

How to Wrangle Multiple Discrete Collections from One Donor: A Case Study of the Subject-based Physical and Digital Consolidation of the Wade Hall Collections

Laura M. Gentry, Erin Ryan, Jessica Rayman, and Martha Bace

ABSTRACT

The authors examined the Wade Hall Consolidation Project at the University of Alabama Libraries Special Collections. The project involved the physical consolidation of more than 1,400 small, discrete collections donated by Wade Hall into larger, subject-based collections along with the merger of 287 existing digital collections to mirror the physical arrangement. This project's goal was to improve access to and discovery of these collections by researchers. During physical consolidation, the archivists created subject-based collections with new finding aids and addressed issues including unclear provenance, legacy descriptions, inaccurate metadata, varying levels of processing, and lack of alignment with current archival best practices and standards. Digital consolidation of existing digital collections coincided with the migration to a new digital asset management system and presented its own challenges, including legacy descriptions, metadata transformation, digital preservation, and dealing with existing metadata shared on the Digital Public Library of America (DPLA) and other multi-institutional digital content aggregators. The authors sought to fill the gap in the literature concerning the consolidation of physical and digital collections and to provide guidance to others considering a consolidation project.

© Laura M. Gentry, Erin Ryan, Jessica Rayman, and Martha Bace.



KEY WORDS

Subject-based arrangement, Consolidation, Legacy description, Metadata harvesting, Digital preservation, Provenance, Finding aids, Digital collections

What do you do when you have more than 1,400 discrete, small collections received over twenty years from one donor? How do you begin to get a handle on it all and make the materials discoverable to users?

The University of Alabama Libraries Special Collections has been a lucky recipient of the largess of the late Dr. Wade Hall. He wrote or edited numerous works relating to Kentucky, Alabama, and southern history, including *The Kentucky Anthology: 200 Years of Writing in the Bluegrass State* (2005) and *Conecuh People: Words of Life from the Alabama Black Belt* (1999).

Dr. Hall was an avid collector, focusing primarily on southern history and culture. Beginning in the early 1990s until his death in 2015, he donated materials to the libraries of the Universities of Alabama and Kentucky and Troy University in southern Alabama, as well as to art museums in Birmingham and Mobile, Alabama, and Columbus, Georgia. The University of Alabama Libraries Special Collections received from him thousands of items, including books, sheet music, recorded music, photographs, and manuscripts, unfortunately with no provenance indicated other than the fact that he donated them. The staff of the University of Alabama Libraries cataloged the published items and processed archival materials to make them publicly available and discoverable by users.¹

This article focuses on Wade Hall's donations of manuscript collections that are each smaller than one linear foot. Originally, Special Collections staff processed manuscript materials as they arrived and identified them according to the creator, resulting in approximately 1,400 small, discrete manuscript collections, the majority of which consisted of only one or two items. At that time, Special Collections relied heavily on students for processing these collections, as there was only one processing archivist. Over the years, those Wade Hall collections relating to women's stories, slavery, African American history, and other trending topics typically led to more research requests and received higher-level processing and detail in their finding aids. Less-popular collections might only receive "placeholder" descriptions and limited subject headings, making it much more difficult for users to discover them.

And therein lies the basis of the original question: how does one manage so many discrete collections and make them more accessible? Our answer was to rearrange these Wade Hall collections thematically into larger subject-based physical collections, create new finding aids, and consolidate existing digital collections to mirror the new physical collections.

Literature Review

In this section, we review larger debates within the literature regarding provenance and subject-based collections. The impetus for the Wade Hall Consolidation Project was to accommodate the changing needs of researchers

in an online environment, most often with the creation of online finding aids and digital collections, a subject the literature documents extensively. This case study explores legacy data and the importance of updating finding aids to align with current standards. The literature rarely speaks specifically to the physical and digital consolidation of existing collections, but a few case studies did offer advice that informed the decision-making regarding the Wade Hall Consolidation Project. Notably absent in the literature is discussion of digital preservation or metadata harvesting of digital collections that are part of consolidated collections.

PROVENANCE AND ARRANGEMENT

Archival decisions about how to describe a collection's provenance are sometimes not as straightforward as they might seem. Jane Zhang points out, in her discussion of the origins of archival theory from its European beginnings in the nineteenth century, that the "principle of provenance mandates that records of the same origin should be processed as one record group and not be intermingled with records of other origins when accessioned, organized, and described into archival collections."² However, the first edition (2004) of *Describing Archives: A Content Standard* (DACS), in Part II's "Introduction to Describing Creators," indicates:

The structure and content of archival materials cannot be completely understood without some knowledge of the context in which they were created. It is insufficient for the archivist simply to include the name of the creator in the title of the description of the materials. Additional information is required regarding the persons, families, and corporate bodies responsible for the creation, *assembly*, *accumulation*, and/or maintenance and use of the archival materials being described.³ [Emphasis added.]

In deciding who a collection's creator is—and what its "origin" is—archivists sometimes have to go beyond simply describing who wrote a letter, for example, to give the full picture and context of how that letter relates to the materials around it and why it is in a repository in the first place.

SUBJECT-BASED COLLECTIONS

The problem with a creator-based arrangement was that the archivists knew very little about the creators of many of the materials, particularly the correspondence. Records for the small Wade Hall collections only denoted the creator by first name and a location. Rearranging these collections with the donor, Wade Hall, as the compiler still aligns with the concept of provenance, but the additional step to arrange the disparate materials into subject-based

collections starts to violate the idea of *respect des fonds*, which would seem to require keeping them separate to preserve context. The purpose of making these collections subject-based is to create context artificially to help the user discover relevant materials. Authors began exploring this tension of creating context without violating the principles of provenance starting in the 1980s. Richard Lytle studied how users approach research in archives by examining subject queries with provenance (Lytle's term for linking subject queries to the administrative histories and biographies of a collection) and context indexing (Lytle's term for matching subject queries with an index or a catalog). Lytle's study found that users tend to seek archival collections on a specific subject, and it exposed the flaws in a provenance-based method of arrangement, as users depend on an archivist to locate collections of interest on a specific subject.⁴ In 1982, Mary Jo Pugh criticized strict adherence to provenance and original order and argued that users want to locate archival materials by specific subjects. At the time Pugh was writing, this was only possible through the publication of a specific subject guide on a topic of interest or through an in-person reference interview between the user and the archivist.⁵ Building upon the work of Lytle and Pugh, Elsie Freeman argued that the archival profession was not oriented to users, was unaware of its user base, was uninformed of how research happens within the archives, and provided inadequate help. She advocated that archivists should perform usability studies.⁶ At the time Freeman, Lytle, and Pugh were writing, users visited the archives in person to view finding aids and materials, and they asked reference questions by mail or by phone, and the usability studies of the next decade reflect this environment. It would take time and technological innovation for the idea of a subject-based arrangement to percolate again to the forefront of archival literature.

ONLINE FINDING AIDS AND DIGITAL COLLECTIONS

The Internet revolutionized the archival profession through the creation of online finding aids and digital collections. In 1997, Thomas Ruller predicted a world where most researchers would interact with archives online, leading to a reimagining of the finding aid and the rise of usability studies in the online environment.⁷ By the 2000s, Wendy Duff, Heather MacNeil, Max Evans, and many others were documenting the archival profession's varied responses to the online environment and its effects on description, usability, and providing reference services.⁸ The transformation of the finding aid to Encoded Archival Description (EAD) was another innovation.⁹ Corey Nimer and J. Gordon Daines III concentrated their usability study on online finding aids, suggesting that the age of search engines had altered users' searching behavior.¹⁰ Elizabeth Yakel and Deborah A. Torres conducted interviews with users of archives and found

them unfamiliar with archival terminology, the structure of finding aids, and search strategies to locate information effectively. Helen Tibbo, Wendy Duff, Catherine Johnson, Donghee Sinn, and Nicholas Soares focused on historians and genealogists, showing the evolution in researchers' information-seeking behaviors while using archives.¹¹ With the rise of online finding aids and digital collections, users did not necessarily have to visit the reference room to interact with the materials and seek help from the archivist.¹² Elizabeth Yakel suggested this shift might change the way archivists keep information "within in its original context or supply a context that enables use in a contextualized way."¹³ The new subject-based arrangement of the Wade Hall collections fits Yakel's idea of a contextualized arrangement. Jennifer Schaffner argued that users often rely on search engines as a means of discovery and search by subjects and key words. Archivists organize collections by provenance, but users are often interested in the "aboutness" of a collection instead of what it is made of ("ofness"). Schaffner concluded that professionals should close this gap between archival descriptive practices and the expectations of users, especially when dealing with minimally described collections.¹⁴

