Dean and Maggie Dickson

2:30 – 3:00 PM
Coffee Break

2:30 – 3:00 PM
Tour of the SHC (optional)

3:00 – 4:00 PM
Discussion

5:00 PM
Dinner (Top of the Hill)

April 12, 2008

8:30 – 9:00 AM
Continental Breakfast

9:00 – 9:30 AM
“Beyond Fast Pictures: Exploring the Potential of Interactive Digitized Collections”
Presenter: Christopher (Cal) Lee

9:30 –10:30 AM
Discussion

10:30 – 10:45 AM
Break
10:45 – 11:00 AM
Exit Survey

11:00 AM – 12:00 PM
Next Steps and Concluding Remarks
Presenters: Deborah Barreau and Laura Clark Brown

12:00 – 1:00 PM
Box Lunch
Appendix D. Exit Survey Data

Question 1. How should the SHC prioritize collections for digitization?

1. The SHC should prioritize collections based on usage and digitize “greatest hits” first.
   
   9 Agree 12 Disagree

2. The SHC should prioritize collections that are less heavily used and digitize its “hidden collections” first.
   
   12 Agree 9 Disagree

3. The SHC should digitize collections thematically.
   
   4 Agree 17 Disagree

4. The SHC should prioritize specific genre types—correspondence, diaries, etc.—within collections.
   
   4 Agree 17 Disagree

5. If portions of a collection cannot be digitized, the SHC should prioritize other collections that can be digitized in their entirety.
   
   7 Agree 14 Disagree

6. The SHC should digitize microfilm holdings of reasonable quality rather than re-image the original documents.
   
   9 Agree 12 Disagree

7. I would pay—through grants, my home institution, or other sources—to digitize a collection or portion of a collection that I need for my research.
   
   15 Agree 6 Disagree
Question 2. Do you have any additional ideas or comments on prioritizing SHC collections for digitization?

- “Materials that are difficult to use or that are at risk of degradation (large maps, folded items, bad paper, etc.) create a standardized process and just begin at the beginning—earliest, most at risk collections, and underserved collections would be my priority.”

- “Create a schedule for digitizing the SHC’s collection, while assigning percentages each year to specific issues—themes, donor interests, grants, fragility, etc.”

- “[Please] make sure that all scans are OCR-compatible, should that become a possibility in the future. (Be "forward compatible")”

- “Regarding microfilm—all depends on "reasonable quality" . . .”

- “Willingness to pay for digitization depends of course on price. :-)”

- “Prioritizing whole collections is not a user-driven strategy (in my perspective) but a sensible management strategy for the library—so my feedback here is only as meaningful as the assumption that going collection-by-collection will be most practical.”

- “As a user, I’d be happy seeing a note explaining why such-and-such is missing.”

- “I was curious about the ‘greatest misses’ idea, which reminded me that archivists have a great deal of knowledge that researchers don't find out unless they individually ask. I would be open to a combined approach: a greatest hits and greatest misses approach at the same time, so that for every huge, familiar collection, a smaller greatest miss collection also becomes available.”

- “A general plan of say 80% based on usage figures (greatest hits) with another 20% based on the staff's sense of hidden treasures makes sense to me.”

- “We can't predict trends, so themes or genres seem the wrong way to go.”

- “Much use depends on when we have key word searching for handwriting. If it's not available and typescripts can [be] key word search[ed] [then it] may make sense to start with huge 20th century collection[s]. Entire collections should be done.”

- “Given that we may be looking at a long-term project it may make sense to start with important collections not available on film. The film could come later.”

- “I feel less need to prioritize collections available as microfilm if scholars can see that film via ILL [Interlibrary Loan].”
• “Readability is crucial: Do what you need to do to get the best images.”

• “Start with A Work through Z or start with Z and work backwards. Perhaps—I think the Beinecke [Beinecke Rare Book & Manuscript Library at Yale University] is using a model like this. When someone pays for digitization of a piece for themselves—they sign rights + post . . .”

• “A combination of greatest hits (including oral history) and of small collections—working toward each other + new collections as a matter of course.”

• “I think that the SHC should pursue a layered strategy that incorporates several paths so that 1) there is a wide range of materials over a broad chronological period and a range of scholars would have access to something, 2) then [there] is a foundation of materials on which to build to which to add. So some ‘greatest hits’ some ‘hidden collections’ . . .”

• “Perhaps survey targeted audiences of specialists to get specific ideas based on topic + content rather than just use.”

• “[Use a] combination of—but do entire collection + if anything left out, ID it . . .”

• “I’d urge working toward defining a balance of priorities some ‘greatest hits’ but also some ‘hidden’ collections known to archivist—some for topical + chronological—so that there would be some different layering strategies.”

• “I think it is very important that the early collections digitized represent different chronological eras. Even more so then varied themes and genres, this would allow a wide range of scholars to use the digitized collections in their research and teaching and to offer feedback regarding the bells + whistles . . .”

• “My priorities would be 1) to begin chronologically—older, handwritten documents, 2) collections that include information about or by African American / enslaved people + within the chronological framework . . .”

• “Link prioritizing to some new collecting (i.e., of HBUC [Historically Black Universities and Colleges] archives) . . .”

• “Let the archivists decide . . .”
Question 3. Which model for viewing the overall contents of a folder or volume do you prefer?

Thumbnail light-box display from which you may select an image to examine more closely (Sir Edmund Barton Papers at the National Library of Australia)
12 (63.2%) PREFER

Thumbnail light-box display adjacent to image viewer (George and Phoebe Apperson Hearst Papers in the Online Archive of California)
4 (21.1%) PREFER

Double-paned viewer with thumbnails in scroll bar adjacent to image viewer
(Alexander Calder Papers at the Smithsonian Institution’s Archives of American Art or Transportation Library Menu Collection at the Northwestern University Library)
2 (10.5%) PREFER

Panning and zooming through Google image cutter (Iliad at the Center for Hellenic Studies, affiliated with Harvard University)
1 (5.3%) PREFER

Comments

• “Whatever you all think is best.”
• “Is a flip model possible?”
• “Ideally a combination of B and D . . .”
• “This is tough—none are ideal—develop own user friendly . . .”
• “Panning and zooming . . .”
Question 4. Which model for viewing the individual documents do you prefer?

Panning and zooming capabilities (Alexander Calder Papers at the Smithsonian Institution’s Archives of American Art or the Iliad at the Center for Hellenic Studies)

14 (66.7%) PREFER

Smaller and larger (high resolution) options with no further zooming capacity (Sir Edmund Barton Papers at the National Library of Australia)

1 (4.8%) PREFER

Toolbar with capabilities for manipulating images, such as pan left/right or up/down and rotate image clockwise/counter clockwise (Transportation Library Menu Collection at the Northwestern University Library)

6 (28.6%) PREFER

Comments

• “NWU toolbar but with contextual information attached . . .”

