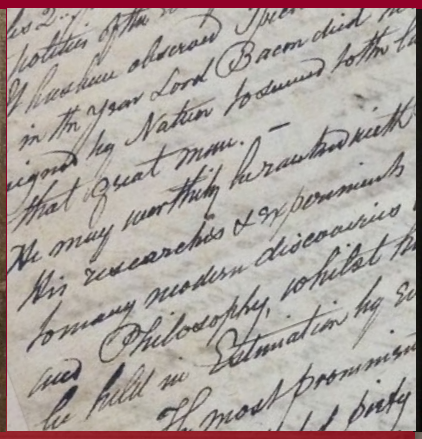
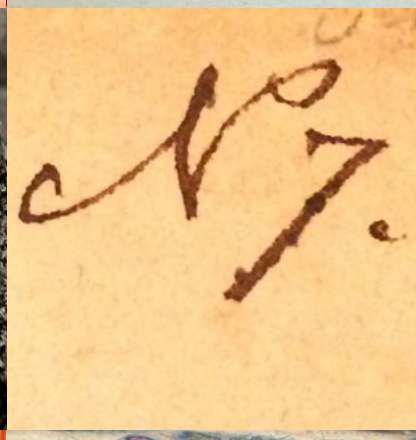
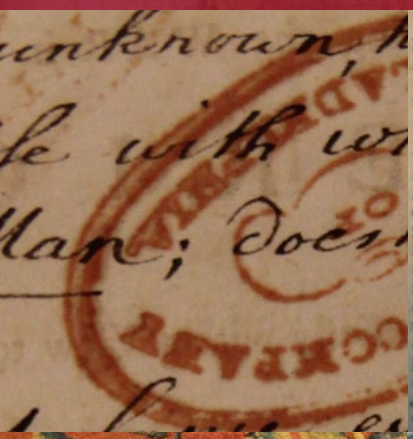
INNOVATION, COLLABORATION *and* MODELS

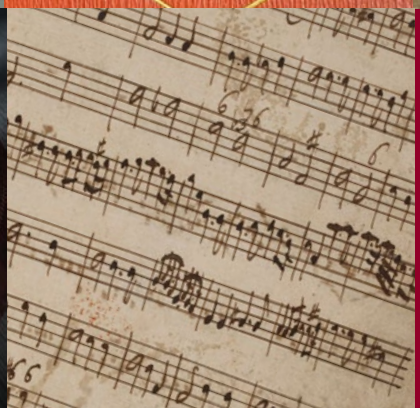
Proceedings of the CLIR Cataloging Hidden Special Collections and Archives Symposium, March 2015

Cheryl Oestreicher, editor







Council on Library and Information Resources

ISBN 978-1-932326-54-3
CLIR Publication No. 169
Published by:

Council on Library and Information Resources
1707 L Street NW, Suite 650
Washington, DC 20036
Web site at http://www.clir.org

Cover design: Jenny Nordstrom. Cover images courtesy of the Provenance Online Project,
Kislak Center for Special Collections, Rare Books, and Manuscripts, University of Pennsylvania.
Map and film reel images via Shutterstock.

Contents

Message from the Editor, Cheryl Oestreicher **1**

Hidden No More, Amy Lucko and Christa Williford **4**

Introductory Remarks, March 13, 2015, Charles Henry **7**

Keynote—Parting the Waters: CLIR's Pathways into the Archive, Jacqueline Goldsby **9**
(Watch video)

Symposium Presentations **18**

Collaborations

All History is Local: Expanding Access to American Jewish Archival Collections
Susan Malbin, Laura Leone, Rachel Miller, Rachel Harrison, Sarah Ponichtera
Christine McEvilly, and Kevin Schlottmann 18

International Collaboration to Reveal Rare Chinese Materials Hidden for Half a Century
Zhijia Shen and Jing Liu 26

The Challenges of Sustaining a Long-Term Collaboration: Reflections on the Philadelphia Hidden Collections Projects
David McKnight and Eric Pumroy 31

Student and Faculty Involvement

Maximizing Partnerships: Faculty Buy-in, Service Learning, and Hidden Collections
Christopher Harter and Elisabeth McMahon 40

Collaboration and Education: Engaging High School Students with EAC-CPF Research
Valerie Addonizio and Christopher Case 48

Engaging Students in Complex Description: Two CLIR Hidden Collections Projects
Lois Fischer Black, Ilhan Citak, Gregory A. Edwards, and Andrew Stahlhut 57

The "Deceased" Preaches His Own Eulogy: Training Students to Provide Access Points on Discovery-Level Records
Felicia Piscitelli, Lisa Furubotten, Anton duPlessis, Alma Beatriz Rivera-Aguilera,
and Ángel Villalba-Roldán 64

Cataloging

Obstacles and Solutions in Establishing Cataloging Standards for Fine Print Collections
Katharine Malcolm and Christen Runge 74

The Churchill Weavers Collection: An American Treasure Uncovered
Jennifer Spence 86

Pennsylvania German Textile Cataloging
Candace Perry 96

Arrangement and Description

The Benefits of Planning: Cataloging the Vertical Files of the Anton Brees Carillon Library
Joy M. Banks and Jaime L. Fogel 104

Discovering the Future: The New York World's Fair Collections of 1939 and 1964 at the Museum of the City of New York and Queens Museum
Annie Tummino 111

Preserving a Montana Senator's Image: The Lee Metcalf Photograph and Film Collections Project
Matthew M. Peek 118

Audiovisual Collections

Accelerating Exposure of Audiovisual Collections: What's Next?
Karen Cariani, Sadie Roosa, Jack Brighton, and Brian Graney 127

Putting Archival Audiovisual Media into Context: An Archival Approach to Processing Mixed-Media Manuscript Collections
Megan McShea 138

Science Collections

The Opportunities of Engagement: Working with Scholars to Improve Description and Access at the Center for the History of Medicine
Emily R. Novak Gustainis 146

The Practice of Privacy
Emily R. Novak Gustainis and Phoebe Evans Letocha 163

Grinnell to GUIDs: Connecting Natural Science Archives and Specimens
Christina Fidler, Barbara Mathé, Rusty Russell, and Russell D. "Tim" White 177

Outreach

Opening up the Urban Archive: Digital Outreach to Urban Studies Scholars
Morgan Gieringer and Jaime Janda 193

Success Beyond Access: CLIR-ing the Way
Harlan Greene, Dale Rosengarten, and Amy Lazarus 202

Concluding Remarks, William Noel **211**
(Watch video)

Epilogue—Hidden Collections for Everyone, Michael Peter Edson **213**

APPENDIX 1: Unconference Sessions **215**

APPENDIX 2: Posters **216**

APPENDIX 3: Learning at Work in the Archives: The Impact of Access to Primary Sources on Teaching and Learning
Kelly Miller and Michelle Morton **217**

Message from the Editor

Cheryl Oestreicher, Head of Special Collections and Archives, and Assistant Professor, Boise State University

The 2015 CLIR Unconference & Symposium was the capstone event to seven years of grant funding through CLIR's Cataloging Hidden Special Collections and Archives program. The two-day event brought together more than 180 past and current grant recipients as well as others interested in the new phase focusing on digitization. The atmosphere was energetic and passionate about the vast array of projects completed or in progress. The more than 25 presentations vividly illustrated the impact of the Hidden Collections grant.

Since 2008, CLIR and The Andrew W. Mellon Foundation have funded 129 projects. To date, institutions have processed at least 2,952 collections, comprising 53,608 linear feet, an additional 4,229 cubic feet, plus 960 boxes of mixed materials. The projects have also created 273,728 item-level records of materials including:

- 50,551 books and manuscripts
- 46,702 audio and audiovisual recordings
- 29,393 items of ephemera
- 27,125 pamphlets
- 15,600 pamphlet plays
- 8,560 maps and map series
- 6,956 artifacts
- 5,537 artworks
- 2,978 architectural drawings

As shown in the Hidden Collections Registry, grant recipients represent academic institutions, government and nonprofit organizations, historical societies, independent libraries, museums, and public libraries. The topics are vast, ranging from activism to book and print history, and from earth sciences to media studies.

We can't know precisely how many researchers have used these newly processed collections and items, but the number is likely to be in the hundreds. At a time when many institutions face chronic processing backlogs and underfunded departments, this program has created jobs, provided resources to catalog collections, and facilitated access to rare materials that will benefit researchers for years to come.

Through my work on a CLIR-funded project and, subsequently, as a grant reviewer, I have greatly broadened my archival knowledge. During my three-year project, Archives from Atlanta: Cradle of the Civil Rights Movement, I was privileged not only to work with amazing civil rights collections, but also to hear how others implement processing procedures and what tools they use. Learning about others' procedures and projects helps me grow as an archivist, and as a manager I share this knowledge and apply many of these practices in my current institution.

In 2010, CLIR hosted its first symposium for grant recipients. At that symposium, archivists presented on topics ranging from architectural records to processing metrics, and from research communities to appraisal. Five years later, the 2015 symposium continued that spirit of connecting and sharing. Attendees learned about others' practices and tools, but also made connections with each other that have led to further collaborations, recognition of expertise, and friendships.

CLIR "aspires to transform the information landscape to support the advancement of knowledge." From the presentations and conversations, it was clear how the program contributes to that goal. In these proceedings, participating archivists and librarians reveal their perspectives, approaches, and ways of sharing information. Readers will see the discovery through the author's eyes.

These proceedings group presentations by theme. *Collaborations* provides examples of multi-institutional projects, including one international collaboration; *Student and Faculty Involvement* reports on practices of engaging students and faculty in processing as well as outreach; *Cataloging* presents projects that experimented with unique ways of classifying unusual formats; *Arrangement and Description* includes a variety of item-level and minimal processing techniques; *Audiovisual Collections* addresses the unique requirements of and advances in providing access to audiovisual items; *Science Collections* focuses on how institutions deal with challenges of science and medical collections, including practical ways to address privacy issues; and *Outreach* describes ways in which projects are engaging current and future patrons.

The Unconference featured in-depth discussions about practical applications, with a choice of workshops in the morning and discussions in the afternoon. Hidden Collections project staff joined with other archivists and information professionals to lead the workshops and discussion sessions. The leaders focused on hands-on tasks, sharing ideas, and reinforcing that we all have similar needs and should work together to learn new skills and implement practices. Notes from the Unconference, like the symposium presentations, are grouped by theme, including data tools, metadata, sustainability, community, promoting collections, donors, and assessment. Each workshop and discussion leader started with their project as a basis and demonstrated how using tools and ideas, such as GIS and K-12 outreach, can be integrated and accomplished. Their purpose was not to just showcase a specific project, but to actively engage participants in discussion and exchange of ideas and experiences.