CASE STUDIES—PHYSICAL AND DIGITAL CONSOLIDATION

Although the literature has advocated for archivists to incorporate the needs of users and the role of an online environment into archival description, it offers few practical examples to show how institutions have implemented these ideas. In 2003, Pam Hackbart-Dean and Sammie L. Morris provided reprocessing case studies of the ever-growing collections of Amelia Earhart at Purdue University and the International Association of Machinists and Aerospace Workers (IAM) at Georgia State University, which touch on many of the issues involved in the Wade Hall consolidation. They outlined four arrangement options for growing collections and the positive and negative outcomes of each. The reprocessing of the Earhart collection mirrors the first option of archivists physically and intellectually integrating additions to expanding collections, while the reprocessing of the IAM collection demonstrates the second option, where archivists intellectually integrate additions within a finding aid but do not physically integrate the material. Another option is to keep new additions physically separate and only add descriptions to the appendix of the finding aid. The final option is to treat additions as separate collections, with each collection having its own collection number, finding aid, and physical containers. That is what had happened in the case of the small Wade Hall manuscript collections. Hackbart-Dean and Morris found that understaffed institutions used this arrangement most often, but identified this approach as the least user-friendly, as researchers must examine multiple collections and disparate finding aids to locate records

of a single creator. Although the four options presented are extremely useful for institutions considering physical reprocessing projects, Hackbart-Dean and Morris skirted the issue of how reprocessing affects previously digitized collections, only mentioning the lack of location information within the finding aid and metadata of digitized items and not going any further with the discussion. The Wade Hall Consolidation Project seeks to fill this gap by addressing how institutions might deal with existing location information in the physical and digital environment as well as larger issues of digital preservation.¹⁵

The University of Alberta's Prairie Provinces Collection is similar in scope to the collections of materials donated by Wade Hall, as it also includes printed, photographic, and manuscript materials that continued to arrive as both individual pieces and small batches over the years. While the previously processed Wade Hall small manuscript collections had legacy descriptions, the staff at the University of Alberta had not accessioned or described their Prairie Provinces ephemera, which allowed them to choose LibGuides as the organizational scheme for the finding aid. The staff initially arranged materials into broad categories (i.e., Photographs, Letters, Memoirs, and Posters), but found a need for additional categories and subcategories to describe this collection adequately. Ultimately, LibGuides proved to be an unwieldy platform, and the staff migrated the content to a traditional finding aid. The University of Alberta's experiences informed the University of Alabama archivists' decision to create separate subject-based collections instead of an overarching Wade Hall manuscript collection with multiple subject-based collections within it.¹⁶

Dickinson College houses one- to two-item collections in filing cabinets, with the only access point being a subject-based card catalog in the reading room. The staff adapted a blog detailing reference requests with tags to create a "catablog" for individual items with description, metadata, and digitized images to create awareness of these collections online and to allow search engines to index them. Web analytics and external linking to the blog from outside websites demonstrate the blog's popularity, and, in turn, its popularity leads to increased reference requests and research visits. Seeing how the tags (e.g., Civil War) within the blog produced artificially created subject-based collections, the staff created separate subject-based blogs, which drew the attention of a noted historian who found them through a search engine. Through search engine indexing, the file cabinets are discoverable.¹⁷ As Dickinson College's example proves, greater discoverability of these small Wade Hall manuscript collections is possible through enhanced description, rearrangement into subject-based collections, and metadata harvesting and search engine indexing of ArchivesSpace and CONTENTdm.

One of the most relevant case studies examines a project that incorporated legacy data in the creation of a subject-based digital collection at the

National Agricultural Library. Christian James and Ricardo L. Puzalan's definition of "legacy data," in which existing data is reused and transformed to create a new collection of data, applies to the way our Wade Hall consolidation team took legacy descriptions and metadata and repurposed them to create new finding aids and metadata for the subject-based Wade Hall manuscript collections. In their 2015 study, James and Puzalan concluded that the inconsistency of the subject terms, different vocabularies, and application of standards in legacy data diminishes the ability of users to search effectively and that contextualization of a subject-based collection allows researchers to make interpretational connections when items are placed side by side instead of organized as disparate collections.¹⁸

CASE STUDIES—LEGACY DATA

The majority of the literature addresses legacy data in connection with a library's adoption of or migration to a specific management system, whether ArchivesSpace, a new digital asset management system (DAMS), or something else.¹⁹ Although these systems may be describing the same collection, an archival collection's finding aid often does so through folder-level description, while item-level metadata is typically the standard for digital collections. In 2013, Jane Zhang and Dayne Mauney sampled the depiction of both archival and digital collections online and identified the embedded, segregated, and parallel models of how institutions represent archival and digital collections.²⁰ This interplay between systems that manage physical collections (i.e., ArchivesSpace) and digital collections (i.e., CONTENTdm) that Zhang and Mauney identified is an issue specifically addressed in the Wade Hall Consolidation Project, but rarely mentioned in discussions of legacy data within the literature. Digital preservation of legacy metadata receives little discussion. In 2014, Todd Bruns, Stacey Knight-Davis, Ellen Corrigan, and Steve Brantley alluded to the challenges of dealing with legacy data and files, including digital preservation concerns, as their institution migrated digital content to a new institutional repository.²¹ This case study of the Wade Hall Consolidation Project addresses legacy metadata and the difficulties inherent with existing digital preservation practices by giving examples of how one institution dealt with this problem during physical and digital consolidation. Jane Darcovich, Kate Flynn, and Mingyan Li remediated legacy metadata from existing digital asset management systems in preparation for a migration as well as inclusion in the Digital Public Library of America (DPLA), a leading example of a multi-institutional digital content aggregator.²² The necessity of maintaining the older uniform resource locators (URLs) for materials already in aggregated sites is another consideration during migration. Persistent URLs (PURLs) are powerful and necessary tools in a shifting digital

environment.²³ The conscious choice to leave metadata as-is for the consolidated Wade Hall digital collections that already exist in multi-institutional digital content aggregators, and rely on the persistent URLs to point users to the new digital asset management system, represents a contribution to the existing literature and a possible strategy to save institutional resources and time.

Consolidation Project: Background and Timing

Why start the Wade Hall Consolidation Project at this particular moment? The overarching impetus for our project was the University of Alabama's focus on becoming a nationally recognized research university with the goal of becoming a Carnegie "R1" research university, which it achieved for the first time in 2018. Responding to this university initiative to strengthen innovation and scholarship, the library addressed staffing, skills, and emerging needs within Special Collections to promote a successful learning and research environment. Three factors converged to set the scene for the Wade Hall project—namely, expanded personnel, migration to a new digital asset management system, and the need to identify users and search strategies.

PERSONNEL

In 2017, the Special Collections department hired a new associate dean who was tasked with supporting research and enhancing the reputation of the unit. This meant hiring two additional processing archivists, restructuring and filling the open special collections and digital initiatives librarian position, and changing the digital asset management system. These additions gave the unit enough resources to take on the consolidation project. In hiring new staff for its Special Collections department, university libraries directly confronted the issue of bringing the university in line with the evolution of professional standards for archivists. Except for one professionally trained archivist, all the other staff had transitioned from archives-adjacent areas, such as library cataloging, or learned everything on the job.²⁴ The newest hires were recent graduates from archival programs where they had been trained on the most current standards, descriptive terminology, and technology used in archives. Moreover, the newer processing archivists had also been trained in archival theory and were aware of an emerging trend in subject categorization, which they were able to bring to the job to enhance what had already been done—all the while stepping lightly so as not to criticize the work of the archivists who had come before them.

MIGRATION TO A DIFFERENT DIGITAL ASSET MANAGEMENT SYSTEM

The University of Alabama Libraries developed an open-source digital asset management system, Acumen, in 2010. Over the years, Digital Services staff digitized 18TB of content, including numerous small Wade Hall collections, and placed it in Acumen. In October 2018, University Libraries selected the cloud-hosted version of CONTENTdm as its new system and publicly launched a small selection of migrated digital collections in February 2019. The decision to move to a vendor-based system arose from several factors, including cost, level of technological development needed, the loss of the original personnel who developed and maintained Acumen, and shifts toward research and accessibility in the library's strategic plan. The timing of the Wade Hall Consolidation Project meant that Digital Services staff could digitally consolidate collections and migrate content at the same time. The digital consolidation also addresses the issue of having too many existing digital collections to fit into the finite number of digital collections CONTENTdm's mobile responsive interface could accommodate without sacrificing appearance and usability. In 2018, CONTENTdm's mobile responsive interface could only technically support 400 digital collections, which was problematic because the University of Alabama had 600 digital collections slated for migration, of which 287 were Wade Hall manuscript collections.²⁵ Finally, the mirroring of the analog and digital collections, with their subject arrangement, will enhance discoverability in an online environment.