• “[Use a] combination . . .”

Question 5. Do you have any additional ideas or comments on interface?

• “Zooming, downloading and printing with citation embedded, map panning with click and drag . . .”

• “[P]lease use intuitive, non-technical language for users—so that it is very clear how to use the navigational tools!”

• “I would definitely want to view thumbnails and full image side-by-side, even if that assumes an extra-wide browser window. Saves time over clicking back and forth; easier to keep track of where I am. I would love to be able to rotate images to view marginalia, but be realistic about how useful this is. If 20% of documents have text going multiple directions, by all means, provide rotating tools. If 5%, let people download the images and open them in offline tools.”

• “I am most interested in being able to move smoothly from item to folder level, but can imagine various interfaces that would provide that. I don’t like having to move to scroll bars or the back button but would prefer to keep the mouse positioned on the document itself.”

• “Interface is crucial. Need to be able to go doc. to doc. easily—with return access to the index. Most scholars will want to at least [glance] through all the docs in a file, not simply select from a thumbnail. I found the Smithsonian thumbnails hard to use.”

• “If using a google tool, allow the user to drag it to right side of screen . . .”
• “If A, provide the small image to indicate specific location of cursor while zooming in on document . . .”
• “[C]lick on folder to open - return to place . . .”
• “Zooming and rotation are especially important for private correspondence in the late eighteenth and nineteenth centuries.”

Question 6. What are your preferences for searching across the SHC’s online finding aids?

1. Basic keyword searching will be my primary means of searching across finding aids.
   
   18 AGREE

   3 DISAGREE

2. For advanced searching, I would use searches limited to collection titles.
   
   7 AGREE

   13 DISAGREE

3. For advanced searching, I would use searches limited to the LC Subject Headings.
   
   3 AGREE

   18 DISAGREE

4. For advanced searching, I would use searches limited to the abstracts.
   
   13 AGREE

   8 DISAGREE

Question 7. What are your preferences for browsing across the SHC’s online finding aids?

1. For subject browsing across finding aids, broad categories worded in natural language rather than LC Subject Headings would be useful.
   
   19 AGREE

   2 DISAGREE

2. For subject browsing across finding aids, lists of all the LC Subject Headings used in the finding aids would be useful.
   
   12 AGREE

   9 DISAGREE

3. I would browse lists of finding aids under geographic place names such as Caswell County, N.C.
   
   17 AGREE

   4 DISAGREE

4. I would browse lists of finding aids under broad date ranges or eras (e.g., eighteenth century, Reconstruction).
   
   15 AGREE

   6 DISAGREE

5. I would browse lists of finding aids limited to those for newly available collections.
   
   12 AGREE

   9 DISAGREE
6. I would browse lists of finding aids limited to those for digitized collections.

14 AGREE 7 DISAGREE

Question 8. Do you have any additional ideas or comments on searching and browsing the SHC's online finding aids?

• “Regarding browsing by date: I would love to have this and would use it if it were useful—but it sounds as though that would not be useful. The library staff needs to make this judgment, I think. The last item—browsing only digitized collections—I'd use for teaching. I would love to see some metadata analysis done on the finding aids themselves. For example, it should be possible to graph the date content of collections by analyzing the number of folders devoted to specific years of decades. How about a graph that would allow me to see which decades are heavily represented in the entire SHC, then click on, say, the 1930s to see which collections' folders led to that bump in the graph?”

• “I know it is difficult, but it would be really helpful to search by small increments of dates —1800-1850 for instance. I realize that the collections cover broad ranges, but some collections are better for certain ranges of dates than others and that info could be included.”

• “Ideally, index to dates (difficult, I know, since SHC will not digitize at item level). But dates are, at times, the fundamental tools of the trade and it is immensely helpful to have an index that, for example, identifies every item in the SHC for a particular year (the value added for context here is incalculable.)”

• “Searching. I can think of different times I would do different things . . .”

• “[W]here there is OCR + take a look at SEASR [Software Environment for the Advancement of Scholarly Research] or some data mining tools if OCR . . .”

• “Have the back option take you back to the place from where you began—not to the top of the list!”

• “I think using the finding aid as the structure for the collections is an excellent idea. Being able to click on folders + have material there would be great.”
Question 9. Should the digital SHC include social Web functionality?

1. The digital SHC should accommodate user-generated content by providing Web 2.0 features and functionality (e.g., social tagging, bookmarking, wikis, and channels).
   15 Agree 6 Disagree

2. I would contribute content—tags, social bookmarks, comments, and wiki entries—to the digital SHC.
   12 Agree 9 Disagree

3. I would use channels—such as those on YouTube—on the digital SHC in order to create instructional resources.
   8 Agree 13 Disagree

4. If I had assistance from SHC staff, I would use channels on the digital SHC in order to create instructional resources.
   8 Agree 13 Disagree

5. I would be willing to identify myself when contributing content to the digital SHC.
   19 Agree 2 Disagree

6. I would examine user-generated content to find relevant material or learn about others' research on the digital SHC.
   15 Agree 6 Disagree

7. I would examine user-generated content from anonymous contributors on the digital SHC.
   1 Agree 20 Disagree

Question 10. Should the Digital SHC include features to help organize your research?

1. I would use a personal account on the digital SHC to store and organize documents.
   14 Agree 7 Disagree

2. I would download documents I needed to examine again and organize and store them outside the digital SHC.
   18 Agree 3 Disagree

3. I would download selected search results into my own bibliographic/citation program, such as EndNote, RefWorks, or Zotero.
   13 Agree 8 Disagree

4. I would print documents from the digital SHC that I needed to examine again.
   20 Agree 1 Disagree
5. Citation information such as collection name and container number should be visible in the printable or downloadable image of the document.

21 Agree

0 Disagree

6. It is sufficient to provide citation information in the finding aid of the digitized collection.

3 Agree

18 Disagree

Question 11. Do you have any additional ideas or comments on social Web functionality or special features for the Digital SHC?

• “Users should be identified by real name and affiliation . . .”

• “I might be interested in anonymous user-generated comments only to the extent that it was the subject of my research . . . so, if I were researching popular culture, for example . . .”

• “Some use[r]s might like accounts based at the digital Southern, but I would continue to use something like Zotero or Endnote and so I see creating this functionality as a low priority re allocation of resources right now.”