Archivists routinely exchange information about their profession through conferences, literature, blogs, social media, and other means. The Hidden Collections grant program has taken that to a new level. CLIR has provided a platform to disseminate knowledge that expands beyond the usual means and allows recipients to share their publicity, outreach, finding aids, workflows, templates, and manuals. This platform collates resources that can benefit all archivists. One can see finished projects as well as the procedures and tools to accomplish them.

As editor of these proceedings, I attended nearly all workshops, discussions, and sessions, albeit some for only a few minutes. I wanted to glimpse as much as possible to ensure the most comprehensive representation of the event. I learned too much to list. Some ideas were brand new and others reinforced my current practices. One suggestion that bears repeating is that one does not need to be an expert before embarking on something new: the only way to become an expert is to jump in and try.

What was apparent throughout the two days was not only participants' depth of archival knowledge but also—and this may be obvious—that we are responsible for a great number of tasks to make our unique materials available: appraisal, acquisition, arrangement, description, and access. To carry out these tasks, we need to understand outreach, cataloging, digitization,

standards, and researchers' needs. When I mentioned this to Michael Edson, who has contributed the epilogue to this volume, he asked if I thought the array of tasks was a product of how fast-paced and demanding our society has become. That specific connection had not occurred to me before, but yes, I do.

As technology develops faster than anyone can keep pace with, our researchers' expectations also grow quickly. We frequently get the question "why can't you digitize everything?" While we know that is a monumentous directive, CLIR's new iteration of the Hidden Collections program is one step in that direction. The accomplishments from the past seven years indicate that CLIR will continue to provide opportunities for more collaboration on and access to hidden collections that will significantly impact the profession and advance research.

Hidden No More

Amy Lucko, Director of Program Administration, Council on Library and Information Resources
Christa Williford, Director of Research and Assessment, Council on Library and Information Resources

When we first accepted responsibility for administering *Cataloging Hidden Special Collections and Archives* in 2008, we appreciated the urgency and gravity of the problem that the program was designed to address: far too many materials languished, undiscovered, in cultural heritage institutions nationwide, and scholarship was at risk because potentially vital resources were unavailable for use. Among the reasons materials remained hidden were lack of sufficient funds, expertise, infrastructure, and standardized approaches to exercising intellectual control. Each of these issues presented serious roadblocks for the professionals charged with making special collections and archives discoverable and usable. CLIR's charge, made possible by the generosity of The Andrew W. Mellon Foundation, was to help remove those roadblocks for some, while seeking to understand these impediments and share ideas about potential ways to surmount them.

While the work of the *Cataloging Hidden Collections* recipients is ongoing, this publication represents the fruition of CLIR's efforts in important ways. As a record of the program's second symposium for grant recipients (the first was held in March 2010), the volume captures the richness and breadth of the professional conversations connected with the program over its eight-year history. The presentations offer many practical ideas for improving the efficiency and efficacy of the work required to make rare and unique materials accessible, ideas that are translatable and replicable across a variety of types of collections, and within institutions of varied sizes and missions. By making these ideas widely available, we hope that they might prove useful to professional communities well beyond those directly involved in the 129 projects CLIR funded between 2008 and 2014.

For us, as program staff, these proceedings are also a reminder of the wealth of opportunities we have had to learn from our program's constituents. Through receiving questions and feedback from applicants, exchanges with our dedicated panel of reviewers, reviewing and preparing grant reports, and conducting site visits to grant recipients, we have heard much that we consider essential for others—scholars, funders, and the general public alike—to appreciate with respect to the "hiddenness" of cultural materials. Beyond the collections themselves, many aspects of the work that goes into making them accessible deserve widespread attention, not just to prompt continued investment in description—although admittedly that will be critical—but also to inspire hope within those of us who believe in the open dissemination of knowledge and culture.

The labor involved in cataloging hidden collections most often takes place in private corners, obscured from public view. Few people outside the cultural heritage professions recognize the breadth and variety of the contributions of librarians, archivists, curators, and researchers in their endeavors to understand and contextualize scholarly materials. When distant from one another in time, space, or both, the labor of each

of these groups can be invisible to the others, despite the interdependence of their work. The essays in this volume reveal innovative developments for the many kinds of intellectual labor required to make collections accessible: the team workflows, specialized vocabularies, and sustainable and adaptable storage of catalog records and finding aids; the optimization of approaches to the description of special formats such as audiovisual materials, fine prints, or textiles; or the assessment of risks related to collection content, such as risks to personal privacy. The sections on cataloging, arrangement and description, audiovisual collections, and science collections are particularly helpful in this regard.

We have been fortunate to see with our own eyes the impact of many of these innovations in the course of making site visits to recipient institutions. Each visit has afforded us a unique opportunity to view the work of library, archival, and museum professionals in situ, coming face to face with the many challenges these professionals must overcome in their daily work. Handling unwieldy, disorganized collections; dealing with mysterious documents or objects that surface unexpectedly; working in cramped and remote locations; making case-by-case decisions about what information to capture for a description and what to let go; adapting standards and systems to new purposes; bridging cultural divides between institutions, professions, and scholarly fields of inquiry—all of these are common difficulties for the professionals who have undertaken the work of *Cataloging Hidden Special Collections and Archives*.

Describing collections, like all kinds of work in cultural heritage institutions, is inherently collaborative. But changes in information and communication technologies in recent decades have expanded the scale of collaboration needed to bring research resources to the attention of end users. Working closely with others from different institutions, professional backgrounds, and even countries is now more the norm than the exception. Most of us recognize the positive effects of such large-scale collaboration and increased standardization in language and data structures; in theory, these developments make description cheaper, faster, and better. Yet in practice there are logistical and financial burdens facing those who pursue these seemingly straightforward goals. In the section of this volume focusing on collaboration, the authors of three papers provide a realistic assessment that should prove helpful to professionals planning to enlarge or deepen their working partnerships.

The section on student and faculty involvement offers similar realistic assessments of the limits of the transformative potential of innovative practices, revealing the practical limitations of engaging these user communities in generating high-quality descriptions. At the same time, these essays expose compelling examples of "hidden learning," often described in the words of student workers, taking place within the context of grant projects. A companion piece to this publication by Kelly Miller and Michelle Morton, "Learning at Work in the Archives: The Impact of Access to Primary Sources on Teaching and Learning," provides additional evidence of the value of scholarly engagement and student learning in the context of work with hidden collections. This study grew out of CLIR's three-year effort to document aspects of scholarly engagement through the activities and experiences of early participants in CLIR's program, titled "Observations on Scholarly Engagement with Hidden Special Collections and Archives." Additional information about this study is available on the program website, and in Miller's and Morton's related piece in ArchiveJournal from 2012.

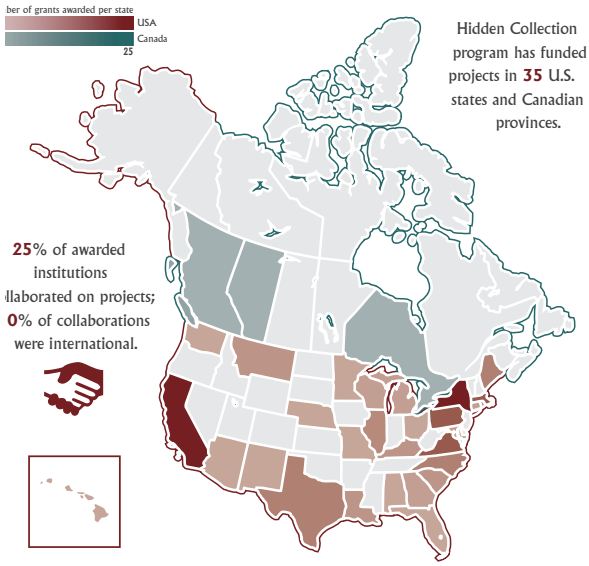
Grants from the *Cataloging Hidden Special Collections and Archives* program have made it possible for scholars, students, and the general public to find and use vast quantities of diverse materials that were not previously discoverable online. This was the program's original purpose, and it has been immensely gratifying to witness its fulfillment over time. But to those of us who have actively participated in the program, the investment of the Mellon Foundation has yielded far more. Grant projects have offered life-changing professional development opportunities for a generation of talented and dedicated staff members new to the cultural heritage professions; they have provided students with unique learning experiences; they have helped forge new collaborative partnerships between institutions; in many institutions, they have prompted the improvement of practices and workflows that are now standard operating procedures. This volume documents these additional benefits of CLIR's program well, and at first-hand. It is our hope that within its pages readers will encounter the full breadth and diversity of activity in today's cultural heritage institutions—all of which is vital to how we understand ourselves both now and in the future—so that they might better appreciate, and perhaps even contribute to, the continuation of this important work.

Introductory Remarks, March 13, 2015

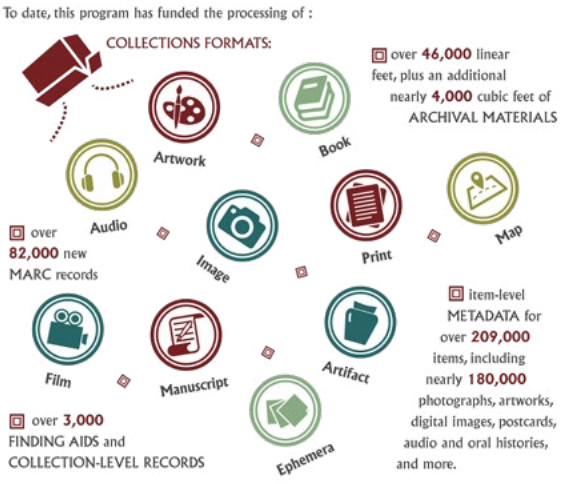
Charles Henry, President, Council on Library and Information Resources

Welcome to the Cataloging Hidden Special Collections and Archives Symposium subtitled "Innovation, Collaboration, and Models." You will be hearing, all day long, some very interesting examples of innovation and collaboration. As we get started, I'd like to thank the program officers and staff at CLIR, who have put together a vibrant agenda, and also thank the Kislak Center and the University of Pennsylvania Libraries for hosting this event.

Very briefly this morning, I'm going to share some statistics and show you what you all have been doing for the last seven years. The Cataloging Hidden Collections program began in 2008 and ran until this year. We've given away 129 grants for a total of $27.5 million. Most of the grants went to academic libraries, but there was a nice dispersal to other kinds of cultural organizations including museums, public libraries, archives, and historical societies.