USERS AND SEARCH STRATEGIES

In support of the university's goal to become a top research institution, the libraries' strategic plan called for the development of innovative library instruction and services for varied users, provided through multiple channels. To fulfill this objective in the strategic plan, Special Collections needed to identify its users and their search strategies before it could determine the best way to alter its current research. As the University of Alabama is a public university, users included community members doing local history or genealogical research, visiting scholars, graduate students, and administrators researching university-specific records. The majority of physical visitors are undergraduate students. Although analyzing exact user statistics on each of the small Wade Hall collections is outside the scope of this article, general experience shows that users searching for collections of the type that Wade Hall's materials represent—that is, collections that are not about a prominent individual or family—tend to search by subject. All the Special Collections staff who assist users in the reading room have had encounters with students who ask for resources about “civil rights” or “slavery” or “women at the university,” for example. Even

before the widespread use of the Internet, Mary Jo Pugh pointed out that subject knowledge about records could not simply be located within the mind of a reference archivist but should exist somewhere attached to the record so that a user might be able to access that knowledge independently.²⁶ Today, patrons often investigate a repository's resources online before they make a physical visit—which gives Pugh's idea even more importance and validates the reasons behind the Wade Hall Consolidation Project.

Methodology

Our methodology outlines the step-by-step processes we followed during this project, as well as how our institution dealt with challenges including legacy description, unclear provenance, digital preservation, and metadata harvesting. Before undertaking a project of this size and scope, the team allocated personnel, created a project timeline, formulated selection criteria, and surveyed existing Wade Hall manuscript collections. After identifying candidates for the subject-based physical and digital consolidation, the team drafted workflows and altered them as problems arose throughout the project.

PERSONNEL

Under the leadership of Special Collections' associate dean, the Wade Hall Consolidation Project began with a team consisting of three processing archivists, the reference services and outreach coordinator, the archival access coordinator, and two staff members in Digital Services, the segment of Special Collections that oversees the digital asset management system. The newly hired special collections and digital initiatives librarian joined the team a few months into the project and serves as the liaison between the archivists and Digital Services.

PROJECT TIMELINE

Special Collections' associate dean provided the team with a two-year timeline to consolidate the items that Hall donated into broader collections organized by category, with the disparate finding aids simplified into one record for each consolidated unit. Digital consolidation would follow once physical consolidation was complete for each collection. At the first meeting, the team established two phases of the project: the planning and preparation phase and the consolidation phase. The planning and preparation phase took approximately five months. The reference services and outreach coordinator identified the

collections that would require consolidation (two months), and the archivists created a workflow and sorted the data into subject-based collections (three months). Once the planning and preparation were complete, the archivists embarked on the consolidation phase, spending 15 percent of their time (six hours per week) on the project for nineteen months. Digital Services spent one month creating metadata revisions templates and adopting existing migration procedures and migrated its first consolidated collection eleven months into the two-year timeline, with a new consolidated collection migrated each month after that point.

CRITERIA AND SURVEY OF WADE HALL COLLECTIONS

The first step was to identify the relevant Wade Hall collections that needed consolidation, including the troublesome one- and two-item collections as well as other collections under one linear foot. Using a native bulk export of JSON-formatted data from ArchivesSpace, Special Collections' reference services and outreach coordinator isolated Wade Hall collections and cleaned the data up to make them useful for analysis. ArchivesSpace's Extent field provided the size of each collection in linear feet, but this measurement did not adequately represent the number of items in small collections. Using the manuscript number, she searched Acumen for an item count for existing digitized items and used the finding aids to determine the collections' true size. She created a size column and entered the number of items, or used "M" (for "many") to designate more than five items in a collection. To aid the team in developing categorical themes, she scraped the JSON data to pull dates and subjects, and she began sorting collections. Through this process, she identified 1,429 physical collections, 287 of which were already digitized.

Using the JSON data, she identified groups based on the collections' subject headings, such as African American, Civil War, and Travel. For those collections lacking subject headings, she identified categories, such as diaries, scrapbooks, and business and finance, and grouped similar collections together. Finally, she took this data and organized them into an Excel spreadsheet with three section tabs: "Just 1–2 letters" collections; "Just 1–2 other" collections that are not letters; and "The rest," which contains collections larger than two items or an unknown number of items (see Figure 1). Although the data may not have included all of the small, consolidation-worthy Wade Hall collections, they became a starting point for the processing archivists, and the spreadsheet served as the master for the project.

26	MSS.2715	Delafield, F	Richard Delafield Letter	1
27	MSS.2377	Welch, Syl	Sylvester Welch Letter	1
28	MSS.2872	Pinnell, J	J. Pinnell Letter	1
29	MSS.2669	Mathews, T	Thomas Mathews Letter	1
30	MSS.1716	Beatty, V.	V. Ormond Beatty letter	1
31	MSS.2847	Anna to Sa	Letter from Anna to Sarah	1
32	MSS.1852	Deforn, Jol	John F. Deforn letter	1
33	MSS.1856			
Just 1-2 letters				Just 1-2 other
				The rest

FIGURE 1. Initial identification of potential physical collections for consolidation in the master spreadsheet

PHYSICAL CONSOLIDATION PLANNING

Such a project involves considerable staff time and resources; the small creator-described collections that already existed were still findable but did not reflect Wade Hall’s intent as the collector. Previous processors involved in creating records for the Wade Hall materials had named all of these small collections after their creators (the “origin”), as in “Samuel Holt Letter” for a single letter written by Holt, with its own resource record and collection number in ArchivesSpace. Upon reexamination of existing processed Wade Hall collections, archivists noted that Wade Hall had often written notes in pencil on the materials he had donated saying things like “Prohibition” or “World War II”—indicating that he intended them to be assembled according to a certain category. The team began to see that providing context to these records meant looking at provenance differently. Instead of seeing the provenance of the Samuel Holt Letter as coming from Holt, a man with no relation to the university and who did not convey his records to the archives here, they saw it as a piece of American life given to us by Wade Hall—Hall’s intentions became the context that gave meaning to the record, showing it as one of many letters describing American southern business transactions. In other words, the earlier processors were not wrong to categorize materials by the creator the way they did, but it could be said that they did not provide the full picture of the context that gives the records meaning and shows how they are assembled.

The previous arrangement of these discrete Wade Hall collections made it difficult for researchers to discover relevant resources in ArchivesSpace. For example, before the consolidation project began, Special Collections held over a hundred small collections from the World War II era. Now these collections all exist as one subject-based unit—the Wade Hall Collection of World War II Materials—that makes it possible for researchers to find all the items within ArchivesSpace or CONTENTdm through a single search. With the consolidation,

people can better gauge the volume of archival holdings on a particular subject, and it is easier for library staff to direct researchers toward other collections that might relate to their topics of interest. For instance, the newly consolidated Wade Hall Collection of World War I Materials complements Special Collections' separate, non-Wade Hall collection of World War I Posters.

One of the archivists' biggest concerns was to preserve necessary information from the original finding aid data. To accomplish this, they added general notes about the old collection numbers and titles to the consolidated collection's finding aid. This would allow researchers who had been familiar with the old collections to search with the same terms they used before and still find the collections they need within their new settings. The team added primary subject headings common among the old collections to the collection-level metadata of each new grouping and left any collection-specific subject headings at the item level. The archivists then streamlined any substantive information in the Abstract, Biographical/Historical note, and Scope and Contents note into a single file-level note in the new metadata. Reexamining the original finding aid data manually did take time, but proved valuable, allowing the archivists to reassess the previous wording and make judgments about its accuracy by comparing the information in the finding aid to the item in-hand.

Physical Consolidation: Specific Challenges

Overall, the physical consolidation went smoothly, though the processing archivists identified two particular challenges: unclear provenance and legacy description.