• “I agree with the usefulness of some user generated features such as bookmarking and tagging, but I don't think wikis are needed. There are existing forums such as H-South for scholars and others who want to engage in conversation about southern history, and adding another one for this site does not seem necessary. Features that directly promote archival goals such as clarifying factual information in collections and annotating collections when appropriate would seem most useful and should be prioritized in designing features.”

• “Providing citation information will be a cost benefit decision. It's not unreasonable for scholars to mark it on their copies as they have for years on photocopies. On the other hand, the sort of citation information provided on Am. Hist Newspapers is very useful.”

• “I don't think content should be anonymous. But ultimately social Web features will be the most revolutionary part of this transformation.”

• “[N]ot anonymously, not wikis or channels, maybe [contribute] tags and comments . . . it seems as if the chance of mis-citing would be high if not on digitized document because of difficulties of moving back + forth online . . .”

• “I'd support more utilitarian aspects related to documents—hope we could discourage more general issue-oriented discussion, but room type interaction . . .”
• “Channels, user accounts, bookmarks, [and] comment features dangerous. Don't cheapen the product with extraneous tools. Stick to the missions of the library.”

• “Balance access with some sense of ordering + setting—don't want to restrict, don't want to be deluged. Thanks—very useful meeting.”

• “I would be interested in downloading the contents of an entire folder, not just on the individual document level. It is extremely important that the SHC define and if possible explain why parts of a collection are not digitized. This should be clear when you enter the finding aid and/or initial description of the collection.”
Appendix E. Agenda for “Legal and Ethical Implications of Large-Scale Digitization of Manuscript Collections”

February 12, 2009

8:30 – 11:00 AM
Continental Breakfast available

9:00 – 9:10 AM
Welcome and Introductions by Richard Szary and Tim West

9:10 – 9:30 AM
“Extending the Reach of Southern Sources: A Report on the UNC University Library’s Mellon Project”
Presenter: Laura Clark Brown

9:30 – 10:30 AM
“Third Party Privacy: Legal and Ethical Obligations”
Presenter: Aprille Cooke McCay

10:30 – 10:45 AM
Break

10:45 AM – 12:00 PM
“Third-Party-Privacy-Case-Studies”
Moderators: Nancy Kaiser and Matt Turi

12:00 – 1:00 PM
Lunch

1:00 – 3:00 PM
“Reconciling Modern Archival Practices and Ethics with Large-Scale Digitization”
Moderator: Merrilee Proffitt
Panelists: Barbara Aikens, Max J. Evans, Tom Hyry, Bill Landis, and Dan Santamaria

3:00 – 3:15 PM
Break

3:15 – 3:30 PM
“Due Diligence, Futile Effort: Pursuing Copyright Holders”
Presenter: Maggie Dickson

3:30 – 5:30 PM
“Orphan Works, Fair-Use, and Risk Management”
Moderator: Merrilee Proffitt
Panelists: Heather Briston, Sharon E. Farb, Peter Hirtle, Bill Maher, and Mary Minow
Appendix F. Agenda for “Moving from Projects to a Program: The Sustainability of Large-Scale Digitization of Manuscript Collections”

March 12, 2009

8:30 – 11:00 AM
Continental Breakfast available

9:00 – 9:15 AM
Welcome and Introductions by Richard Szary and Tim West

9:15 – 10:00 AM
“From Investigation to Implementation: Reports from Two Concurrent Digitization Projects in the Southern Historical Collection”
Presenters: Laura Clark Brown and Maggie Dickson

10:00 – 10:15 AM
Break

10:15 AM – 12:00 PM
“Business Models for Large-Scale Digitization Programs”
Presider: Christopher (Cal) Lee
Presenters: Liz Bishoff, Oya Rieger, and John Wilkin

12:00 – 1:00 PM
Lunch

1:00 – 3:00 PM
“Sustaining Large-Scale Digitization Programs”
Moderator: Will Owen
Panelists: Liz Bishoff, Oya Rieger, and John Wilkin

3:00 – 3:15 PM
Break

3:15 – 4:30 PM
“Workflows for Large-Scale Digitization at the Archives of American Art”
Presenters: Barbara Aikens, Toby Reiter, and Karen Weiss
Appendix G. Decision Matrix

Introduction

The decision matrix is a complex series of questions that is applied to each of the SHC’s 4,600 existing manuscript collections and to its future holdings. The matrix questions determine the priority of each individual collection in the stream of large-scale digitization directed by the SHC. Aimed at an audience of scholars, this stream seeks to provide online access to entire manuscript collections containing original and unique materials. Priority is given to collections that contain any documents addressing the “great subtext of American history” and of southern history in particular—race and race relations.

However, the intent of the large-scale digitization model is to digitize the entirety of the prioritized collections rather than extracting those singular items focused on race. For example, a large collection may contain a small pocket of documentation involving treatment of freedpeople. A prioritized collection documenting four generations of an affluent white family in North Carolina from 1813 to 1956 would therefore be digitized in its entirety. The result of this approach is the broad coverage of geography, chronology, material types, creators, and content sought by the SHC’s scholars who rejected prioritization schemes that gave priority to, for example, time periods, geographical locations, or document types (e.g., typescripts with OCR possibilities).

The first set of questions determines whether the collection will proceed through the matrix. Answers to the remaining sets of questions in Part I and Part II of the matrix determine an individual collection’s priority for digitization. Part III of the matrix is reserved for quantitative data about each collection in the priority list.

<table>
<thead>
<tr>
<th>PART I</th>
</tr>
</thead>
</table>

**Original Materials**

Is the majority of the collection or substantive portions of the collection comprised of original materials held in the SHC rather than copies (e.g., microfilm, typescript copies, photostatic copies, photocopies, mimeographs)?

- □ Yes
- □ No

**Documentation**

Does the collection contain any materials related to race or race relations in the American South?

- □ Yes
- □ No
Provenance and Rights

1. Was the collection either a donation or purchase?
   □ Yes □ No

2. If the collection is on deposit, is the material essentially orphaned with no feasible means of locating the depositor?
   □ Yes □ No

3. Is the donor or depositor known to be or likely to be deceased? Is the collection essentially orphaned?
   □ Yes □ No

4. If the donor or an active heir survives the collection, is the donor or heir amenable and favorably disposed to digitization?
   □ Yes □ No

5. If the collection is on deposit and not orphaned, would the depositor be amenable to digitization?
   □ Yes □ No

6. Is the entire collection unrestricted?
   □ Yes □ No

7. Is the majority of the collection: [Check all that apply.]
   □ 120 years old or older?
   □ by creators deceased for 70 years or longer?
   □ by easily identifiable copyright holders likely to allow digitization?
   □ unlikely to provoke copyright infringement claims?