Through this grant program, 25 percent (33) of the funded projects were collaborative, involving at least two institutions. This is something we really had hoped to see in the beginning, and we encouraged collaboration throughout. One year, we had three wonderful proposals to catalog Civil Rights materials from four different institutions. The review panel got back to these institutions and said, "If you put this together, it'd be a tremendous proposal." They did and it was. That got us moving on trying to encourage collaboration whenever we could. Three of the collaborative projects funded were international collaborations involving institutions in both the United States and Canada. You can see some remarkable statistics here describing funded projects' collective achievements to date: metadata for over 209,000 items, including 180,000 photographs, art works, digital images, and postcards; 3,000 finding aids, and 82,000 MARC records. This is impressive: you all have been very busy.



What have we heard through interviews, surveys, and conversations with you all? What has come out of this, aside from the numbers and the statistics? To me, this is a marvelous catalog, if you will, of reactions and evidence of the importance of this program. First, jobs for new professionals came through these grants. Research guides were created. New courses came into being. There was a really strong pedagogical aspect to this project that we hadn't thought of much in the beginning, but grew over time. Many of you have participated in numerous presentations at symposiums and conferences. And, not least, the projects have led to abiding and binding institutional partnerships and collaboration.

Today, we have 172 people attending: 75 of you are contributing as presenters over the two-day event. Together, participants in this symposium represent 55 recipient institutions and consortia and 62 unique Hidden Collections projects. I am very grateful for all of your work.

Now I'd like to transition to our keynote speaker today, Professor Jacqueline Goldsby. Jackie is chair of the African American studies program at Yale University. My staff and I know Jackie best through her enormous, important contributions to the Hidden Collections review panel, on which she has served for many years. To this day, we continue to talk about the remarkable insight and compassion that she evidenced for the value of our cultural heritage: not just the value of cultural heritage for teaching and research, but the value of our cultural heritage for a vibrant intellectual society.

When the staff and I sat down in 2008 to write the first proposal for the Mellon grant that gave rise to this program, we cited a remarkable effort that was under way at the University of Chicago. This effort entailed a collaboration of librarians, graduate students, and faculty who would go around to institutions in the Chicago area, uncover materials and collections pertinent to African studies, and then catalog them using acceptable standards. The goal of this activity was to make these wonderful materials that were heretofore hidden accessible to the public. That project was headed by Jackie at the time and we said in the grant proposal, "This is a beautiful piece of work and the model on which we want to build." Here we are today. It is a great pleasure to welcome Professor Jacqueline Goldsby.

Keynote
Parting the Waters: CLIR's Pathways into the Archive

Jacqueline Goldsby, Professor of English & African American Studies, Yale University

Standing at the Water's Edge

The title of my talk deserves an explanation—or, rather, a confession. I wish I could credit my inspiration to Taylor Branch's magisterial study of the Civil Rights Movement, because that allusion would suggest I believe that CLIR's Hidden Collections Cataloging Program succeeded because it advances information access as a *civic right*.[1] Now, I do think that case can be made. Moreover, that view informs my remarks today. However, that intellectual genealogy didn't inspire my talk's title.

Instead, my title stems from a guilty pleasure. Every March, I watch Cecil B. DeMille's *The Ten Commandments* (1956). I might rationalize this ritual by claiming to study Cold War allegories of anti-communist statecraft. But that's not true. My mother started me on this habit when I was in elementary school. She was a huge film buff, admired Charlton Heston, and relished well-choreographed visual spectacles. So, I have to admit, the image of Heston—arms thrust wide open against the roiling black sky; intoning his commands in fine, prophetic mode; Yul Brynner, Edward G. Robinson, and the other faux Egyptians shocked and awed as the waters of the Red Sea part: that image immediately sprang to mind as the icon I needed to focalize this talk.

I concede that the religious allusion or the film's kitsch excess (maybe both) might offend secularists and modernists in the audience. But the scene's staging of the proverbial "making a way out of no way" strikes me as a visceral but useful emblem that acknowledges the special effects that CLIR's funding played in the success of the cataloging initiative.

When, in 2005, I first began my own archival project in Chicago, Mapping the Stacks, I read Barbara Jones' ARL white paper, *Hidden Collections, Scholarly Barriers* (2003), and Mark Greene and Dennis Meissner's manifesto "More Product, Less Process" (2005). These field-defining arguments placed my goals for Mapping the Stacks in a sobering context. As Jones and Greene and Meissner made clear, hidden collections imposed a staggering burden upon archives and cultural heritage repositories because uninventoried boxes of materials, overcrowded shelving space, understaffed processing units, patrons clamoring for materials—often unaware that ones more suitable for their projects were just several linear feet away—defied archivists' efforts to devise efficient strategies to recover uncataloged manuscript collections. Not to mention moving image, photography, and born digital collections, which pose their own daunting challenges.

The financial, human, and technological capital required to relieve and reverse those pressures needed to be shrewdly mobilized, on the one hand, and commanding in its own right, on

1 I refer to the first volume in Branch's extraordinary trilogy chronicling the Civil Rights Movement, *Parting the Waters: America in the King Years, 1954-63* (NY: Simon & Schuster, 1989).

the other hand. CLIR's multi-year, multi-million dollar funding capacity bordered on something we might call miraculous, partly because it summoned money from the Mellon Foundation to seed grants that weren't then presently available. More importantly, because CLIR targeted hidden collections explicitly, its funding operated with both fine precision and wide-ranging scale that fostered breakthrough strategies that repositories can now use not only to unhide collections, but to keep those holdings visible and accessible throughout a collection's life cycle.

Judging from their titles and abstracts, the posters and papers presented at yesterday's unconference and today's symposium examine more precisely how these new protocols function *in situ*, on the ground. I begin my remarks with this broader appreciation: the inroads that now lead into hidden collections were made possible because of CLIR's large-scale, longer tail funding capacity.[2] The $27.5 million distributed to the 129 projects across the U.S. and Canada has created not just one but multiple pathways for knowledge production to reach new shores.

Indeed, it's not an exaggeration to proclaim that the collections unhidden through CLIR's Cataloging initiative can—and will—transform the landscapes of research, teaching, and public engagement in humanist studies here in the U.S. and around the world. Truly: the range of materials recovered is logistically staggering and intellectually thrilling. I, for one, can't wait to steer my students and colleagues to the riches that await us in CLIR's Hidden Collections Registry.

Even though this database is still a work-in-progress and might be refined further (Christa Williford advised me), in its current state the registry is an extraordinary resource because of its main differences from resources like Archives USA.[3]

First, the collections federated in CLIR's Registry are, implicitly, the freshest, "rawest" primary sources available, precisely because they've been previously hidden and hardly researched. That's not the case for collections indexed in Archives USA. Second, the collections federated in CLIR's registry are searchable along multiple pathways (seven, compared to Archives USA's two).[4] These search routes hold remarkable potential as user-friendly portals. Which ones and why?

The index of project titles is quite helpful, for instance, because those thematic cues provide a sharper sense of a given collection's likely content. The three subject portals (along with the keyword search box) function in their expected ways but with an important twist: unfamiliar individuals, organizations, practices, or events can be discovered and then linked to broader, more commonly recognized topics.[5]

For instance: I might not know that Margaret Bush Wilson was a Civil Rights Movement activist, but because her papers are indexed under "Civil Rights," CLIR's Registry places Wilson's career in that historical field.[6] With that data point in view, others crop up, ready to plot:

2 A political scientist might liken CLIR's funding model to federal government block grants. Applying for yearly support from The Andrew W. Mellon Foundation, CLIR functioned as a state that redistributed Mellon monies to local entities, to advance Mellon's policy interests in hidden collections vis-à-vis its Scholarly Communications program. For a recent and useful account of this funding approach, see Dilger and Boyd 2014.

3 In its cataloging grant proposal, CLIR stresses the registry's "complementary" functions to Archives USA (p. 4). The point I'm making here is that the differences between the databases make them compatible.

4 Archives USA allows users to search by collection and repository name.

5 To explore these features, see the Hidden Collections Registry's home page at: http://www.clir.org/hiddencollections/registry#c12=all&b_start=0.

6 For Margaret Bush Wilson's papers, see http://www.clir.org/hiddencollections/registry/hc.0953.

because immigration rights collections are also housed under "Civil Rights," a student might begin to think comparatively about, say, the shape of the movement along regional lines (U.S. South, Southwest, Far West).[7] Open to that suggestion, a researcher might think about the politics of segregation differently when, from Wilson's Papers, she's able to identify within the "Civil Rights" rubric that the California State University system's "Activism, Culture and Diversity in Southern California" project includes the McFauling Collection, which holds records concerning Japanese Americans' relocation from the internment camps following World War II. How might African American wartime migration and housing settlement patterns in Los Angeles, Long Beach, and other southern California port cities make sense against the backdrop of Japanese-Americans' removal from their neighborhoods and property?[8] As an English professor, when I teach a novel like Julie Otsuka's *When the Emperor was Divine* (2002)—a haunting novel about a family's forced relocation from Berkeley, California, to Camp Topaz in Utah—addressing questions like these could set very interesting, unexpected cultural frames around that novel. CLIR's Hidden Collection Registry encourages this kind of expansive pedagogical experimentation.

The Registry's potential is so rich that I pledge to use this site in my research seminars and thesis advising at Yale, and I'll encourage my colleagues to do the same. Can you imagine the A.P. high school, college, and graduate-level instructors who could also send their students to this resource to explore possible research projects? Can you envision documentarians who might comb the Registry for leads to un-listened-to sound recordings and never-before-seen prints, photographs, and films to inspire multimedia projects?

I'm sure you can imagine these scenarios. They're probably already happening in your organizations. If so, that's fantastic. The point I want to make here, though, is a request. I'd respectfully encourage all of you—and CLIR—to channel your inner Charlton Heston-as-Moses and lead the publics you serve to the Hidden Collections Registry more assertively. This shouldn't be a hard story to sell. The recovery of so many original, fascinating, inspiring, never-or-hardly-used archival collections—and the labors archivists and librarians expended to organize them—is a mediagenic story that should be spread as widely as possible. The work that you've accomplished deserves publicity on the scale of a Cecil B. DeMille spectacle!

I'm highlighting the Registry for this effort instead of the individual social media sites established for individual projects (a sampling of which can be accessed through the Project Related Resources web page), because the Registry centralizes and focuses attention on the whole, collective lot of unhidden collections that are now known and available for use. A federated publicity strategy aimed at directing a more varied and mass public to converge upon this shared portal will drive a broader range of users to your repositories, I'm certain. Full-to-crowded reading rooms will be a good problem to have on your hands, not simply to build larger constituencies, but also to align new allies and advocates for the ongoing work that collections development requires in the twenty-first century.

7 For the broad listing of civil rights-related collections, see http://www.clir.org/hiddencollections/registry#b_start=0&c12=civil-rights.