UNCLEAR PROVENANCE

Over fifty years ago, Barbara Kaiser profiled donors, like Wade Hall, who donate in varying amounts and over a long period of time. She posited that the type of relationship initially established between the repository and the donor will determine "to a considerable extent the number and nature of ensuing problems." Kaiser advised repositories to communicate their policies to donors early in the donation process.²⁷ Upon first establishing a relationship with a donor, an institution should clarify, among other things, how soon a donated collection is likely to be processed, whether the institution will discard unwanted materials, what sort of storage it can provide, and the kinds of restrictions it will apply to the donated materials. Some of the problems Special Collections archivists had as they tried to reorganize Wade Hall's donated materials stemmed from earlier curators not clarifying these policies at the time of intake—particularly in the area of restrictions on sensitive materials.

Wade Hall generally purchased his collections from estate sales, flea markets, auctions, rare materials dealers, and other sources. He rarely had any direct connection with the creators of these items—which he acquired largely in Kentucky and the lower Midwest, although the materials came from all over the United States and beyond. Some of the letters in the collection are relatively recent items from the 1980s and 1990s that contain private information from writers who had no idea their discarded correspondence would find its way into a public archival collection at the University of Alabama. While University Libraries has deeds of gift for the materials from Hall, Special Collections also realizes that the murky provenance of these items requires a unique approach to the repository's usual restrictions practice. To respect the privacy of the creators and their family members, Special Collections established a policy that restricts access to sensitive Wade Hall materials for seventy-five years from the date of creation. Previous archivists had restricted some items, though no consistent policy existed on how long the restrictions would apply. Throughout the consolidation process, Special Collections staff reexamined previously restricted Wade Hall materials to determine if the restrictions should be retained or if they could now be lifted.

LEGACY DESCRIPTION: INACCURACIES AND ASSUMPTIONS

The team saw many inconsistencies in the ArchivesSpace records, but time had never been allocated to correct them. Such inconsistencies came about when staff imported records into ArchivesSpace from Archivist's Toolkit, or they existed because information input by earlier staff and students lacked uniformity and standardization, and because processing procedures had changed over the years. Because almost every item was reprocessed, the archivists corrected inaccuracies and assumptions found in the old finding aids. For example, some past finding aid authors hypothesized, without verification, that items were

<p>Original (collection-level) Scope and Contents note from the finding aid for Phillip Thompson Letter, MSS.1687:</p> <p>This collection consists of a single letter addressed to Thompson; the letter itself is so badly damaged that the author's signature is missing. The content relates to a court case involving the author's estate and his relations with his children. There is a reference to the town of Morganfield, suggesting that it was written in Kentucky, but no evidence beyond that as to either the author's or the recipient's location. The author's grammar is very poor, and deciphering the letter is a considerable challenge.</p>
<p>New Scope and Contents note for Letter to Phillip Thompson, now a file within the consolidated Wade Hall Collection on Business and Finance, MSS.4267:</p> <p>This file contains a letter addressed to Phillip Thompson. The letter, which is missing a signature, relates to a court case involving the author's estate. The top of the letter bears a penciled note written by donor Wade Hall that mentions Morganfield, Kentucky.</p>

FIGURE 2. Original and updated Scope and Contents notes for the Letter to Phillip Thompson, showing changes to the inaccurate and presumptive description

produced in certain locales. Archivists removed such speculations, allowing researchers to determine the locale by examining the evidence on their own. They also removed editorial commentary such as “the handwriting is unintelligible,” “poor grammar,” “spelling is atrocious,” or “the language of the letter could not be determined,” and substituted neutral language throughout the description. One such example was the Phillip Thompson Letter (see Figure 2). The processing archivist made several changes to its Scope and Contents note, such as removing gendered references to the unidentified letter writer (i.e., “his” children).

The team also used this opportunity to identify the language of materials wherever possible and to correct typos, misspelled names, wrong dates, and other mistakes in the legacy descriptive metadata. This consolidation project was a chance to revisit every item and make sure it is described according to current archival standards.

LEGACY DESCRIPTION: LOCAL SUBJECTS

In the past, Special Collections created its own nonstandardized local broad subject headings, such as “Architecture and Landscape” and “Civil Rights and Human Rights,” which staff members sometimes used in their finding aids instead of authorized Library of Congress subject headings. The consolidation project archivists, upon finding a suitable substitute, deleted these local subject headings and replaced them with authorized Library of Congress subject headings as new finding aids were created in ArchivesSpace for the consolidated collections. By standardizing subject headings and names, the archivists intended to make it easier for future researchers to locate collections similar to one another.

“Daily Life and Family” was the most problematic catch-all local subject heading as it appeared in more than 600 of these collections, and it was often the only subject heading. The content of these collections did not fit into a clearly defined category. The overwhelming majority of the “Daily Life and Family” materials were letters in which ordinary Americans discussed their everyday life in a certain place and time. The solution was to create a new umbrella collection called the Wade Hall Collection on American Life, dividing it into geographical series (i.e., Northeast Region, Midwest Region) based on the creator’s location and arranging it in chronological order. This new arrangement provides users with examples of American life in a particular geographic region or during a specific time and allows for easier cross-comparison.

Physical and Digital Consolidation Workflows

PROCESSING PLAN/WORKFLOW

Once the archivists had time to go over the master spreadsheet produced by Special Collections' outreach coordinator and the accompanying notes, the team created two working documents for brainstorming ideas—one focused on workflow and the other on subject categories. Using the workflow document, the team proposed different approaches to this large project, outlined the responsibilities of each team member, and discussed potential rearrangement problems. After identifying frequently occurring categories from the master spreadsheet, the team recorded them in the second document and culled it to create the official list for the subject-based, consolidated collections. The team identified fourteen categories (e.g., Civil War, Diaries and Scrapbooks, and Travel and Tourism) that would form the consolidated collections. These categories came from both the assigned subject headings associated with the original small collections and the categories identified by Special Collections' outreach coordinator during the creation of the master spreadsheet.

The two brainstorming documents enabled the team to establish an official workflow, which consisted of two phases: planning and consolidation. The planning phase involved the redistribution of collection data into the new subject-based collections. The team created a new spreadsheet for each subject-based collection with tabs that aligned with the master spreadsheet, plus an additional tab for restricted items, and moved each original collection to one of the subject spreadsheets using the provided metadata.

After creating new subject-based collections, the archivists moved into the consolidation phase and prepared to distribute the materials for physical arrangement. Instead of dividing all the collections at once, the team decided to start small, with each archivist taking on two collections from the subject list. This choice allowed the team to develop best practices, estimate timeframes for processing collections, and discover other challenges that could surface during the intellectual arrangement and creation of finding aids in ArchivesSpace. Each new subject-based collection required the archivists to pull materials from various locations in the stacks, organize the materials both physically and intellectually into the new subject-based collection, and transfer pertinent finding aid data from the previous small collections into the new, subject-based finding aid. The archival access coordinator then edited, approved, and published the new finding aids. Once physical consolidation was complete, digital consolidation was the final step in the project.

Digital Consolidation: Issues

Digital consolidation had its own set of challenges. For all digital content, Digital Services utilizes a standardized file-naming convention, which incorporates a letter and a four-digit number based on the type of material (manuscripts would be u0003), followed by the manuscript collection number (1987), to create a unique digital identifier (u0003_0001987 at the collection level and u0003_0001987_0000001 at the item level). The collection-level identifier is the unique key to locating and managing everything associated with a digital collection, including metadata, digitized images, persistent URLs, and digital preservation. Little overlap exists other than a shared set of collection numbers between the digital collection’s collection-level identifier (u0003_0001987) and the physical collection’s finding aid in ArchivesSpace (MSS.1987). The only real link between the finding aid and the digital collection is the EAD Identifier. Before Special Collections had a public-facing version of ArchivesSpace to display finding aids, Digital Services took an exported EAD and used a script to transform it into a PDF finding aid for display within Acumen’s interface. The archivists added information into the EAD Identifier field in ArchivesSpace. Within the finding aid, the EAD Identifier (u0003_0001987) served as the bridge between the physical and digital collections, as it incorporated the collection number to serve as a collection-level identifier in Acumen.