8. Is the collection at low risk for violation of living individuals’ privacy, particularly concerning financial, student, and health information?
   □ Yes □ No

Use

1. Has the collection been used and cited by scholars?
   □ Yes □ No

2. Is the collection likely to be used and cited by scholars if digitized?
   □ Yes □ No

Arrangement and Description

1. Is this collection unlikely to require either substantial re-processing or massive physical preparation prior to digitization?
   □ Yes □ No

2. If the collection is newly accessioned or includes recent additions, has the entire collection been processed?
   □ Yes □ No
3. Does the collection have an online finding aid that includes a container list?
   □ Yes  □ No

SCORING PART I

Originals and Documentation

A "no" to the first question concerning the presence of original materials will result in a removal of the collection from further consideration in the matrix.

If the answer to the question concerning documentation is "yes," the collection proceeds for further evaluation in Decision Matrix Part I.

If the answer is "no" the collection is ranked as a Tier 4 collection for this stream.

Scale

The scale is 100 points.

Point Values

The remainder of Decision Matrix Part I consists of three sections, and each section has sets of questions:

Provenance and Rights (5 sets of questions)

Use (2 sets of questions)

Arrangement and Description (3 sets of questions)

A "yes" in each of the sets of questions within the sections earns the collection ten points.

Tiers

A **Tier 1** designation moves a collection directly into Decision Matrix Part II.
Tier 1 Ranking = 80 to 100 points.

A **Tier 2** designation places a collection in a hold for Decision Matrix Part II.
Tier 2 Ranking = 50 to 70 points.

A **Tier 3** designation essentially rules out a collection for this first scholar stream.
Tier 3 Ranking = 5 to 40 points.

A **Tier 4** designation rules out a collection for this first scholar stream.
Tier 4 Ranking = 0 points.
PART II

[For True or False questions, circle answer.]

Donor Restrictions

1. T or F The collection has no donor or depositor imposed restrictions on access.

2. T or F The collection has no donor or depositor imposed restrictions on reproduction.

3. T or F The collection does not require a user to seek permission to publish from a third party--other than the Library.

4. [Check one box.]
   □ If the collection has any of the above restrictions, the deposit or gift agreement is such that it might be possible for the director of the SHC to lift the restrictions.
   Or
   □ Despite the restrictions, the donor or depositor will likely allow digitization and open online access to some or all of the collection.

Privacy and Sensitive Materials

1. If the collection contains any of the following types of materials that may infringe on the privacy of living third parties, these materials are easily segregated from the bulk of the collection and not inextricably woven throughout the collection. [Check all that apply.]
   □ Student records (e.g., report cards, transcripts, recommendation)
   □ Refereed journals (e.g., correspondence and manuscripts of rejects)
   □ Refereed grants (e.g., grant proposals)
   □ Financial materials (e.g., salary information, banking, investments, Social Security numbers, credit card numbers)
   □ Medical and health information (e.g., medical bills, physician journals, insurance records)
   □ Legal records (e.g., case files, arbitration files)
   □ Business records (e.g., pay stubs, personnel records, grievance files)
   □ “Sensitive” subject matter concerning an identified third party (e.g., sexually explicit information, documents revealing illicit activity such as drug use)
2. T or F If any of the above are present, the absence of those materials in an online version of the collection would not significantly decrease the collection's research value (i.e., without the material in question, this collection should still be a high priority for digitization).

Provenance and Acquisition

1. T or F A gift agreement or loan agreement is on file.

2. T or F If the collection is a loan, the removal of the collection or any portion of the collection is highly unlikely.

3. T or F If the collection is a loan, transfer of ownership may be a possibility if the collection was digitized.

4. T or F If the collection is a loan, removal of the physical collection is less likely if the collection was digitized.

5. T or F If the donors, depositors, or immediate heirs are living and in contact with the SHC, they may be amenable to digitization, including open online presentation of the materials.

Content

1. The collection contains substantive documentation in which of the following SHC subject strengths? [Check all that apply.]
   - □ American Civil War
   - □ Antebellum plantation era
   - □ Business
   - □ Civil rights
   - □ Communities (e.g., Delta Providence Farms; Penn Center; Mound Bayou)
   - □ Family
   - □ Jewish life and culture
   - □ Journalism
   - □ Labor
   - □ Law
   - □ Politics
   - □ Public health
□ Reconstruction
□ Religion and religious communities
□ Sectionalism
□ Slavery
□ Social activism
□ Sociology
□ University of North Carolina at Chapel Hill
□ Writers

2.  **T or F**: The collection is related to other collections by topic or by creators.

3.  These related collections are located at: [*Check all that apply.*]
   - □ Southern Historical Collection
   - □ Southern Folklife Collection
   - □ University Archives
   - □ North Carolina Collection
   - □ Rare Book Collection
   - □ Duke University
   - □ North Carolina State University
   - □ North Carolina Central University
   - □ North Carolina Department of Cultural Resources

**Use**

1.  **T or F**: Based on circulation records, the collection has been used by scholars.

2.  **T or F**: Based on known publications—past, current, and forthcoming—the collection has been used and cited by scholars.

3.  **T or F**: Based on frequent copying or digitization requests, the collection is currently in demand.

4.  **T or F**: If the collection has been here a while but has not been used much by scholars, it is more likely to be used by scholars if digitized.

5.  **T or F**: If the collection is relatively new, it is likely in staff’s estimation to be used by scholars.
Copyright and Intellectual Property

1. **T or F** For portions of collections presumed to remain under copyright protection, the copyright holders are likely unknowable or untraceable.

2. **T or F** If the copyright holders are known or knowable (e.g., as indicated in a statement of copyright on the items; or assumed by authorship such as a novelist's manuscript), we will likely obtain permission to digitize.

3. **T or F** If the collection contains the following copyright protected materials, it is possible to exclude these materials from the digitized collections without disrupting the continuity, integrity, and context of the collection.

4. Which of the following post-1923 published works does the collection contain? [Check all that apply.]
   - ☐ newspapers and magazines (issues or clippings)
   - ☐ pamphlets
   - ☐ scripts
   - ☐ books

5. Which of the following unpublished works in the collection have identified authors or known literary estates? [Check all that apply.]
   - ☐ literary manuscripts
   - ☐ speeches
   - ☐ articles
   - ☐ letters

6. Does the collection contain commercially produced sound recordings or moving images?
   - ☐ Yes
   - ☐ No

7. Does the collection contain photographs under 120 years old with credit lines of photographers either living or deceased fewer than seventy years?
   - ☐ Yes
   - ☐ No
Microfilm and Previous Digitization

1. T or F Assuming that the SHC will give higher priority to those collections without access surrogates, the existence of available microfilm or digital objects for large portions of the collection will lower the collection's priority in the queue.