8 See http://www.clir.org/hiddencollections/registry/hc.0797.

The Relational Archive[9]

CLIR's Cataloging grant program and the Hidden Collections Registry are remarkable for a third—and final—reason I want to discuss.

As anthropologist and cultural theorist Ann Laura Stoler observes in her brilliant study of archive-making in the nineteenth-century Dutch Indies empire, *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*, "transparency is not what archival collections are known for" (2010, 8), and this problematic has been the focal point of what she calls "the archival turn"—that is, the shift in scholarly emphasis from relying on archives as *resources* for study to critiquing them as *objects* of study (2010, 46-47, 52). I'm sure you're familiar with the body of criticism that defines this turn—I refer here to such works as Michel Foucault's *The Archaeology of Knowledge* (1980), Jacques Derrida's *Archive Fever* (1996), Michel Rolph Trouillot's *Silencing the Past* (1995), Carolyn Steedman's *Dust* (2001), and Diana Taylor's *The Archive and the Repertoire* (2003) among others. These works offer bracing, important challenges to the institutional formations of libraries and archives that have sustained opacities, exclusions, and suppressions of various kinds.[10] However, I want to suggest that, taken together—CLIR's Cataloging initiative, the innovative work you've done arranging and describing those collections, and the Hidden Collections Registry—proffer a different theory of the archive, what I call the relational archive. Its conceptual underpinnings, on the one hand, and social-political capacities, on the other hand, are important to name, because those features may help us confront the sobering but necessary fact that this conference itself portends: how to move forward processing still hidden collections now that CLIR Cataloging grant program has ended?

To address that pressing question, let me first define what I mean by "relational archive." My use of the phrase and concept is inspired by visual studies theorist Nicolas Bourriaud's *Relational Aesthetics* (2002). In that study, Bourriaud explains why contemporary art during the 1990s and 2000s moved away from traditional plastic mediums (such as painting, sculpture, and drawing) to embrace performance-based works in which acts of human contact and sociability were not simply the subject of the work but the form of the artwork itself. A recent, well publicized example would be Marina Abramovic's 2010 performance-installation at MOMA, "The Artist is Present." Abramovic sat in a chair in a large, empty room for three months straight, while visitors lined up for the chance to sit directly across from her, for as long as they chose. The viewers could speak to Abramovic; she would not reply to them.[11] Bourriaud would call artwork like Abramovic's "relational" because it actively solicits the viewer in making the art work function and visible. Indeed, the process of engagement itself—the relationships and transactions that occur between the artist, viewer-participant, and the art object-event constitute the "work" of art as such. In

9 I've addressed this point in an earlier presentation, "Mapping the Stacks: Five Years Later," Opening Colloquium Lecture, "The Past's Digital Presence: Database, Archive, & Knowledge Work in the Humanities," Yale University, February 19, 2010.

10 Other exemplary critiques include Gordon 2008, Lepore 2001, and Hartman 2008. Stoler 2010 offers a cogent overview and complex interrogation of the conceptual rigidities of these approaches, particularly within post-colonial studies (see 46-47, 52).

11 For recent accounts of Abramovic's work, see MOMA's summary of its retrospective of her career, "The Artist is Present" at http://www.moma.org/visit/calendar/exhibitions/965, and the usefully bewildered take by art reviewer Richard Dorment (2014). Bourriaud's book is threaded through with so many lucid, engaging examples it's hard to single any one out. All of them make his point quite astutely.

an era where communications technology and global capital atomize us as much as they link us together, "dialogue," Bourriaud argues, "grants form a productive status…. As part of a 'relationist' theory of art, inter-subjectivity does not only represent the social setting for the reception of art, […it] also becomes the quintessence of artistic practice" (22).

I see Bourriaud's concept of relationality at work when I study the Hidden Collection Registry's subject tag cloud, for instance.[12] That visualization calls to mind the many applications I reviewed in which repositories promised to design social tagging protocols, to involve nonprofessionals in the intellectual labor of identifying and categorizing the materials of a given collection. Likewise, the range and comparative scales of subject areas visualized by the tags' font sizes bespeaks a new sensibility as to what counts as research-worthy knowledge objects, as do the formats of those knowledge objects. In turn, this diversity mirrors the variations we see in the kinds of repositories that were eligible to seek CLIR funding in the first place. Ranging from academic special collections to local historical societies to other cultural heritage nonprofit organizations, "the archive" encompasses the multifarious and amorphous forms it's always been.

Finally, in one of the most remarkable relational patterns, collaborative grants linking institutions together were not uncommon in CLIR's Cataloging initiative. Over the years, joint proposals composed a small but steady percentage of the total number of Cataloging grant applicants.[13] However, funded collaborations represented an eyebrow-raising proportion: 40% in 2008; 26.6% in 2009; 23.5% in 2010; 15.8% in 2011; 22.7% in 2012; 18.1% in 2013; 31.6% in 2014.[14] And yet, prevailing wisdom insists that digitization singularly realizes the relational archive's ideal of shared conceptual engagement.

For instance, in a provocative lecture at the ARL's Fall 2009 Forum on the fate and function of special collections in contemporary times, Mellon Foundation Program Officer Donald Waters argued that digitization will transform "special collections into common ones" because such endeavors can promote collection sharing instead of institutional competition for primary source materials. Moreover, Waters observed, to the extent that digitization destabilizes the idea of institutional ownership, other definitions of value can inform digital-based archival collections.[15] I don't disagree with Waters' propositions at all. Interestingly, though, these very same principles were asserted in CLIR's Cataloging grant proposal—I refer here to its explicit criteria of "interoperability" that required applicants to participate in an iterative exchange process when developing their proposals (2-3) and the "cyberinfrastructure" that was imagined to "facilitate building…virtual organizations

12 Scroll to the bottom of the Registry's home page for the subject tag cloud: http://www.clir.org/hiddencollections/registry#c12=all&b_start=0 .

13 The total collaborative grants submitted (funded or not), run as follows: 2008, 11/118; 2009, 11/169; 2010, 7/145; 2011, 8/72; 2012, 9/100; 2013, 12/75; 2014, 15/92. I'm grateful to Amy Lucko and Christa Williford for preparing and providing these statistics. See e-mail communication from Christa Williford, 25 Feb. 2015.

14 The number of awarded collaborative grants runs as follows: 2008, 6 out of 15; 2009, 4 out of 15; 2010, 4 out of 17; 2011, 3 out of 19; 2012, 5 out of 22; 2013, 4 out of 22; 2014, 6 out of 19. I've exchanged with Amy Lucko and Christa Williford why these patterns may have evolved. First, they rightly pointed out to me that funded collaborations were steady but not numerically dominant. Furthermore, the grant cycle's timing often frustrated efforts to build consortiums from scratch, as it were. Third, in the wake of budget cuts, repositories may have been more reluctant to take on ambitious, collaborative projects. E-mail communication with Amy Lucko, 20 Feb. 2015 and Christa Williford, 25 Feb. 2015.

15 Comparable aspirations inform CLIR's Digitization Initiative. See CLIR's grant proposal, "Digitizing Hidden Collections and Archives: Enabling New Scholarship through Increasing Access to Unique Materials" (2014), http://www.clir.org/hiddencollections/program-proposals/2015-digitizing-hidden-collections-proposal-pdf.

that transcend geographic and institutional boundaries, an interlocking of technical and social elements" (3-4).[16] Put another way, processing and cataloging hidden collections, differently but no less than digitization, articulate a relational view of the archive, too. This paradigm shift in the social logic and role of the archive rates, for me, as one of the key achievements of CLIR's Cataloging initiative.

Theories and Miracles in the Real World

If the statistics I cited earlier don't offer robust enough proof of concept, we can—and therefore should—assess cataloging's transformative power in culturalist terms. Here, Ann Stoler's insight into what counts as an "archival event" is worth recalling:

> …[we should] think about archival events with and against Foucault's compelling injunction to treat them as 'reversals of a relationship of forces, the usurpation of power, the appropriation of a vocabulary turned against those who once used it.' Such an approach undoes the certainty that archives are stable 'things' with ready-made and neatly drawn boundaries. But the search for a dramatic 'reversal,' 'usurpation,' and successful 'appropriation' can hide 'events' that are more muted in their consequences, less bellicose in their seizures, less spectacular in what they reframe. Here I treat archival events more as moments that disrupt (if only provisionally) a field of force, that challenge (if only slightly) what can be said and done, that question (if only quietly) 'epistemic warrant,' that realign the certainties of the probable more than they mark wholescale reversals of direction. (51)

I quote Stoler at length because her insights allow us to appreciate how needles move—which is to ask: what might come next in our common work?

We can't be naïve about the contexts in which archival work now occurs. The vexing developments that shape current educational policy also threaten the momentum CLIR's Cataloging grant program has built. The creative disruptions triggered by constant technological change; the diversification (arguably, fragmentation) of learning publics; and, most devastatingly, the disinvestment in humanist research, teaching, and learning by local, state, and federal government put a sharper point on the imperative I raised at the outset of my talk: how do we keep this work—and the funding that supports it—moving forward?

Self-consciously adapting the conceptual rubric of the relational archive can serve as the impetus and guide we need to extend the gains achieved by CLIR's Cataloging Hidden Collections Program. The "archival common sense" (to use Stoler's apt phrase) it introduces as an institutional practice marks a strategic pathway forward, for two reasons. First, the concept of relationality places clarifying pressure on prevailing critical theories that posit "the archive" to be always already subjugating, dominating, or "imperial." Second, those accounts, whose historical critiques I support but whose trans-historicizing impetus I take issue with, should be distinguished from the relational archive that CLIR's Hidden Collections initiative has set in motion. Because the kind of archive we need to justify and fund, *in and for our present historical moment*, has found viable, compelling, inspiring forms through the practices that your projects have tested and CLIR's Cataloging grant program supported.

16 Quotes are cited from CLIR's original 2007 proposal, "Cataloging Special Collections and Archives: Building a New Research Environment." This document can be found at http://www.clir.org/hiddencollections/program-proposals/2008HiddenCollectionsProposal.pdf.

The strategies you devised to "unhide" hidden collections have transformed "the quintessence of archival practice," to adapt Bourriaud's phrase that I cited earlier. Facilitating outreach to and collaboration with nonprofessionals to process collections; redefining what counts as an archival repository and its knowledge-objects; developing technologies and practices that dislodge traditional claims of originality, rarity, accessibility, and use of knowledge-objects; collating that descriptive information within the Registry's single, shared portal: this model of the archive—which your projects launched—invites library professionals, scholars, students, and the lay public to forge partnerships that can foster and sustain collections management work for the short and long term. Why? Because the relational archive acknowledges we all belong to it. The relational archive functions when we all labor in and on its behalf. The relational archive requires our collective capital—human and financial—to make its form possible. Tackling the problem of hidden collections has been key to this transformation. Precisely because CLIR's cataloging initiative has come to its end, we can see relationality for what it is: an archival mode and form whose time has come, whose time is now.