FINDING AIDS

Throughout the years in Special Collections, processing archivists most often described collections to the folder level and used *Describing Archives: A Content Standard* (DACS) to govern descriptive format, while Digital Services included item-level metadata description using a defined metadata schema (MODS, Dublin Core, etc.) within its digital collections. As part of the physical consolidation process, archivists planned to create one single collection number for each subject-based collection, thus eliminating all the previous identification numbers for the smaller individual collections under its umbrella and breaking the fragile link between the physical and digital that had existed through the collection number and the EAD Identifier in ArchivesSpace. The problem was that the EAD Identifier field was only available at the collection level, not at the file level within ArchivesSpace (see Figure 3). The compromise

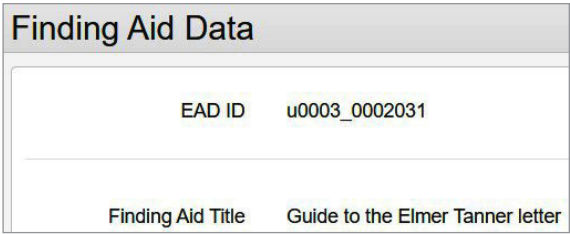


FIGURE 3. Use of the EAD Identifier in the original finding aid

The image shows a web-based interface for a digital finding aid. It features a 'General' section with the following fields: 'Persistent ID' (ac0aa824f9c683afbdf1f337e68bdae3), 'Label' (Local Identifier), 'Type' (General), and 'Publish?' (False). Below this is a 'Sub Notes' section, which contains a 'Text' note. The note's content is 'u0003_0002031'. At the bottom of the note, there are two tabs: 'Raw' and 'Formatted', with 'Formatted' being the active tab.

FIGURE 4. Inclusion of the Local Identifier as a note within the finding aid

was to put a general note at the file level in the finding aid and rename its Local Identifier, a name already in use by Digital Services to describe its digital content (see Figure 4).

FILE STORAGE AND DIGITAL PRESERVATION

Now that the new consolidated collection finding aid maintained the Local Identifiers for previous individual collections, Digital Services needed to determine how the introduction of a new collection number for a consolidated collection affected its file storage of images and metadata on the server and the digital preservation of previous individual digital collections, each with its own collection-level identifier. To understand what changing the collection number would do, it is important to examine the current file storage and digital preservation practices already in place. As part of its digital preservation practices, the University of Alabama Special Collections engages in bit preservation by managing multiple copies of the data, abiding by fixity, and ensuring the data are secure from corruption, damage, or deletion.²⁸ After the creation of a new digital collection, Digital Services stores images and metadata on a network drive before moving the digital content onto a Linux server. Once uploaded to the server, Digital Services generates MD5 checksums for the digital content as a means of ensuring the data are the same and monitors the fixity of the content.²⁹ Digital Services ingests the digital content into the Alabama Digital Preservation Network (ADPN), the statewide LOCKSS network, which replicates the data to seven storage sites.³⁰ With the physical consolidation uniting items thematically

under one collection number, the new collection number did not match the old collection number of the existing digitized materials on the server and within ADPN. The choice was either to rename the digital files and metadata and create a duplicate set of files with the new collection number within the server and ADPN, or to leave the digital files and metadata stored under the original collection number. The first option violates best practices of digital preservation, especially fixity. Renaming the files and consolidating the preserved metadata files into one set of metadata files would be time-intensive and prone to human error. Copying and transferring files makes them more susceptible to damage and possible loss of bit-level integrity. This option would also waste storage space with the creation of duplicate files. The second option was to leave digital files and metadata under their original collection number, which seemed the least disruptive option for file storage and digital preservation. With the choice of the second option, Digital Services had to determine how to tie the new collection number back to the original collection number with a “breadcrumb” to navigate within its file storage and digital preservation systems.

After deciding that each individual digitized Wade Hall collection would keep its existing collection-level identifier within its file storage and digital preservation systems, Digital Services brainstormed how to incorporate a new collection number into its metadata template while retaining the original collection’s relevant metadata. A new metadata field, the Parent Collection Identifier, will house the new collection number and will serve as a unifying metadata field for previously digitized collections and any materials from the consolidated collection that are digitized in the future. As materials were physically consolidated into new subject-based collections, the archivists created a spreadsheet (see Figure 5) to update the following metadata fields: Collection Name, Item Location (Box and Folder), Parent Collection Identifier, and Finding Aid URL. The

Old Collection Name ▼	New Collection Name ▼	Item Location ▼	Parent Collection ▼	Local Identifier ▼	Finding Aid ▼
Bill H. Axelby Letter	Wade Hall collection of World War II materials	Box 4253.004, Folder 016	u00003_0004253	u00003_0002039_0000001	https://archives.lib.ua.edu/repositories/3/recources/4836
Harold E. Little letter	Wade Hall collection of World War II materials	Box 4253.004, Folder 015	u00003_0004253	u00003_0002040_0000001	https://archives.lib.ua.edu/repositories/3/recources/4836
Wyndham G. Allen letters	Wade Hall collection of World War II materials	Box 4253.001, Folder 014	u00003_0004253	u00003_0002042_0000001	https://archives.lib.ua.edu/repositories/3/recources/4836

FIGURE 5. Archivists’ spreadsheet of updated metadata fields of Collection Name, Item Location, Parent Collection Identifier, Local Identifier, and Finding Aid for three collections

Local Identifier represents the unique identifier of a collection at the item level, which Digital Services uses to locate all digital content filed under its previous collection number.

Digital Services will use the existing item-level metadata from each individual collection (i.e., Bill H. Axelby letter, u0003_0002039) and add in the updated metadata fields from the archivists' spreadsheet for the new collection (i.e., Wade Hall collection of World War II materials, u0003_0004253) (see Figure 6).

Digital Services will archive the updated metadata under its former collection number, with the Parent Collection Identifier field preserving a crosswalk from the legacy metadata in the old collection to the new collection.

METADATA HARVESTING AND MULTI-INSTITUTIONAL AGGREGATORS

Several collaborative websites, including Alabama Mosaic, a statewide digital collection consortium EBSCO's Electronic Discovery Service (EDS); WorldCat;

Title	Letter from Bill H. Axelby, Fort Benning, Georgia, to Mother, Jersey City, New Jersey, November 6, 1944
Creator	Axelby, William H. (Bill H.) (Correspondent)
Contributors	Axelby, F. D., Mrs. (Addressee)
Type	Text
Date	1944-11-06
Language	eng
Format	electronic; image/jpeg; 3 p.
Provenance	Gift of Wade Hall, 2007.
Funding Note	The digitization of this collection was funded by a gift from EBSCO Industries.
Subject	United States. Army; World War, 1939-1945; Axelby, William H. (Bill H.)--Correspondence; Axelby, F. D., Mrs.--Correspondence; letters (correspondence)
Coverage	Fort Benning (Ga.); United States--Georgia--Chattahoochee County
Repository	The University of Alabama Libraries Special Collections
Collection Name	Wade Hall collection of World War II materials
Parent Collection Identifier	u0003_0004253
Item Location	Box 4253.004, Folder 016
Local Identifier	u0003_0002039_0000001
PURL	http://purl.lib.ua.edu/117324
Finding Aid	https://archives.lib.ua.edu/repositories/3/resources/4836
Copyright and Terms	Images are in the public domain or protected under U.S. copyright law [Title 17, U.S. Code], and both types may be used for research and private study. For publication, commercial use, or reproduction, in print or digital format, of all images and/or the accompanying data, users are required to secure prior written permission from the copyright holder and from archives@ua.edu. When permission is granted, please credit the images as Courtesy of The University of Alabama Libraries Special Collections.

FIGURE 6. Consolidated metadata for Bill H. Axelby Letter that is now part of the Wade Hall Collection of World War II materials; arrows indicate new or updated fields

and Digital Public Library of America had harvested Acumen using Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH). The harvest included numerous Wade Hall digitized items. The Wade Hall Consolidation Project and the migration to CONTENTdm overlapped in terms of timing, and it was not advantageous to switch metadata harvesting to CONTENTdm until an overwhelming majority of the content had been migrated. Luckily, the University of Alabama Libraries have maintained persistent URLs to their digital content. After a digital collection has been migrated, Digital Services updates the PURL database to point to CONTENTdm, which provides a seamless experience for any user who clicks on a PURL anywhere on the Web. The most problematic website was Civil War in the American South (<http://american-south.galileo.usg.edu>), which the Association of Southeastern Research Libraries (ASERL) launched in 2011 to commemorate the sesquicentennial of the Civil War. The University of Georgia's Digital Library of Georgia hosts that website and, in turn, shared that metadata again when its institution became part of DPLA. These harvested websites contained only a specific subset of digital collections dealing with the Civil War, including collections that were part of the Wade Hall Consolidation Project. Digital Services weighed the amount of effort and time it would take to fix the metadata harvesting (with little or no return for both Digital Services and the personnel of these two websites) and determined it was not worth it (see Figure 7).

Although seemingly counter to best practices, this conscious choice not to fix metadata harvesting may be another option for institutions that use persistent URLs and are migrating to a new DAM.