2. T or F Neither the entire collection nor all the substantial portions of the collection pertaining to African Americans or race relations are available on microfilm.

3. T or F Neither the entire collection nor all the substantial portions of the collection pertaining to African Americans or race relations have been digitized.

Processing and Finding Aid

1. T or F The collection does not require any re-processing or substantial re-housing.

2. T or F The entire finding aid is online.

3. T or F The container list is at the folder level.

Size of the Collection

1. How many linear feet?

2. How many items?

3. How many folders?

4. How many volumes?

5. How many oversize items?

6. Approximately how many scans?
Scope and Content of the Collection

1. What is the chronological scope?
2. What is the date range?
3. What are the bulk dates?
4. Which of the following eras are well represented in the collection? [Check all that apply.]
   - Colonial and Early National
   - Antebellum
   - Civil War and Reconstruction
   - New South
   - Jim Crow
   - Long Civil Rights
5. What is the geographical scope of the collection? [Check all that apply.]
   - North Carolina
   - South Carolina
   - Virginia
   - Georgia
   - Alabama
   - Mississippi
   - Louisiana
   - Florida
   - Tennessee
   - Arkansas
   - Kentucky
   - Texas
   - District of Columbia or Maryland
6. Which of the following does the collection contain? [Check all that apply.]
   - Correspondence
   - Diaries/Journals
   - Ledgers/Daybooks
   - Maps/Plats
   - Drawings/Sketches
   - Photographs
   - Moving images
   - Sound recordings
   - Born digital
Accessions

1. In what year was the original collection accessioned?
   Is this approximate or exact?

2. Acquisition of the collection: [Check one.]
   □ The original accession forms the heart of the collection.
   □ The bulk of the collection came in as additions.

Additions

1. When was the most recent addition made?
2. How many more additions are expected?
3. What size will the additions be?
4. Do the existing and expected additions contain material suitable and desirable for digitization?
   □ Yes □ No
5. Are the additions processed?
   □ Yes □ No

Previous Microfilming and Digitization

1. What portion of the collection is on microfilm?
2. Was the microfilm commercially produced?
   □ Yes □ No
3. If the film was produced in-house, what year was the collection filmed?
4. Are printing negatives available?
   □ Yes □ No
5. What portion of the collection has been digitized?
6. Are the digital objects available online?
   □ Yes □ No
7. Is any part of the collection available in digital form on these sites? [Check all that apply.]
   □ Documenting the American South
   □ Other University Library website
   □ Non UNC-CH website
8. If located on a non-UNC-CH website, is access open?
   □ Yes □ No
Appendix H. Proposal for a Sustainable Program of Digitization

It is difficult to anticipate all financial resources that will be required for a long-range sustainable program of digitization. The cost of technological equipment changes rapidly, and the staff and equipment resources required for a given year will vary depending on the quantity of material to be digitized in that timeframe. However, the SHC staff has anticipated annual needs based on the possibility of scanning roughly 20 hours per week (88,550 individual scans per year) and has developed a base plan for staffing and workflows that would facilitate digitization at that pace. This program could be expanded with additional scanning equipment, staff, and storage capacity to accommodate a more ambitious pace of digitization.

PROGRAM MISSION

GOALS

STAFF ROLES

WORKFLOWS

MINIMUM RESOURCE REQUIREMENTS

PROGRAM MISSION

The SHC seeks to create a Digital SHC that offers unprecedented access to the rich documentary holdings currently available only to those individuals who could travel to UNC-CH. Digital facsimile editions of SHC collections will allow users anywhere in the world to conduct research without leaving home or to preview collections prior to a visit to use the original, physical documents.

GOALS

Immediate Long-Range Goal

- To digitize more than 1,030 collections totaling more than 5 million items.
- To present each digitized collection online in conjunction with the collection's finding aid.
- To provide free and open access to the digitized collections.

Short-Term Goal (3 – 5 years)

- To digitize 500 collections (each one linear foot or less) and together total approximately 200 linear feet.
- To present each digitized collection online in conjunction with the collection's finding aid.
- To provide free and open access to the digitized collections.
Note: The SHC plans to seek funding to allow for a more rapid digitization of these initial 500 collections and to establish a defined timeline for digitizing the prioritized 1,030 collections.

**Staff Roles**

**Digital SHC Coordinator**

- Projects and tracks digital storage requirements and usage with the CDLA project management librarian.
- Supervises the student assistant who physically prepares the collections prior to digitization.
- Maintains tracking instruments for collection materials, processes, and digital storage.
- Schedules the conservation assessment for collections with the Conservation Lab.
- Schedules the production and equipment assessment for collections with the CDLA Digital Production.
- Coordinates movement of physical collections with the scanning technician.
- Compiles and maintains guidelines and manuals for the Digital SHC program as a whole and for the preparation and scanning of particular types of manuscript materials.
- Reviews the collection’s finding aid and, if necessary and in conjunction with the supervising technical services archivist, modifies the finding aid to bring it up to modern standards and proper encoding.
- Links the digital objects to the finding aid.
- Communicates and consults with the director of the SHC and the Senior Research and Instruction Librarian for the SHC.

**Collection Preparation Assistant**

- Physically prepares the collection for digitization.
- Ensures correct order and arrangement of materials.
- Identifies manuscripts needing conservation attention that were not pretreated and refers these items to the Conservation Lab through the digital SHC coordinator.
- Provides a rough count of images for space and time projections.
- Identifies (in consultation with the digital SHC coordinator) special handling needs and unusual material types, and notes directions for the scanning technician, on a copy of the finding aid (which is sent to the Digital Production Center).
Scanning Technician

- Uses (under the direction of the project management librarian) the appropriate equipment for image capture of the manuscripts.
- Uses file-naming conventions.
- Captures and maintains the technical metadata.
- Coordinates physical movement of the collections with the digital SHC coordinator.
- Identifies manuscripts needing conservation attention that were not pretreated and refers these items to the Conservation Lab through the digital SHC coordinator.
- Contributes to the manuals and guidelines developed for scanning processes and conventions.

Director, SHC

- Oversees the program budget and resource allocations.
- Develops and contributes (in cooperation with the head of research and instructional services for the SHC and the digital SHC coordinator) to the priorities list.