Acknowledgments

I owe deep thanks to Charles Henry and the staff at CLIR—Amy Lucko, Christa Williford, and Nicole Ferraiolo—for inviting me to offer these reflections on the extraordinary achievements of CLIR's Cataloging Hidden Collections Program.

I also want to acknowledge the work of all the curators, archivists, librarians, and project staff members who serve as the stewards of the cultural heritage materials that made CLIR's program an outstanding success. This is one time I'll hazard to speak categorically, on behalf of researchers, teachers, and students everywhere, as well as the reading-thinking-learning publics at-large to say: thank you for the vital and essential work you do to make knowledge production, teaching, and learning possible.

References

Bourriaud, Nicolas. 2002, trans. [1998, orig. in French]. *Relational Aesthetics,* trans. by Simon Pleasance and Fronza Woods with the participation of Mathieu Copeland. Dijon, France: Les Presses du Réel.

CLIR. Hidden Collections Registry. Available at http://www.clir.org/hiddencollections/registry#c12=all&b_start=0.

CLIR. Project-Related Resources. Available at http://www.clir.org/hiddencollections/resources.

Dilger, Robert Jay and Eugene Boyd. 2014. *Block Grants: Perspectives and Controversies*. Washington, DC: Congressional Research Service. Available at https://www.fas.org/sgp/crs/misc/R40486.pdf.

Dorment, Richard. 2014. Marina Abramovic review: 'I hated every second but I can't deny its power.' *The Telegraph*, June 12. Available at http://www.telegraph.co.uk/culture/art/art-reviews/10895104/Marina-Abramovic-review-I-hated-every-second-but-I-cant-deny-its-power.html.

Gordon, Avery. 2008. *Ghostly Matters: Haunting and the Sociological Imagination.* Minneapolis: University of Minnesota Press.

Greene, Mark, and Dennis Meissner. 2005. More Product, Less Process: Pragmatically Revamping Traditional Processing Approaches to Deal with Late 20th Century Collections." *American Archivist* 68(2): 208-63.

Hartman, Saidiya. 2008. Venus in Two Acts. *Small Axe* (June): 1-14.

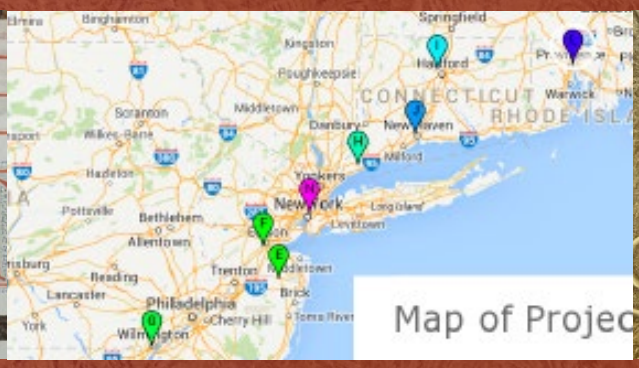
Jones, Barbara M. 2003. *Hidden Collections, Scholarly Barriers: Creating Access to Unprocessed Collection Materials in North American Research Libraries*. A white paper for the Association of Research Libraries Task Force on Special Collections. Available at http://www.arl.org/storage/documents/publications/hidden-colls-white-paper-jun03.pdf.

Lepore, Jill. 2001. Historians Who Love Too Much: Reflections on Microhistory and Biography. *Journal of American History* 88: 129-44.

Stoler, Ann Laura. 2010. *Along the Archival Grain: Epistemic Anxieties and Colonial Common Sense*. Princeton, NJ: Princeton University Press.

Waters, Donald. 2009. "The Changing Role of Special Collections in Scholarly Communications." Presented at "An Age of Discovery: Distinctive Collections in the Digital Age," Fall Forum hosted by the Association of Research Libraries and Coalition for Networked Information, Washington, D.C., October 14, 2009. Available at http://msc.mellon.org/msc-files/specialcollectionsvalue.pdf.

Collaborations



All History is Local: Expanding Access to American Jewish Archival Collections

Susan Malbin, Co-Project Director, American Jewish Historical Society

Laura Leone, Co-Project Director, Center for Jewish History

Rachel Miller, Project Manager, Center for Jewish History

Rachel Harrison, Project Archivist, Center for Jewish History

Sarah Ponichtera, Project Archivist, Center for Jewish History

Christine McEvilly, Data Implementation Specialist, American Jewish Historical Society

Kevin Schlottmann, Data Implementation Specialist, Center for Jewish History

Abstract

The Expanding Access to American Jewish Archival Collections initiative is a multistep collaborative project of the Center for Jewish History (the Center) and the American Jewish Historical Society (AJHS)—one of the Center's five partner institutions. The aim of the initiative is to enhance access to American Jewish archival collections at local Jewish historical societies and museums and within academic special collections on the eastern seaboard. This paper focuses on phase one of the project, funded by the Council on Library and Information Resources, and points to future project phases.

The Expanding Access to American Jewish Archival Collections initiative is a multiphase collaborative project of the Center for Jewish History (the Center) and the American Jewish Historical Society (AJHS)—one of the Center's five partner institutions. The aim of the initiative is to enhance access to American Jewish archival collections at local Jewish historical societies and museums and within academic special collections on the eastern seaboard. The primary goal of the initiative is to create a place for local histories in the larger narratives of American Jewish history by boosting online representation and collection visibility so that local collections become widely accessible.

In its construction and process, the initiative represents an expansion of the collaborative model that currently exists between the Center and its five partner organizations. The partners of the Center—American Jewish Historical Society, American Sephardi Federation, Leo Baeck Institute, Yeshiva University Museum, and YIVO Institute for Jewish Research—are independently managed entities with unique collecting policies related to particular facets of Jewish history. The Center offers a unifying infrastructure for the five organizations in its provision of access-oriented services, which include archival processing, preservation and digitization services, and maintenance of a single OPAC (Online Public Access Catalog) and digital asset management system. The Center's collaborative environment relieves the partners of much of the burden of collection management, freeing resources for other activities central to the partners' respective missions. It also presents a promising model for resource sharing among local Jewish cultural heritage institutions.

Phase one of the initiative entailed completing a series of repository site visits focused on relationship building, information gathering, and the migration of a test batch of collection-level records into AJHS's Portal to American Jewish History. The Center for Jewish History and the American Jewish Historical Society are grateful for the generous support of CLIR (the Council on Library and Information Resources) for the project's first phase.

Selecting Project Participants

The first step was to complete a preliminary survey of local Jewish historical societies and a geographic layout of Northeast and Mid-Atlantic historical societies as a sampling of small, "hidden" repositories of archival material relevant to the American Jewish experience. AJHS provided a previously compiled list of Jewish historical societies in America. The team worked from this list to select 12 institutions as potential project participants: Jewish Historical Society of Greater Bridgeport (CT), Jewish Historical Society of Greater Hartford (CT), Jewish Historical Society of Fairfield County (CT), Jewish Historical Society of Greater New Haven (CT), Jewish Historical Society of Delaware (DE), Jewish Heritage Center of the North Shore (MA), Jewish Historical Society of Western Massachusetts (MA), Jewish Historical Society of North Jersey (NJ), Jewish Historical Society of MetroWest New Jersey (NJ), Jewish Historical Society of Central Jersey (NJ), Trenton Jewish Historical Society (NJ), and Rhode Island Jewish Historical Association (RI).

The project directors and manager then made cold calls to each of the institutions on the list and explained the proposed project. The calls included basic introductions, information about CLIR support, an outline of the project's goals and the objectives of phase one, and a description of the role of a project participant. After the call, the project team e-mailed a project summary and memo of understanding to the prospective participant. In some cases, institutions immediately handed off communication to their archivist or librarian. In one case, the same person fulfilled all roles—and more particular discussions were arranged with the proposed visiting archivists. Some institutions agreed to participate immediately. Others requested more information or time to discuss the project with stakeholders at their institutions.

As the project team continued to communicate with prospective participants, team members learned the value of refining the project pitch. Although prospective participants were receptive and willing to learn more, the team found that conveying too much information about the project up front could be overwhelming and, in some cases, off-putting. It was important to state the project team's goals, then listen to the prospective participants' responses, needs, envisioned roles, and other concerns. Subsequent communication played a crucial role in cementing the relationships and developing concrete plans, such as site visit dates and data transfer procedures. In addition, conveying a clear understanding of the AJHS portal's function as a data aggregator was vital. It was important that participants understood their metadata would be searchable via the portal, but would remain identified with their institution. Preserving identity proved to be a salient issue for a few participants. The need to sensitively address this concern was a valuable lesson for the project team.

Several factors affected prospective participants' decisions about whether to participate: staff changes, prior commitments to conflicting projects, and internal resource limitations. In two cases, institutions either did not have 501(c)3 status, a requisite for participation, or had recently transferred their collections to other repositories. For two institutions, the prospect of contributing

data to AJHS's Portal to American Jewish History presented problems of identity, control, and ownership. In one of those cases, leadership did not understand that an aggregated portal would drive more traffic back to the institution's own site. In another, the society did not interpret the project as a collaborative venture, and believed its role in the project would be construed as that of a follower rather than a leader. The leadership at these two institutions decided that they had no need of the services the project could offer.

However, most of prospective participants responded positively, and a few went so far as to recommend additional repositories to contact based on regional connections and collaborations. The project directors and manager expanded the geographic area down the east coast to include additional regional Jewish historical societies or institutions with relevant collections. The team spoke by phone with 20 repositories. Ultimately, the following 13 organizations agreed to participate:

- Beth Ahabah Museum and Archives (Richmond, VA)
- Charlotte Jewish Archives at the Jewish Historical Society of Greater Charlotte (Charlotte, NC)
- Jewish Buffalo Archives Project, University at Buffalo (SUNY) and Bureau of Jewish Education of Greater Buffalo (Buffalo, NY)
- Jewish Heritage Foundation of North Carolina and Duke University's Rubenstein Library (Durham, NC)
- Jewish Heritage Museum of Monmouth County (Freehold, NJ)
- Jewish Historical Society of Central Jersey (New Brunswick, NJ)
- Jewish Historical Society of Delaware (Wilmington, DE)
- Jewish Historical Society of Fairfield County (Stamford, CT)
- Jewish Historical Society of Greater Hartford (West Hartford, CT)
- Jewish Historical Society of Greater New Haven (New Haven, CT)
- Jewish Historical Society of Western Massachusetts (South Deerfield, MA)
- Rhode Island Jewish Historical Association (Providence, RI)
- University of North Carolina Chapel Hill, Southern Historical Collection (Chapel Hill, NC)

The Jewish Buffalo Archives Project served solely as a data collaborator. (The organization was in the middle of a move and preferred to forgo the site visit component of phase one.) One factor of success for the project was the fact that the final list of participants represented an even more diverse spectrum of institutions and collections than originally envisioned. The participant list included small, volunteer-run institutions to large collections affiliated with universities.