Migration and Digital Consolidation: Workflow

Digital migration of a physically consolidated Wade Hall collection began with the process of setting up a new collection in CONTENTdm with the consolidated collection name and the finding aid's abstract from ArchivesSpace for the digital collection's main landing page. The collection landing page contains the longer Scope and Contents note as well as the hyperlinked finding aid in the public version of ArchivesSpace. Using the archivists' spreadsheets as a crosswalk between the old and new collection numbers, the next step was to pull the images and metadata for all consolidated collections using the previous collection-level Local Identifier to locate the files on the server. Digital Services staff ran several Python scripts to pull the digitized content in batches, convert the TIFs to JPEGs, and organize the image files into folders to prepare them for ingest into CONTENTdm. They chose to use JPEGs within CONTENTdm to reduce the amount of time to upload as well as to save on storage costs of this cloud-hosted platform. Digital preservation of the components of the digital

Letter from C. W. Hale to father, July 28, [1864]	
Item: 2 of 2 Prev Record Next Record Search Results Save Record	
Click here to view the item	
Creator:	Hale, C. W. Father
Title:	Letter from C. W. Hale to father, July 28, [1864]
Description:	A letter from C. W. Hale at Ft. Ward written to his father. Almost the whole of the letter relates to the health of Hale and his comrades in arms.
Types:	Letters (correspondence) Text
Subjects:	letters (correspondence) Hale, C. W.--Correspondence Father--Correspondence Soldiers--Virginia Fort Ward (Alexandria, Va.) United States. Army. Ohio Infantry Regiment, 67th (1861-1865)--Military life Soldiers--Health and hygiene War and families--United States United States--History United States, Virginia United States, Ohio
Collection:	C. W. Hale letter
Institution:	William Stanley Hoole Special Collections Library
Contributors:	William Stanley Hoole Special Collections Library (University of Alabama)
Original Material:	C. W. Hale Letter, William Stanley Hoole Special Collections Library
Rights and Usage:	http://rightsstatements.org/vocab/NoC-US/1.0/ To obtain permission to publish or reproduce, please contact the W. S. Hoole Special Collections Library
Related Materials:	http://american-south.org/ Prev Record Next Record

FIGURE 7. Comparison between the same item’s metadata in ASERL’s Civil War in the South (above) and CONTENTdm (below)

Title	Letter from C. W. Hale to father, July 28, [1864]
Creator	Hale, C. W. (Correspondent)
Type	Text
Date	1864-07-28
Language	eng
Format	electronic; image/jpeg; 4 p.
Abstract	A letter from C. W. Hale at Ft. Ward written to his father. Almost the whole of the letter relates to the health of Hale and his comrades in arms.
Provenance	Gift of Wade Hall, 2004
Funding Note	The digitization of this collection was funded by a gift from EBSCO Industries.
Subject	Hale, C. W.--Correspondence; Father--Correspondence; Soldiers--Virginia; United States. Army--Military life; Soldiers--Health and hygiene; War and families--United States; letters (correspondence)
Coverage	Fort Ward (Alexandria, Va.); United States--History--Civil War, 1861-1865; United States--Virginia--Fort Ward
Repository	The University of Alabama Libraries Special Collections
Collection Name	Wade Hall collection of Civil War materials
Parent Collection Identifier	u0003_0004273
Item Location	Box 4273.002, Folder 013
Local Identifier	u0003_0001646_0000001
PURL	http://purl.lib.ua.edu/20764
Finding Aid	https://archives.lib.ua.edu/repositories/3/resources/4873

Local Identifier	Reference URL					
u0003_0001644_0000001	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/82					
u0003_0001645_0000001	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/9					
u0003_0001645_0000002	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/47					
u0003_0001645_0000003	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/23					
u0003_0001645_0000004	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/42					
u0003_0001645_0000005	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/68					

FIGURE 8. CONTENTdm export of a previous one-item collection (u0003_0001644) and five-item collection (u0003_0001645) as shown by the Local Identifier of each item

collections takes place outside of CONTENTdm through archiving the content in ADPN. Another set of rules-based Perl and Python scripts took the existing MODS metadata and transformed it into Dublin Core metadata in a tab-delimited text file. Because this metadata is based upon the old collection arrangement, Digital Services staff edited the metadata in the Collection Name, Collection Number, and Item Location (Box and Folder) fields, and added the new metadata fields Parent Collection Identifier and Finding Aid URL. Posing the most difficulty to the digital consolidation was the Abstract metadata field, as this field sometimes describes at the collection level (“The John Doe collection . . .”) and sometimes at the item level (“John Doe wrote to Jane Doe concerning . . .”). Because the scripted metadata transformation process cannot discern the specificity of the Abstract field, Digital Services staff had to inspect the contents manually and make a decision to retain or purge the metadata in the field. Once the metadata was finalized, Digital Services staff uploaded the metadata tab-delimited text file and images to CONTENTdm, indexed the collection, and published it, making it live.

At this point in the process, the current collection- and item-level persistent URLs in CONTENTdm pointed to Acumen. To update the persistent URLs, Digital Services staff exported the CONTENTdm collection metadata and retained only the Local Identifier and the CONTENTdm reference URL metadata fields in a tab-delimited text file (see Figure 8).

Digital Services ran a Python script to match the Local Identifier from the export and the PURL database (see Figure 9) and to replace the existing Acumen URL with the CONTENTdm reference URL (see Figure 10).

Local Identifier	PURL	URL				
u003_0001644_0000001	http://purl.lib.ua.edu/20750	http://acumen.lib.ua.edu/u0003/00001644/0001				
u003_0001645_0000001	http://purl.lib.ua.edu/20751	http://acumen.lib.ua.edu/u0003/00001645/0001				
u003_0001645_0000002	http://purl.lib.ua.edu/20752	http://acumen.lib.ua.edu/u0003/00001645/0002				
u003_0001645_0000003	http://purl.lib.ua.edu/20753	http://acumen.lib.ua.edu/u0003/00001645/0003				
u003_0001645_0000004	http://purl.lib.ua.edu/20754	http://acumen.lib.ua.edu/u0003/00001645/0004				
u003_0001645_0000005	http://purl.lib.ua.edu/20755	http://acumen.lib.ua.edu/u0003/00001645/0005				

FIGURE 9. Simplified view of the PURL database showing that the persistent URL should redirect to the Acumen URL

Local Identifier	PURL	URL
u003_0001644_0000001	http://purl.lib.ua.edu/20750	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/82
u003_0001645_0000001	http://purl.lib.ua.edu/20751	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/9
u003_0001645_0000002	http://purl.lib.ua.edu/20752	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/47
u003_0001645_0000003	http://purl.lib.ua.edu/20753	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/23
u003_0001645_0000004	http://purl.lib.ua.edu/20754	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/42
u003_0001645_0000005	http://purl.lib.ua.edu/20755	http://cdm17336.contentdm.oclc.org/cdm/ref/collection/p17336coll58/id/68

FIGURE 10. Python script matches the Local Identifier from the CONTENTdm export and PURL database and replaces the existing URL with the CDM Reference URL

The updating of the persistent URLs ensured that researchers who previously cited University Libraries’ digital collections—such as the ones on ASERL’s Civil War in the South and DPLA—would redirect to CONTENTdm instead of Acumen. As it also simultaneously updates existing library websites and LibGuides that point to the digital collection, this process saves staff time that would otherwise be spent fixing broken links.

Discussion

Better planning would have made this consolidation project easier to execute in a shorter amount of time. The decision to update legacy description as needed was made fairly early in the project, but the archivists unintentionally limited the scope of their edits to correcting inaccurate metadata and typographical errors. As the archivists began rewriting finding aids, they quickly realized that some of the terminology used in the old Scope and Contents notes was dated and possibly offensive (e.g., the use of “slaves” rather than the term “enslaved persons”) or was speculative or subjective (e.g., the use of phrases such as “the language could not be determined” and “spelling is atrocious”). Bringing the language of the Scope and Contents notes to a more acceptable form was every bit as important as correcting inaccurate metadata and typographical errors, but it meant revisiting some previously completed collections.

One of the most time-intensive errors was not sorting all the collections listed on the master spreadsheet into the larger subject-based collections. Previous archivists had attempted to consolidate several Wade Hall collections already and had grouped them into larger collections such as the Wade Hall Vertical File, the Wade Hall Miscellaneous Letters Collection, and so on. None of these quasiconsolidated collections had any particular theme, and, together, they consisted of about 122 individual items. The miscellaneous letters collection was the largest combined collection, with seventy-five items. The only collection-level description was “Miscellaneous letters from across the United States and around the world,” along with a date range of 1837–1965. Originally, the archivists on the consolidation team put these themeless collections aside until later. About halfway through the project, one of the archivists examined

each item in detail, dated it, and described it in a Scope and Contents note, then dispersed the items into the subject-based collections where they belonged.