Senior Research and Instruction Librarian for the SHC

- Maintains, develops, and contributes (in cooperation with the director of the SHC and the digital SHC coordinator) to the priorities list.
- Advises the digital SHC coordinator and the director of the SHC on matters of intracollection selection and prioritization.
- Contributes to the manuals and guidelines developed for scanning processes and conventions.
- Gathers user feedback and coordinates responses to users.
Archival Technical Services Staff

Head, Special Collections Technical Services and Manuscripts Processing Librarian

• Train the digital SHC coordinator on technical services work.
• Supervise the technical services work—including finding aid revision and posting of digitized content—of the digital SHC coordinator.
• Develop and maintain finding aid style sheet.
• Oversee SHC’s use of content management system.
• Coordinate technological troubleshooting with University Library Systems Department.
• Contribute to the manuals and guidelines developed for scanning processes and conventions.

Conservation Lab Staff

Conservator and Assistant Conservator

• Conduct conservation assessments for collections.
• Oversee conservation treatment on collection materials.
• Advise SHC and Digital Production Center staff on handling of manuscripts with conservation concerns during the prepping and scanning processes.

Conservation Assistant

• Repairs documents (under the supervision of the conservator and assistant conservator) identified prior to scanning.

CDLA Digital Production Center Staff

Manager, CDLA Digital Production Center and CDLA Project Management Librarian

• Conduct the production and equipment assessments for collections.
• Hire, train, and supervise the scanning technician.
• Schedule the equipment.
• Maintain quality control systems.
• Contribute to the manuals and guidelines developed for scanning processes and conventions.

CDLA Digitization Technician

• Performs any necessary image editing and other postdigitization processing.
WORKFLOWS

Prescanning Workflow

- Digital SHC coordinator consults the decision matrix for the next collection in priority listing.
- Digital SHC coordinator analyzes the finding aid to determine whether it is encoded with a container listing. If the finding aid requires modification, the digital SHC coordinator will proceed with that work while the other workflows progress.
- Digital SHC coordinator pulls collection and schedules conservation assessment with the Conservation Lab.
- If the collection needs treatment, Conservation Lab will schedule treatment; or if the work required is more extensive than resources can accommodate at that time, will postpone digitization. In that case, the digital SHC coordinator will proceed with another collection.
- If the collection is approved by the Conservation Lab for digitization without pretreatment, the digital SHC coordinator schedules the equipment and production assessment with the Digital Production Center. The Digital Production Center then schedules the collection for the appropriate equipment.
- Digital SHC coordinator projects the amount of time needed for the collection to be closed and officially closes the collection for that period, notifying the Senior Research and Instruction Librarian for the SHC and placing a notification on the online finding aid and a sign in the stacks location.
- Following the equipment and production assessment, the collection preparation assistant straightens the collection—e.g., with the finding aid, ensures correct order of documents and folders; removal of metal fasteners where necessary—and adheres to any strictures or suggestions emerging from the assessments.
- With the digital SHC coordinator, the collection preparation assistant identifies special handling and unusual material types, and notes directions for the scanning technician on a copy of the finding aid, which is sent to the Digital Production Center with the collection.
- Collection preparation assistant calculates a rough image count of the collection and communicates that information to the digital SHC coordinator.
- Collaborating with the CDLA project management librarian the digital SHC coordinator projects and records digital storage needs for the collection. The CDLA project management librarian provides the digital storage requirements to the Library Systems Department.
- Digital SHC coordinator coordinates with the scanning technician to transport the collection to the Digital Production Center.
- Digital SHC coordinator uses the tracking tool to log assessments, processes, and movements of the collection materials.
• Digital SHC coordinator relays the progress being made on digitizing the collection to the Senior Research and Instruction Librarian for the SHC.

Scanning Workflow
• Scanning technician (trained by staff in the Digital Production Center and by conservators) uses the scheduled equipment to scan an image of every item as directed by the digital SHC coordinator.
• Scanning technician uses the established file-naming conventions for the scans and the established procedures for file management and drive usage.
• Scanning technician captures and maintains the technical metadata as directed by the digital SHC coordinator and the project management librarian.
• Scanning technician identifies manuscripts needing conservation attention that were not pretreated and refers these items to the Conservation Lab through the digital SHC coordinator.
• Scanning technician arranges for secure transport of collection materials to and from the Digital Production Center workspace with the digital SHC coordinator.
• Scanning technician logs work time and production for each collection.

Postscanning Workflow
• Digitization technician in the Digital Production Center performs any necessary image editing and other postdigitization processing.
• Digital SHC coordinator oversees the reshelving of the collection.
• Digital SHC coordinator reopens the collection, removing the sign from the stacks and the notification on the online finding aid.
• Digital SHC coordinator loads the scans and accompanying metadata into the content management system.
• Digital SHC coordinator follows procedures (to be established) for relating the collection's finding aid to its digital objects in content management system, mounts the modified finding aid to the Web, and alerts the head of technical services to changes that may be needed in the collection's records.
• In coordination with the head of technical services, the digital SHC coordinator communicates any technical failures or problems to the Systems Department.
• In coordination with the Senior Research and Instruction Librarian for the SHC and the director of the SHC, the digital SHC coordinator submits a brief form report on digitization for the collection's control file.
MINIMUM RESOURCE REQUIREMENTS

The minimum resource requirements are based on calculations for 20 hours of scanning for 50 weeks in a year.

Personnel

Program Staff

- Digital SHC coordinator
- Scanning technician
- Collections preparation assistant
- Conservation assistant

Permanent Staff and Associated Departments

- Director of SHC
- Senior Research and Instruction Librarian for the SHC
- Archival Technical Services Section
- Digital Production Center
- Conservation Lab
- University Library’s Systems Department

Software and Equipment

- Imaging
- Image Processing
- Zeutschel overhead scanner (20 hours per week)
- Other imaging devices as needed for oversize and variable formats (e.g., slides, negative film.)
- Work station for digital SHC coordinator
- Desktop computer

Printer

Content management system

Supplies

- Conservation tools (e.g., spatulas)
- Conservation repair tissue
- Archival bond paper (all sizes)
- Archival folders (legal size)

**Annual Projected Digital Storage Space Requirements**

**File Types Produced**

400 dpi JPEG 2000 (RGB color)

**Average File Size**

12 MB

**Projected Number of Scans**

88,550 scans per year (based on 20 hours of a scanning technician on one scanner)

**Network Storage Plan**

JPEG 2000 files on production server while a collection is in process;

JPEG 2000 files to Digital Archive (server space designated for long-term storage of electronic files) as collections are completed.