After confirming the list of project participants, the Center assigned two project archivists to conduct site visits to 12 institutions (one archivist per visit). During the site visits, which averaged about two to three hours, archivists recorded information about the history, collections, projects, and priorities of each institution. Following the visits, archivists compiled a two- to three-page site visit report on each repository. Summaries of the individual reports can be found at the project web page.

The two archivists who conducted on-site surveys of the 12 repositories found a number of trends among project participants. Although the

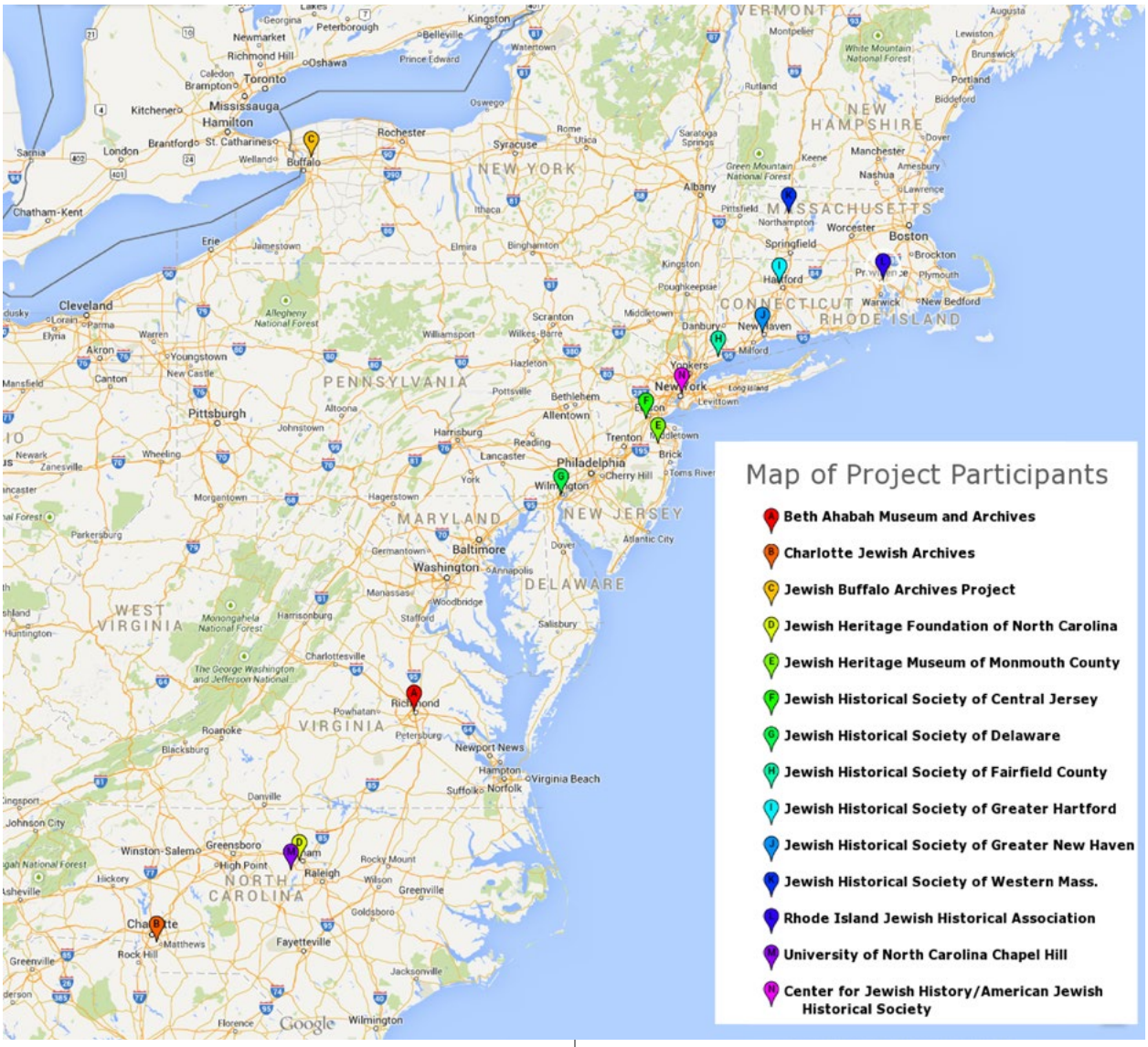


Fig. 1: Map of project participants

institutions visited during phase one are highly diverse—from lone arrangers to university archives—all reported a desire to raise their public profile and increase user traffic and public awareness of their holdings, as well as to engage in more outreach and programming activities. Users of these institutions are often interested in doing genealogical research. They are often faculty and students from local universities, local authors and historians, or staff from Jewish organizations. A number of the repositories are Jewish historical societies housed in spaces that they share with other organizations such as local Jewish Federations, Jewish community centers, larger historical societies, university library special collections, or synagogues and museums. The repositories often benefit from this physical proximity because of shared resources, visitors, and occasionally access to staff and volunteers. Many of the repositories have small staffs who work part time and have numerous responsibilities beyond working with collections. Several

rely on volunteers and interns who may not have access to formal archival training. At the smaller repositories, both staff and volunteers would benefit from training in current archival methods, standards, and practices.

In terms of holdings, the smallest repository comprises eight linear feet, and the largest around 1,000 linear feet. Many range from 300 to 500 linear feet. The dates of holdings range from the 1700s to the present, with most from the nineteenth and twentieth centuries, often corresponding to the dates of mass Jewish immigration to this country. The collections are in a variety of languages, mostly English, Hebrew, Yiddish, with some in German, Russian, and Polish. Staff members' ability to work with these languages ranges widely.

Many of the repositories' collecting policies focus on local Jewish history. Along with genealogical materials, the surveyed repositories collect manuscripts, personal papers, Jewish institutional records, farm records, account books and financial records, ledgers, photographs, books, cemetery records, rabbis' sermons, synagogue records, graphic materials, historic ketubahs, artifacts, correspondence, brochures and programs, albums and scrapbooks, oral and video histories, microfilm, audio and video cassettes, DVDs, electronic records, meeting minutes, memoirs and diaries, slides, historic newspapers, yearbooks, city directories and phone books, family trees, obituaries, maps, and drawings. Oral and video histories are particularly well-represented, although few of them have associated transcripts.

The institutions have a range of intellectual control over their collections. Several have inventories that list the contents of folders and boxes but lack the other elements of a formal finding aid. Others have accession records but no contents lists. Photographs account for many research requests and tend to be better described. Other than those working in the largest of participating institutions, the staff of the repositories mentioned an interest in receiving help with processing their archival backlogs and creating online finding aids. Several repositories have organizational systems set up by previous staff, many of whom were not archivists, and are using these systems even though they may not prefer or even fully understand them. Among the processed materials, some are arranged alphabetically by subject and use local terms, some are in collections, and some are just in alphabetical folders. At most of the repositories, some of the collection information resides with the staff and can be hard to access from outside the institution. Many finding aids and inventories exist only on paper or on a local computer as a Word or Excel document or PDF, and should be encoded in EAD (Encoded Archival Description) to provide online access and searchability. Several small- to mid-size institutions provide summaries or inventory lists online, though the data are not structured or easily searchable. Large academic institutions have their data online and in a searchable format, but the records do not consistently have access points indicating that the material is related to the American Jewish experience.

Several repositories requested guidance on developing and improving their collecting and accession policies, increasing their collecting activities, and determining what to keep of their own institutional records. They also expressed interest in how to do outreach for accessioning Jewish collections, how to better capitalize on what they already have, and in getting assistance with grant writing. Some repositories expressed a wish for guidance on how to best conduct oral histories and how to make them

more available to users through online hosting. Repositories had myriad questions about general archival practice, technology trends, digitization and born-digital materials, and audiovisual materials. A few repositories mentioned needing access to a conservator for consultations or for actual conservation work. The Jewish historical societies are also interested in increasing their membership, particularly among younger members, and hope that greater access to their archival collections will help them do this.

Overall, the site visits completed in phase one established strong relationships between the Center, the American Jewish Historical Society, and project participants. Visits allowed for in-depth dialogue and created space for informal exchange of knowledge and questions. Face-to-face sessions led to new and unexpected lines of inquiry and discoveries that would have not happened as easily by phone or e-mail. The site visits were an integral component of the project, allowing all project participants to consider how they might collaborate on collection management needs and access issues.

Ingesting Collection-Level Records into AJHS's Portal

After establishing collaborative relationships, the initiative's next step toward increasing access to American Jewish collections was the ingestion of collection-level records into AJHS's Portal to American Jewish History. AJHS's portal is a metadata aggregator that enables researchers to perform complex searches across American Jewish archival collections currently residing at more than 10 geographically dispersed repositories. It is currently set up on a Drupal Collective Access platform. Drupal is an open source website content management system, and CollectiveAccess is an open source collection information management

Fig. 2: Jewish Heritage Museum of Monmouth County (photo credit: Sarah Ponichtera)

system (or metadata database). Using a Drupal plug-in, CollectiveAccess feeds information into Drupal, allowing users to search the database seamlessly throughout the site.

In 2013 Whirl-i-gig, the independent contracting firm that created CollectiveAccess, set up AJHS's system on an Amazon cloud server. Since then, AJHS has maintained the site and continues to import data from new repositories. In the past few months, AJHS moved the system to an in-house Center for Jewish History server to improve control over the site, reduce costs, and leverage the Center's resources. The use of Drupal and CollectiveAccess gives AJHS the opportunity for extensive future growth. Drupal can be modified and extended with many freely available plug-ins; it can be extensively modified to present results with faceted browsing and thumbnails of images. CollectiveAccess is a robust platform

that lets AJHS create new fields as needed, handle authority records, and manage subject and genre terms for website browsing facets.

As part of phase one of the Expanding Access to American Jewish Archival Collections project, data implementation specialists ingested 104 collection-level records into the portal. The records originated from three project participants: the Jewish Historical Society of Fairfield County (Stamford, CT); the Jewish Buffalo Archives Project, University at Buffalo (SUNY) and Bureau of Jewish Education of Greater Buffalo (Buffalo, NY); and the Jewish Historical Society of Greater Hartford (West Hartford, CT). Different approaches were used to prepare each set of data for ingest. For their metadata, all three repositories signed Creative Commons licensing agreements with the Center and AJHS, with an eye toward easing future data sharing, especially with the Digital Public Library of America.