This change in the physical arrangement had ramifications for the archivists who had to edit finding aids for consolidated collections they had already completed, and for Digital Services staff, who had to edit the metadata, specifically the Item Location field (Box and Folder), essentially redoing the process again for each individual collection within a consolidated digital collection. The archivists reunited dispersed groups of letters by the same individual that people had previously filed in different folders or identified as miscellaneous fragments, each with its own finding aid and manuscript number, as in the case of the Menial Horton Kaiser Letters. Someone had once processed the last two pages from one of his letters as a separate collection and placed them in their own file. The problem was that by the time the archivists found the missing pages, the rest of the letters were in the Wade Hall World War II Collection. Digital Services deleted the existing letter in CONTENTdm, digitized the rediscovered pages, and revised the metadata before uploading it again to CONTENTdm. The archivists learned that this review would have been better performed during the planning phase rather than halfway through the consolidation phase. After having to redo a few migrated digital collections, Digital Services postponed archiving revised metadata until the conclusion of the consolidation project.

Conclusion

While this type and scale of project will probably not happen often in any repository's lifetime, the frequency rests on an institution's previously established relationship with long-term donors. A repository's relationship with donors depends on the rapport they establish early on, and it should be in line with the institution's collection development policy. As soon as it becomes apparent that a donor is planning multiple donations of similar materials over a considerable length of time, the repository should work with the donor to ensure the materials are organized in such a way as to make the collections useful and discoverable to researchers. All parties involved in the process should help determine detailed plans for the disposition of the materials that include guidelines, procedures, and expected outcomes. In terms of resources, Special Collections staff should determine the level of arrangement and description for donated collections based on the importance of a donation, the significance of the subject matter, and the availability of staffing and resources necessary for both processing and digitization.

The question becomes this: are the consolidated Wade Hall manuscript collections more discoverable by users? The new consolidated collections have not been publicly available long enough to gather viable statistics on their use,

which is a possible area of future research. However, the reference services and outreach coordinator says that she now finds it easier to pull together Wade Hall materials for her classes, thus exposing researchers to these materials.

This consolidation case study will help fill the gap in the literature concerning the consolidation and migration of physical and digital collections and perhaps provide some guidance on the issues of provenance, subject-based collections, legacy data, metadata harvesting, and digital preservation to others considering a large—or small—consolidation project. By consolidating these many small collections into larger, subject-based collections, the team moved closer to Wade Hall's original intent for his donations: to make these materials discoverable by researchers of the southern and American experiences.

NOTES

- ¹ For more on the scope of the Wade Hall collections, see Amy Hildreth Chen, *Miracle and Mystery: A Guide to the Wade Hall Collections*, <http://purl.lib.ua.edu/172988>.
- ² Jane Zhang, "Archival Context, Digital Content, and the Ethics of Digital Archival Representation," *Knowledge Organization* 39, no. 5 (2012): 332–39, <http://dx.doi.org/10.5771/0943-7444-2012-5-332>.
- ³ Quoted in Larry Weimer, "Pathways to Provenance: DACS and Creator Descriptions," *Journal of Archival Organization* 5, nos. 1–2 (2008): 33–48, https://doi.org/10.1300/J201v05n01_03. The "Introduction to Archival Authority Records" section in the second edition of DACS covers the same points. See Society of American Archivists, *Describing Archives: A Content Standard* (Chicago: Society of American Archivists, 2013, revised 2015), http://files.archivists.org/pubs/DACS2E-2013_v0315.pdf, captured at <https://perma.cc/3JNZ-AE8A>.
- ⁴ Richard Lytle, "Intellectual Access to Archives," *American Archivist* 43, no. 1 (1980): 64–75, <https://doi.org/10.17723/aarc.43.1.31031523737j16n7>.
- ⁵ Mary Jo Pugh, "The Illusion of Omniscience: Subject Access and the Reference Archivist," *American Archivist* 45, no. 1 (1982): 33–44, <https://doi.org/10.17723/aarc.45.1.21861730132n24vx>.
- ⁶ Elsie Freeman, "In the Eye of the Beholder: Archives Administration from the User's Point of View," *American Archivist* 47, no. 2 (1984): 111–23, <https://doi.org/10.17723/aarc.47.2.a373340078502136>.
- ⁷ Thomas J. Ruller, "Open All Night: Using the Internet to Improve Access to Archives: A Case Study of the New York State Archives and Records Administration," *The Reference Librarian* 26, no. 56 (1997): 161–70, https://doi.org/10.1300/J120v26n56_12.
- ⁸ Wendy Duff and Penka Stoyanova, "Transforming the Crazy Quilt: Archival Displays from a User's Point of View," *Archivaria* 45 (Spring 1998): 44–79; Heather MacNeil, "Picking Our Text: Archival Description, Authenticity, and the Archivist as Editor," *American Archivist* 68, no. 2 (2005): 264–79, <https://doi.org/10.17723/aarc.68.2.01u65t6435700337>; Max J. Evans, "Archives of the People, by the People, for the People," *American Archivist* 70, no. 2 (2007): 387–400, <https://doi.org/10.17723/aarc.70.2.d157t6667g54536g>; Cara S. Bertram, "Putting the User First: The Importance of the Reference Archivist in Online Projects," *Archival Issues: Journal of the Midwest Archives Conference* 40, no. 1 (2019): 7–23, <https://doi.org/archivalissues.11084>.
- ⁹ For more on EADs, see Kris Kiesling, "EAD as an Archival Descriptive Standard," *American Archivist* 60, no. 2 (1997): 344–54, <https://www.jstor.org/stable/40294442>; Anne J. Gilliland-Swetland, "Popularizing the Finding Aid: Exploiting EAD to Enhance Online Discovery and Retrieval in Archival Information Systems by Diverse User Groups," *Journal of Internet Cataloging* 4, nos. 3–4 (2001): 199–225, https://doi.org/10.1300/J141v04n03_12; J. Gordon Daines III and Cory L. Nimer, "Re-imagining Archival Display: Creating User-Friendly Finding Aids," *Journal of Archival Organization* 9, no. 1 (2011): 4–31, <https://doi.org/10.1080/15332748.2011.574019>.