**Space Required for Web Access**

1,062,600 MB (1.06 terabytes)

**Space Required for Digital Archive**

1,062,600 MB (1.06 terabytes)
Appendix I. List of Scholars Interviewed (2008 Annual Meeting of American Historical Association, phone, at UNC-CH)

Edward L. Ayers, PhD
President
University of Richmond

Bruce Baker, PhD
Senior Lecturer
Department of History, Royal Holloway, University of London

William Blair, PhD
Professor / Director
Richards Civil War Era Center
Department of History, Pennsylvania State University

Judkin Browning, PhD
Assistant Professor
Department of History, Appalachian State University

Stephanie M. H. Camp, PhD
Associate Professor
Department of History, Rice University

Peter Carmichael, PhD
Eberly Family Professor of Civil War Studies
Department of History, West Virginia University

James Cobb, PhD
Professor / Spalding Distinguished Research Professor
Department of History, University of Georgia

Barbara Hahn, PhD
Assistant Professor
Department of History, Texas Tech University

Christine Heyrman, PhD
Robert W. and Shirley P. Grimble Professor of American History
Department of History, University of Delaware

Tracy K'Meyer, PhD
Associate Professor / Co-Director
Oral History Center
Department of History, University of Louisville

James Leloudis, PhD
Associate Professor / Associate Dean of the Honors Program / Director of James M. Johnston Center for Undergraduate Excellence
Department of History, UNC–Chapel Hill
William A. Link, PhD
Richard J. Milbauer Professor
Department of History, University of Florida

Matthew Mason, PhD
Assistant Professor
Department of History, Brigham Young University

Michael O'Brien, PhD
Professor of American Intellectual History
Jesus College, Cambridge University

Paul Quigley, PhD
Lecturer in American History
School of History, Classics, and Archaeology, University of Edinburgh

Adam Rothman, PhD
Associate Professor
Department of History, Georgetown University

Philip Troutman, PhD
Assistant Professor
University Writing Program, George Washington University

John Michael Vlach, PhD
Professor of American Studies and Anthropology / Director
Folklife Program
Department of Anthropology, George Washington University

Jeffrey Young, PhD
Lecturer / Associate Director
Honors Program, Georgia State University
Appendix J. List of Workshop Participants

Edward E. Baptist, PhD
Associate Professor
Department of History, Cornell University

Deborah Barreau, PhD
Assistant Professor
School of Information and Library Science, UNC-Chapel Hill

Laura Clark Brown, MA, MLIS
Project Director, “Extending the Reach of Southern Sources”
Southern Historical Collection, University Library, UNC-Chapel Hill

Tad Brown
President
Watson-Brown Foundation, Inc.

W. Fitzhugh Brundage, PhD
William B. Umstead Professor of History / Director of Graduate Studies
Department of History, UNC-Chapel Hill

Orville Vernon Burton, PhD
Professor / Senior Research Scientist
National Center for Supercomputing Applications
Department of History, University of Illinois at Urbana-Champaign

Alexander X. Byrd, PhD
Assistant Professor
Department of History, Rice University

Hugh Cayless, PhD
Head, Technology Research and Development
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Daniel J. Cohen, PhD
Assistant Professor / Director of Center for History and New Media
Department of History and Art History, George Mason University

Jackie Dean, MSLS
Manuscripts Processing Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

Maggie Dickson, MSLS
Watson-Brown Project Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

Laura F. Edwards, PhD
Professor
Department of History, Duke University
Crystal Feimster, PhD
Assistant Professor
Department of History, UNC-Chapel Hill

Marcie Cohen Ferris, PhD
Assistant Professor
Curriculum in American Studies, UNC-Chapel Hill

Kirill Fesenko, MS
Director, Carolina Digital Library and Archive / Co-Principal Investigator
“Extending the Reach of Southern Sources”
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Gaines Foster, PhD
T. Harry Williams Professor
Department of History, Louisiana State University

Raymond Gavins, PhD
Professor
Department of History, Duke University

Glenda E. Gilmore, PhD
Peter V. and C. Vann Woodward Professor
Department of History, Yale University

Lorri Glover, PhD
Associate Professor
Department of History, University of Tennessee Knoxville

Jacquelyn Dowd Hall, PhD
Julia Cherry Spruill Professor / Director, Southern Oral History Program
Department of History, UNC-Chapel Hill

Lynn Holdzkom, MSLS
Assistant Curator of Manuscripts / Head of Technical Services / Manuscripts Cataloging Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

John C. Inscoe, PhD
University Professor
Department of History, University of Georgia

Nancy Kaiser, MA, MSLS
Project Archivist
Southern Historical Collection, University Library, UNC-Chapel Hill

Stephen Kantrowitz, PhD
Associate Professor
Department of History, University of Wisconsin-Madison
Robert R. Korstad, PhD  
Kevin D. Gorter Associate Professor of Public Policy Studies and History  
Hart Leadership Program  
Department of History, Duke University

Christopher (Cal) Lee, PhD  
Assistant Professor  
School of Information and Library Science, UNC-Chapel Hill

Adriane Lentz-Smith, PhD  
Assistant Professor  
Department of History, Duke University

Sarah Michalak, MLS  
University Librarian / Associate Provost  
University Library, UNC-Chapel Hill

Robert J. Norrell, PhD  
Bernadette Schmitt Chair of Excellence  
Department of History, University of Tennessee Knoxville

Jennifer Ritterhouse, PhD  
Associate Professor  
Department of History, Utah State University

David Silkenat, PhD  
Research Assistant, “Extending the Reach of Southern Sources”  
Southern Historical Collection, University Library, UNC-Chapel Hill

Mark M. Smith, PhD  
Carolina Distinguished Professor  
Department of History, University of South Carolina

Natasha Smith, MSLS  
Head, Documenting the American South / Head, Digital Publishing and Project Management  
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Steven M. Stowe, PhD  
Professor  
Department of History, Indiana University Bloomington

Richard V. Szary, MA  
Director of Wilson Library / Associate University Librarian for Special Collections  
University Library, UNC-Chapel Hill

Helen R. Tibbo, PhD  
Professor  
School of Information and Library Science, UNC-Chapel Hill
Kate Douglas Torrey
Director
University of North Carolina Press, UNC-Chapel Hill

Matthew Turi, MA, MSLS
Manuscripts Reference Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

David Walbert, PhD
Editorial Director and Lead Developer / Designer
LEARN NC, UNC-Chapel Hill

Harry Watson, PhD
Professor / Director
Center for the Study of the American South
Department of History, UNC-Chapel Hill