Acquiring the Jewish Buffalo Archives Project data was straightforward. Excitingly, the process can be used to obtain data from any HTML finding aid published from a digital asset management (DAM) system in a scalable and automated way. Buffalo provided a list of formulaic links to HTML finding aids and the HTML produced by Buffalo's DAM, XTF, is regular and well-structured. Therefore, staff could easily scrape all the HTML and extract the needed data via xQuery. Each collection has a MARC record in the University of Buffalo system, and most have HTML finding aids that are linked from the MARC record. Buffalo provided AJHS and the Center a list of collections with records numbers. From Cygwin Unix command-line shell, the Center archivist used a single "curl" command to scrape all the HTML finding aids into one file. The list of IDs was taken from the spreadsheet provided by Buffalo. The "curl" command allowed staff to use one link, with variable IDs, to scrape the URLs for each finding aid. The command wrote the HTML it found at each URL into one large file.

The resulting document was opened in oXygen XML Editor, with doctype declarations removed and an XML root element and XML declaration added. The resulting valid XML file was imported into the XML database software BaseX. After some study of the HTML, an xQuery script was written to extract the metadata fields required by the portal's CollectiveAccess software and those used for public display. This process produced CSV data that could be imported using Excel. For the exact script used, see the project webpage.

The Jewish Historical Society of Greater Hartford provided EAD files by e-mail. The EAD files were exported from Archivists' Toolkit, and were thus uniform in terms of tag and data placement.

Fig. 3: Record in the Portal to American Jewish History

The Center's data implementation specialist analyzed the EAD files, imported them into the XML database software BaseX, and ran an xQuery script to extract metadata fields for the portal's CollectiveAccess software. For the exact script used, see the project webpage.

The Jewish Historical Society of Fairfield County provided AJHS and the Center with very rich container list information in an Excel format, exported from the collection management software PastPerfect. The records were not organized at the collection level, but included title information that indicated the provenance of a particular container. For instance, a sample container-level row contained this information:

Temple Sinai of Stamford, CT Bulletins from 2000-2001.	Temple Sinai Bulletins	Bulletins	478

The Center and AJHS's data implementation specialists suggested manually creating collection-level records, and the society agreed. The final collection-level record, which was imported into the portal, looked like this (record slightly abbreviated):

Aaterials on Temple Sinai of Stamford, CT. Includes monthly ulletins from 1959-1960, 1978-2004 (with gaps and missing issues). Also includes booklets, annual meeting reports, various historical ocuments. Subjects include memorial services, High holiday service, membership directories, adult education, congregation history, bar nd bat mitzvahs, cantors, Rabbi Marc Disick, Rabbi Liz Rolle, Rabbi avid Wolfman, Rabbi Sharon Sobel, Rabbi Samuel Silver, Rabbi tephen Pearce, Rabbi David Pearlman, Rabbi Jay Tel Rav, and Rabbi ohn Franken. ID numbers 473-524, 316b	Temple Sinai archival and bulletin collection	Temple Sinai	1954-2012	JHSFC_473-524	473-524, 316

This manual process was time-consuming, but the resulting records were contextually appropriate for the portal.

Encouraged by phase one of the project, the Center and AJHS have established ambitious goals for continuing with the Expanding Access to American Jewish Archival Collections project. Future goals for the project include:

- Becoming a Digital Public Library of America (DPLA) Judaica content hub
- Migrating additional collection-level records into the AJHS portal
- Encoding legacy finding aids and inventories
- Processing selected participant collections or training staff to process collections
- Hosting already digitized materials online in the Center's DAM, providing both access and long-term preservation
- Digitizing selected material and hosting that material in the Center's DAM
- Augmenting the development of the AJHS Portal by:
    - Monitoring and checking in with participating repositories to update feeds as necessary
    - Evaluating usage using Google Analytics
    - Building interpretative exhibits and interactive social media tools

In the future, the initiative will expand to include more and varied types of institutions including those in the South and Southwest United States. A social media component in the portal would also enable self-selecting historical societies, Jewish community centers, and individuals to upload material and participate. In addition to allowing participants to share highlights of their collections and to make these items discoverable at little cost, the portal will allow participants to situate their currently hidden local histories within larger narratives of the American Jewish experience. Information and updates about the next phases of the Expanding Access to American Jewish Archival Collections can be found on the project website.

International Collaboration to Reveal Rare Chinese Materials Hidden for Half a Century

Zhijia Shen, Director, East Asia Library, University of Washington

Jing Liu, Chinese Language Librarian, Asian Library, University of British Columbia

Abstract

This report discusses Discovering Modern China, an international collaborative project supported by CLIR (Council on Library and Information Resources) and conducted by the University of Washington Libraries and the University of British Columbia Libraries (Canada) to reveal the hidden treasures of Chinese language materials. These include rare books, special collections, archival materials, and so on, printed from the fourteenth century to the early twentieth century. We address collaborative project planning and management, resource sharing, and staff exchanges and training among the University of Washington, the University of British Columbia, and academic libraries in China. We also demonstrate how international collaboration can optimize libraries' capacity and resources to accomplish difficult tasks, and we illustrate the challenges of working with the policies, laws, and regulations of different countries.

Among the best libraries specializing in East Asian studies in North America, the University of Washington (UW) Libraries and University of British Columbia (UBC) Library are major regional, national, and international bases of information and research about East Asia. Through more than half a century of collection development, the two libraries have built outstanding collections including unique Chinese rare books, archives, and other special materials. Many of these special materials, however, have yet to be cataloged because of lack of funding and staff expertise.

In 2013, UW and UBC received jointly a grant from the CLIR Cataloging Hidden Special Collections and Archives program to conduct an 18-month project called Discovering Modern China: University of Washington and University of British Columbia Collections. The goal was to catalog these hidden treasures and make them discoverable to users and scholars worldwide. This project was also one of the first two CLIR-funded international collaborative efforts involving a Canadian university library. The grant enables UW and UBC to collaborate on project planning and resource sharing, staff exchanges and training, and collections and user services sharing. It also enables us to receive expert consultations and personnel from academic libraries in Asia. As a result, UW and UBC libraries will create 2,000 original bibliographic records in OCLC WorldCat and in the Union Catalog of Rare Books of the China Academic Library and Information System (CALIS), an academic library consortium in China. Since its launching, the project has already created several hundred records and increased the awareness of and access to these treasures, which benefits not only the scholars of the two institutions, but also scholars worldwide. In an international joint effort such as this, the process is just as important as the result. Collaboration has been the core of this project from the start.

Identifying Hidden Treasures

The UW and UBC library collections contain rare and unique materials originally owned by private individuals, families, scholars, and collectors. Many special collections came to our libraries through various routes during troubled times in early twentieth-century China. Throughout the past half century, however, access to the proposed collections has been extremely limited, and materials have been kept in storage, unorganized. On-site browsing has been the only method of access.

The two institutions' Chinese collections identified for this project complement each other in their uniqueness. For example, the UW collections include rare historical archives and unique special collections such as wooden fish books, while the UBC collections are known for several noteworthy private collections of traditional thread-stitch-bound rare books. Both collections are rich in rare or unique books and are recognized for their significant research value. To plan for this project, we connected with world-renowned scholars who have browsed and benefited from the hidden treasures in the two libraries. They are familiar with both institutions' hidden collections, which have benefited their research. These scholars—from the United States, Canada, China, and Europe—not only provided strong recommendations but also helped with selection and planning. Some of them are currently leading international research teams and are excited to see that we are taking an international collaborative approach to unveiling hidden treasures to the world.

Our project comprises three phases. Phase I (6/1/2014 to 11/30/2014) focused on hiring and training student assistants to search collections, hiring project catalogers and a rare book consultant, and setting up project equipment and space. Phase II (12/1/2014 to 3/31/2016) has focused on cataloging and assessing collections. Phase III (4/1/2016 to 5/31/2016) will focus on project evaluation.

At the time of this presentation, we have completed phase one. In that phase we searched numerous uncataloged materials to identify unique and highly valuable titles. We hired and engaged graduate students of Asian studies and history to conduct bibliographical searching and sorting, to create an inventory list of unique titles, and to mark their preservation needs. For both institutions, especially UW, phase one of the project has yielded a significant number of potential rare books for the CLIR consultant librarian to evaluate.

Collaboration in Applying Cataloging Guidelines and Rules

The UW and UBC project teams closely follow the Research Library Group's (RLG) *Cataloging Guidelines for Creating Chinese Rare Book Records in Machine-Readable Format* as well as the Program for Cooperative Cataloging RDA (Resource Description and Access) hybrid bibliographic records guidelines. Created to reconcile Anglo-American cataloging rules with cataloging practices of centuries-old traditional Chinese rare books, these guidelines presented an international collaborative model before us. The East Asian library community in North America has been working to apply these rules to the MARC 21 format. Because we do not plan to confuse the guidelines with the RDA's new cataloging rules, we were mindful about maximizing sustainability and interoperability by adding hybrid RDA fields and relationships in the online bibliographic records created for the project.

Charlene Chou, head of technical services at the UW East Asia Library, has prepared questions and

comments derived from the rare book cataloging practices of the project, and submitted them to the Committee of Technical Processing of the Council on East Asian Libraries. Charlene is leading the committee and its rare book working group.

The Importance of Communication

Both the UW and UBC project teams selected unique and rare materials from the hidden collections to create original bibliographic records. We established a clear scope of work and workflow through close communication. The two teams take day trips to visit each other and to inspect collection conditions and work space setup. We conduct regular Skype meetings and use instant messaging and image sharing. Both teams launched project websites[1] and posted project status reports, exciting discoveries, and scholarly engagement activities. We also use Google Drive to share working documents. Although we follow our respective library's cataloging policies, we also conduct vivid discussions about best practices. The collaborative experience has enhanced the quality of our work. The well-documented working papers we have created will be an important reference for future efforts and collaboration.

Sharing Staff Expertise between Institutions

The UW East Asia Library has a very strong technical services unit led by Charlene Chou, an internationally known cataloging librarian. The head librarian, Zhijia Shen, specializes in Chinese history and international scholarly engagement. To ensure high-quality work, UW provides training for all CLIR project staff via telecommunication and in-person visits. Training and technical workflow are fully documented and available on the UW technical services staff website. The project also has sought international expertise including top-notch scholars from China as core project staff. Professor Boyue Yao, a renowned Chinese rare-book specialist and librarian at the Beijing University Library in China, joined us at UW's East Asia Library in January 2015 for a 13-month appointment as the CLIR project librarian for Chinese rare books and cataloging. He also joined UBC's Asian Library as a Chinese rare-book consultant for two months in April and May 2015. Mr. Ya Min Wu, former assistant director of Liaoning Provincial Library and a seasoned Chinese rare-book librarian, was appointed CLIR rare-book cataloger at UBC for 12 months as of December 2014. Both librarians are well-known Chinese rare-book experts. They bring in-depth knowledge and expertise to our CLIR project.