- ¹⁰ Cory Nimer and J. Gordon Daines III, "What Do You Mean It Doesn't Make Sense? Redesigning Finding Aids from the User's Perspective," *Journal of Archival Organization* 6, no. 4 (2008): 216–32, <https://doi.org/10.1080/15332740802533214>. For more usability studies, see Morgan Daniels and Elizabeth Yakel, "Seek and You May Find: Successful Search in Online Finding Aid Systems," *American Archivist* 73, no. 2 (2010): 535–68, <https://doi.org/10.17723/aarc.73.2.p578900680650357>; Christopher Prom, "User Interactions with Electronic Finding Aids in a Controlled Setting," *American Archivist* 67, no. 2 (2004): 34–68, <https://doi.org/10.17723/aarc.67.2.7317671548328620>; Rachel Walton, "Looking for Answers: A Usability Study of Online Finding Aid Navigation," *American Archivist* 80, no. 1 (2017): 30–52, <https://doi.org/10.17723/0360-9081.80.1.30>.
- ¹¹ Elizabeth Yakel and Deborah Torres, "AI: Archival Intelligence and User Expertise," *American Archivist* 66, no. 1 (2003): 51–78, <https://doi.org/10.17723/aarc.66.1.q022h85pn51n5800>. For more on information-seeking behavior of different types of researchers, see Wendy Duff and Catherine Johnson, "Where Is the List with All the Names? Information-Seeking Behavior of Genealogists," *American Archivist* 66, no. 1, (2003): 79–95, <https://doi.org/10.17723/aarc.66.1.1375uj047224737n>; Helen Tibbo, "Primarily History in America: How U.S. Historians Search for Primary Materials," *American Archivist* 66, no. 1 (2003): 9–50, <https://doi.org/10.17723/aarc.66.1.b12037011g718n74>; Wendy M. Duff and Catherine A. Johnson, "Accidentally Found on Purpose: Information-seeking Behavior of Historians in Archives," *Library Quarterly* 72, no. 4 (2002): 472–96; <https://doi.org/10.1086/lq.72.4.40039793>; Donghee Sinn and Nicholas Soares, "Historians' Use of Digital Archival Collections: The Web, Historical Scholarship, and Archival Research," *Journal of the Association of Information Science and Technology* 65, no. 9 (2014): 1794–809, <https://doi.org/10.1002/asi.23091>.
- ¹² For more on the shift away from mediation in reference services, see Sigrid McCausland, "A Future without Mediation? Online Access, Archivists, and the Future of Archival Research," *Australian Academic & Research Libraries* 42, no. 4 (2013): 309–19, <https://doi.org/10.1080/00048623.2011.10722243>; Wendy Duff, "Archival Mediation," in *Currents of Archival Thinking*, ed. Terry Eastwood and Heather MacNeil (Santa Barbara, CA: Libraries Unlimited, 2010), 115–37.
- ¹³ Elizabeth Yakel, "Thinking Inside and Outside the Boxes: Archival Reference Services at the Turn of the Century," *Archivaria* 49 (2000): 140–60.
- ¹⁴ Jennifer Schaffner, "The Metadata Is the Interface: Better Description for Better Discovery of Archives and Special Collections, Synthesized from User Studies," in *Making Archival and Special Collections More Accessible* (Dublin, OH: OCLC Research, 2015), 85–97, <https://www.oclc.org/content/dam/research/publications/2015/oclcresearch-making-special-collections-accessible-2015.pdf>, captured at <https://perma.cc/2XZ4-VGYX>.
- ¹⁵ Pam Hackbart-Dean and Sammie L. Morris, "Case Studies in Managing Collections that Grow," *Archival Issues: Journal of the Midwest Archives Conference* 28, no. 2 (2003): 105–20, <https://www.jstor.org/stable/j40048364>.
- ¹⁶ Sharon Farnel, Robert Cole, Robert Desmarais, Spencer Holizki, and Jeff Papineau, "Where There's a Will There's a Way: Using LibGuides to Rescue Paper Ephemera from the Bibliographic Underbrush," in *Description: Innovative Practices for Archives and Special Collections*, ed. Kate Theimer (Lanham, MD: Rowman & Littlefield, 2014), 115–29; "Collection FC 3234.2 P732—Prairie Ephemera Collection," The University of Alberta Libraries, <https://discoverarchives.library.ualberta.ca/index.php/prairie-ephemera>, captured at <https://perma.cc/9LB3-ZJBP>.
- ¹⁷ James Gerenscer, "Opening the Black File Cabinets: Describing Single Items for Discovery and Access," in *Description: Innovative Practices for Archives and Special Collections*, 163–78.
- ¹⁸ Christian James and Ricardo L. Punzalan, "Legacy Matters: Describing Subject-Based Digital Historical Collections," *Journal of Archival Organization* 12, nos. 3–4 (2015): 198–215, <https://doi.org/10.1080/15332748.2015.1150104>.
- ¹⁹ For a few relevant examples of dealing with legacy data in systems, see Laura A. Knodel, "The Evolution of Archival Description at the Southern Historical Collection" (master's thesis, University of North Carolina at Chapel Hill, 2004), <https://doi.org/10.17615/3p52-3j03>; Christopher Prom, "Optimum Access? Processing in College and University Archives," *American Archivist* 73, no. 1 (2010): 146–74, <https://doi.org/10.17723/aarc.73.1.519m6003k7110760>; Kathryn Hujda, Caitlin Marineau, and Amanda Wick, "Maximum Product, Even Less Process: Increasing Efficiencies in Archival Processing Using ArchivesSpace," *Journal of Archival Organization* 13, nos. 3–4 (2016): 100–113, <https://doi.org/10.1080/15332748.2018.1443549>; William H. Hardesty, "Linear

- to Granular: Preparing Descriptions for an ArchivesSpace Migration,” *Archival Issues: Journal of the Midwest Archives Conference* 39, no. 2 (2019): 46–61, <https://doi.org/archivalissues.11074>; Anna Neatrou, Jeremy Myntti, Matt Brunsvik, Harish Maringanti, Brian McBride, and Alan Witkowski, “A Clean Sweep: The Tools and Processes of a Successful Metadata Migration,” *Journal of Web Librarianship* 11, nos. 3–4 (2017): 194–208, <https://doi.org/10.1080/19322909.2017.1360167>.
- ²⁰ Jane Zhang and Dayne Mauney, “When Archival Description Meets Digital Object Metadata: A Typological Study of Digital Archival Representation,” *American Archivist* 76, no. 1 (2013): 174–95, <https://doi.org/10.17723/aarc.76.1.121u85342062w155>.
- ²¹ Todd A. Bruns, Stacey Knight-Davis, Ellen K. Corrigan, and Steve Brantley, “It Takes a Library: Growing a Robust Institutional Repository in Two Years,” *College & Undergraduate Libraries* 21, nos. 3–4 (2014): 244–62, <https://doi.org/10.1080/10691316.2014.904207>.
- ²² Jane Darcovich, Kate Flynn, and Mingyan Li, “Born of Collaboration: The Evolution of Metadata Standards in an Aggregated Environment,” *Visual Resources Association Bulletin* 45, no. 2 (2019): 1–12, <https://online.vraweb.org/index.php/vrab/article/view/35/33>.
- ²³ For an overview of persistent URLs (PURLs), see William Y. Arms, “Uniform Resource Names: Handles, PURLs, and Digital Object Identifiers,” *Communications of the ACM* 44, no. 5 (2001): 68, <https://doi.org/10.1145/374308.375358>.
- ²⁴ “Guidelines for a Graduate Program in Archival Studies,” Society of American Archivists, 2016, <https://www2.archivists.org/prof-education/graduate/gpas>, captured at <https://perma.cc/4VD2-8WYM>. The guidelines also note that “Institutions often administratively connect archives, libraries, and special collections. Departments relating to the information science profession can share physical space, technical resources, staff, stakeholders, and skill sets.”
- ²⁵ CONTENTdm resolved this technical limitation in August 2019 to allow for unlimited digital collections but recommends 500 digital collections for optimum performance of its mobile responsive interface.
- ²⁶ Pugh, “The Illusion of Omniscience.”
- ²⁷ Barbara Kaiser, “Problems with Donors of Contemporary Collections,” *American Archivist* 32, no. 1 (1969): 103–7, <https://doi.org/10.17723/aarc.32.2.tg7m216122163338>.
- ²⁸ For an overview of the basics of digital preservation, see Trevor Owens, “Managing Copies and Formats,” in *The Theory and Craft of Digital Preservation* (Baltimore: John Hopkins University Press, 2018), 103–27.
- ²⁹ Paula De Stefano et al., “Checking Your Digital Content: What Is Fixity, and When Should I Be Checking It?,” NDSA, 2014, <http://hdl.loc.gov/loc.gdc/lcpub.20136551171>.
- ³⁰ For more on ADPN and LOCKSS, see Aaron Trehub, Corey Davis, Mark Jordan, Cinda May, and Sam Meister, “LOCKSS Networks Community-Based Digital Preservation,” in *Digital Preservation in Libraries: Preparing a Sustainable Future*, ed. Jeremy Myntti and Jessalynn Zoom (Chicago: ALA Editions, 2019), 217–40.

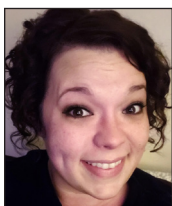
ABOUT THE AUTHORS



Laura M. Gentry is the special collections and digital initiatives librarian at the University of Alabama Libraries Special Collections. She received her MLIS and a graduate certificate in records and information management from Wayne State University in 2014 and an MA in history from the University of Alabama in 2008. Previously, she worked as a special collections librarian at the Birmingham Public Library.



Erin Ryan is a processing archivist at the University of Alabama Libraries Special Collections. She received her MSLS, with a concentration in archives and records management, from the University of North Carolina at Chapel Hill in 2017. Previously, she worked as an editor for *Celebration* magazine, a liturgical resource, and wrote articles for the *National Catholic Reporter* and the *Charlotte Observer*. She has an MA in church history from Saint John's School of Theology in Collegeville, Minnesota.



Jessica Rayman is a processing archivist at the University of Alabama Special Collections. She received her MSLS with a concentration in archives and records management from the University of North Carolina at Chapel Hill in 2017. She also holds a BA in history from Georgia Gwinnett College.



Martha Bace was a processing archivist until March 2020 when she retired after eighteen years in the University of Alabama Libraries. She received her MLS from George Peabody College for Teachers in Nashville, Tennessee, in 1979. Previously, she worked as the technical services librarian at Southern Arkansas University and as a cataloging and reference librarian at the University of Montevallo. She holds a BA in history from Austin College in Sherman, Texas.