Tim West, MAT, MSW
Director, Southern Historical Collection / Curator of Manuscripts / Co-Principal Investigator
“Extending the Reach of Southern Sources”
Southern Historical Collection, University Library, UNC-Chapel Hill

Heather Williams, JD, PhD
Associate Professor
Department of History, UNC-Chapel Hill

Charles R. Wilson, PhD
Kelly Gene Cook, Sr. Chair in History
Department of History, University of Mississippi
Appendix K. Grant Advisers and Symposia Participants

Barbara Aikens, MA
Chief, Collections Processing
Archives of American Art, Smithsonian Institution

Wendy Bernhardt
Copy Editor
Bernhardt Editing Services Inc

Jodi M. Berkowitz, MSLS
Legacy Finding Aids Archivist
Southern Historical Collection, University Library, UNC-Chapel Hill

Liz Bishoff, MLS
Director, Digital & Preservation Services
Bibliographical Center for Research

Heather Briston, JD, MSI
Richard and Mary Corrigan Solari University Historian and Archivist
University of Oregon

Laura Clark Brown, MA, MLIS
Project Director, “Extending the Reach of Southern Sources”
Southern Historical Collection, University Library, UNC-Chapel Hill

Rachel Canada
Administrative Assistant
Southern Historical Collection, University Library, UNC-Chapel Hill

Hugh Cayless, PhD
Head, Technology Research and Development
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Joyce Chapman
Graduate Student
School of Information and Library Science, UNC-Chapel Hill

Jackie Dean, MLS
Manuscripts Processing Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

Maggie Dickson, MSLS
Watson-Brown Project Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

Max J. Evans, MA
Director, Collections and Research
The Church of Jesus Christ of Latter-day Saints History Department
Sharon E. Farb, JD, PhD
Associate University Librarian, Collection Management and Scholarly Communication
University of California, Los Angeles

Kirill Fesenko, MS
Director, Carolina Digital Library and Archive / Co-Principal Investigator
“Extending the Reach of Southern Sources”
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Peter Hirtle, MA, MLS
Senior Policy Advisor and Intellectual Property Officer
Cornell University Library

Lynn Holdzkom, MSLS
Assistant Curator of Manuscripts / Head of Technical Services /Manuscripts Cataloging Librarian
Southern Historical Collection, University Library, UNC-Chapel Hill

Biff Hollingsworth, MSLS
Collecting and Public Programming Archivist
Southern Historical Collection, University Library, UNC-Chapel Hill

Tom Hyry, MLIS
Head, Manuscript Unit
Beinecke Rare Book & Manuscript Library, Yale University

Nancy Kaiser, MA, MSLS
Project Archivist
Southern Historical Collection, University Library, UNC-Chapel Hill

Bill Landis, MILS
Head, Arrangement and Description
Manuscripts and Archives, Yale University

Bill Maher, MA, MLS
University Archivist
University of Illinois at Urbana-Champaign

Aprille Cooke McKay, JD, MSI
Digital Preservation Specialist
Inter-University Consortium for Political and Social Research
University of Michigan

Mary Michalak, MLS
University Librarian / Associate Provost
University Library, UNC-Chapel Hill

Mary Minow, JD, AMLS
Consultant
Librarylaw.com
Will Owen, MA, MLS  
Head of Library Systems  
University Library, UNC-Chapel Hill

Merrilee Proffitt  
Senior Program Officer  
OCLC Programs and Research

Oya Reiger, MIS, MPA  
Associate University Librarian  
Cornell University Library

Toby Reiter  
Web Developer  
Archives of American Art, Smithsonian Institution

Dan Santamaria, MSI  
Assistant University Archivist for Technical Services  
Seeley G. Mudd Manuscript Library, Princeton University

Jessica Sedgewick, MSLS  
Manuscripts Processor (African American Collections)  
Southern Historical Collection, University Library, UNC-Chapel Hill

David Silkenat, PhD  
Research Assistant, “Extending the Reach of Southern Sources”  
Southern Historical Collection, University Library, UNC-Chapel Hill

Tim Shearer, MSLS  
Web Development Coordinator  
University Library, UNC-Chapel Hill

Holly Smith, MA, MS  
Overholser Archival Fellow for African American Studies  
Southern Historical Collection, University Library, UNC-Chapel Hill

Natasha Smith, MSLS  
Head, Documenting the American South / Head, Digital Publishing and Project Management  
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Fred Stipe  
Manager, Digital Production Center  
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Richard V. Szary, MA  
Director of Wilson Library / Associate University Librarian for Special Collections  
University Library, UNC-Chapel Hill
Matthew Turi, MA, MSLS  
Manuscripts Reference Librarian  
Southern Historical Collection, University Library, UNC-Chapel Hill

Rita Van Duinen, MSLS  
Project Management Librarian  
Carolina Digital Library and Archive, University Library, UNC-Chapel Hill

Kathryn Walbert, PhD  
Independent Scholar

Karen Weiss, MLS.  
Information Systems Manager  
Archives of American Art, Smithsonian Institution

Tim West, MAT, MSW  
Director, Southern Historical Collection / Curator of Manuscripts / Co-Principal Investigator  
“Extending the Reach of Southern Sources”  
Southern Historical Collection, University Library, UNC-Chapel Hill

John P. Wilkin, MA, MLS  
Associate University Librarian for Library Information Technology and Technical and Access Services  
University Library, University of Michigan
Notes

1 A finding aid is a description of a collection of archival material. It provides information about a collection and its creators and an outline of a collection’s contents.

2 All the collections considered as case studies included sensitive material that was deeply interwoven throughout the collection. Redacting only the sensitive portions of each collection would be tremendously difficult and cost-prohibitive due to the staff time involved. Following the symposium, SHC staff examined the feasibility of redacting material by having a professional archivist attempt redaction with the Delta Health Center Records family case studies. Redacting a single folder took three hours, in addition to the time it took to duplicate the original documents. Clearly, redaction on a folder-by-folder basis would require far more time and resources than would be feasible for SHC’s large-scale digitization efforts.

3 Thomas E. Watson was a Georgia populist politician and newspaper publisher.


5 Archival processing is the technical work of arranging, preserving, describing, and cataloging manuscript collections.

6 In the future, the SHC could form a relationship with a commercial publisher interested in digitizing microfilmed copies of the SHC’s holdings, with the stipulation that the SHC could still digitize the original documents and make them broadly available through the Digital SHC.

7 Future grants or initiatives within departments at UNC-CH may create opportunities for partnerships that would lead to more rapid digitization of resources from the SHC that might not have been prioritized as highly based on the decision matrix.

8 Peretz, Ibid.