To maximize the impact of the project on cataloging, the principal catalogers in central technical services at UW and UBC Libraries have discussed training in NACO (the Library of Congress's Name Authority Cooperative Program), to enable the project to contribute name authority records.

In fact, this project has prompted partnerships that reach beyond our Chinese hidden collections. More broadly, it has enhanced communication and collaboration between the two libraries in areas such as preservation and conservation, digitization, staff exchange, training, and future grant proposals.

Although copy cataloging is not part of the CLIR project that we proposed, both UW and UBC libraries have encountered tough decisions about how to deal with it. For this project, both institutions did searching, selecting, and inventory listing of all the unique and rare materials that have no holdings showing in OCLC WorldCat. At UBC, copy cataloging is done by union-member

1 Available at http://depts.washington.edu/ealclir/wordpress and http://guides.library.ubc.ca/clir.

catalogers, and cannot be handled by student assistants or non-union members. Because the UBC library has no Chinese copy catalogers on staff, those books must be left behind locked doors. This is a challenge. The library needs to come up with a new strategy to tackle the copy cataloging resulting from the CLIR project.

The UW experience in handling copy cataloging provides a good model. In phase one we developed a parallel plan for the CLIR hidden collections project. While searching for hidden treasures, copy cataloging is performed by highly trained student assistants and paraprofessionals. Copy-cataloged materials go right into circulation and become available for interlibrary loan requests. Thanks to the strong technical service unit at UW's East Asia Library, searching, selecting, and copy cataloging are smoothly integrated into the workflow, and satellite projects have been created in support of the core CLIR project. This integrated working environment has optimized the project's impact, which reaches well beyond the original proposal to catalog 1,000 Chinese rare books.

Resource Sharing Worldwide

Thanks to Professor Yao, the CLIR project librarian from the Beijing University Library, the UW-UBC CLIR project has joined the Union Catalog of Rare Books of CALIS (the China Academic Library and Information System). The UW and UBC libraries have become its first members outside China. Twenty-five of China's top academic libraries belong to CALIS. Many records in the system are not available in OCLC WorldCat. Valuable bibliographical information can be easily derived from a similar edition in the CALIS Union Catalog for our original cataloging. Its customized interface for rare-book-specific information is user friendly. Records can be converted between CN MARC and MARC 21. Joining CALIS's Union Catalog not only has sped up the cataloging for our CLIR project, it has also promoted our newly unveiled hidden treasures to scholars in China. At the same time, it has allowed UW and UBC professors and students access to the Chinese rare-book collections held by 25 of the best academic libraries in Mainland China, Hong Kong, and Macau through a half-million online rare-book records and more than 270,000 images.

In phase two of the project, we have encountered many issues of forgery and authentication. To achieve our goal of creating accurate cataloging records and preparing for sharing the hidden treasures widely, we plan to scan one to three images from each rare book to provide crucial information for edition authentication. The images will be linked to the online cataloging records, which will enhance the research potential of these materials and benefit scholars and researchers worldwide. Meanwhile, we hope to receive user-contributed information and metadata about these rare books, which will not only help with our cataloging tasks, but will also make connections to other similar holdings.

The UW-UBC project has optimized our libraries' capacity and resources to accomplish difficult tasks through international collaboration. As of February 13, 2015, the second month of the second phase, we already had cataloged hundreds of unique and extremely rare titles, contributed authority records, and triggered several conservation plans. We discovered rare sets, some volumes of which were held in libraries in China and others in our libraries. For example, a rare book set held at UBC was missing a few volumes, which were considered permanently lost. They have now been identified at Beijing University Library. The genealogy of our rare books and

their collecting history over the centuries have attracted the research interest of professors and graduate students. It is now possible that scattered volumes of original collections may be identified and linked together virtually. Once the project is completed, we hope that some of the efforts of the original collectors will be realized, albeit across a long time span and wide geographic distance. This project has allowed our hidden collections at UBC and UW to enter a new world, gain new life, and enjoy continuity of heritage.

Challenges of International Collaboration

The benefits and impacts of our international collaboration are profound and long-lasting. However, such collaborations can also be challenging. We have run into obstacles working with different countries' policies, laws, and regulations, and with communicating within different administrative and organizational structures and reporting systems. For example, we selected for UBC a seasoned cataloger from Taiwan National University Library, but failed to obtain her permit to work in Canada. We anxiously waited for Professor Yao's application for a Canadian visa, which ended up being complex, and delayed his arrival in Seattle. When the UBC team brought Professor Yao to UBC for a one-day visit after visiting UW, he ran into tough questions from border immigration officers such as "Why can't you work remotely for UBC by video conferencing?" and "Do rare book consultants really need to touch the actual items?" These are just some examples.

The CLIR project has enabled us to discover and make available to users the treasure collections in our libraries. Even more importantly, it has provided us an invaluable opportunity to learn about international collaboration. As libraries and information services become increasingly globalized, there will be more need for international collaborations. We hope the lessons learned and experiences gained from our project will benefit future CLIR applicants and projects internationally.

The Challenges of Sustaining a Long-Term Collaboration: Reflections on the Philadelphia Hidden Collections Projects

David McKnight, Director, The Rare Book and Manuscript Library, University of Pennsylvania

Eric Pumroy, Associate Chief Information Officer and Seymour Adelman Director of Special Collections, Bryn Mawr College

Abstract

During the past five years, the Philadelphia Area Consortium of Special Collections Libraries (PACSCL) received two CLIR Hidden Collections grants that tested the capacity of this voluntary collaboration of academic and cultural institutions to develop and sustain a unified methodology and provide a single point of public access to their manuscript and archival collections. While the projects have been successful in achieving their goals, there continue to be unresolved issues around the development and maintenance of the project database. This two-part paper looks at these projects as a means of examining the challenges of building a sustainable multi-institution technology-based program, and does so by looking at the issues from the perspective of both the host institution and the participants.

The Philadelphia Area Consortium of Special Collections Libraries (PACSCL) received two Hidden Collections grants to process archival and manuscript collections in the region: one in 2008 for $500,000, and a second, smaller grant in 2011 for $249,000. These are the most recent of a long series of collaborative grant-funded projects that this consortium of 36 academic institutions, historical organizations, and museums has received since its founding in the 1980s, including grants from the Pew Charitable Trusts in the early 1990s to address uncataloged book and manuscript collections; a grant from the William Penn Foundation in the late 1990s to set up online catalogs for the smaller institutions in the consortium; and a grant from The Andrew W. Mellon Foundation in 2005 to survey unprocessed manuscript collections and identify priorities for further attention. This last project, which ended in 2008, served to identify most of the collections that were processed in the two CLIR grants.

The two Hidden Collections grants were highly successful in meeting their goals of reducing the backlogs of important unprocessed collections in the region, and in fact, they processed more than had originally been called for in our proposals. In the first grant project, 133 collections totaling 4,000 linear feet were processed across 21 institutions. In the second, smaller project, 45 collections totaling 1,685 linear feet were processed across 16 institutions, some of which were not part of the first project. In all, 178 collections totaling 5,865 linear feet were processed during the two CLIR Hidden Collections projects and involved collections from 24 institutions.

As important as the processing was, it was only a piece of a larger and more ambitious plan to create a finding aids database that would provide access to all historical collections in the region. The University of Pennsylvania Library was in the process of setting up such a database for its own finding aids at the time we were developing

the first proposal and generously offered to expand it to include finding aids from PACSCL institutions. The proposal for the initial grant explicitly stated that the Hidden Collections project finding aids would only be the beginning point for the database, which we saw eventually becoming the critical site for research in the Philadelphia area. Seven years and two projects later, the database has become much richer, but the number of institutions that routinely add new finding aids to it is not as large as we had hoped. The purpose of this paper is to look at this project as a test case of what it means to collaborate on a project that involves building and maintaining a technology infrastructure. The first part will consider how the database has been used and the factors that have determined institutions' level of involvement, and the second will examine the implications of a collaborative technology project for the host institution.

Part I: The Participants' Perspective

Eric Pumroy, Bryn Mawr College

In many respects, the finding aids database has been an impressive success. As of spring 2015 there are about 2,800 finding aids in it, compared with the 178 finding aids created during the two Hidden Collections projects. Our great surprise is that the largest contributor is a program that did not exist when we got started: The Hidden Collections Initiative for Pennsylvania Small Archival Repositories, funded by grants from The Andrew W. Mellon Foundation and developed and managed through the Historical Society of Pennsylvania. As of March 2015, there were more than 900 finding aids from about 150 repositories represented in the finding aids database, and nearly all of these were collections that were invisible to the larger research community until the project and the database provided a way to catalog and publicize them. The other major contributors are the University of Pennsylvania with more than 700 finding aids, the Hagley Museum and Library with 671 finding aids loaded in the fall of 2014, and the rest of the PACSCL institutions with about 500. It is this last group that I want to look at more closely. Aside from the University of Pennsylvania and Hagley Museum and Library, the PACSCL members that contributed the greatest number of finding aids were Drexel University, Haverford College, Bryn Mawr College, the Free Library of Philadelphia, and the College of Physicians of Philadelphia. Contrary to our expectations, most others have not continued submitting finding aids. In a survey conducted at the end of the first project in 2012, 17 institutions, or about three-quarters of the participants, said that they were very pleased with the database and would continue to submit finding aids. In late 2014 and early 2015 we did a follow-up survey of all PACSCL members, and of the 17 respondents who had participated in the CLIR grants, 8 said that they planned to continue reporting (for the most part, they are the ones that have been reporting), and 8 said that they were uncertain.

The reasons for the uncertainty and lack of continued submissions vary by institution, but most of them come down to issues with technology. The way the system works is that an institution creates a finding aid using Archivists' Toolkit, and then uploads the finding aid to its own web server where Penn's automated harvester finds it and loads it into the finding aids database. Several of the smaller institutions reported a lack of IT support, which makes running Archivists' Toolkit and uploading finding aids difficult. Others had to deal with institutional IT policies that blocked access to the harvester. There are also problems with the database itself that held institutions back from

loading finding aids, particularly in the way it requires collection numbers to be formatted. And finally, a few institutions' finding aids were dropped from the database when their websites moved and the harvester could no longer locate their finding aids. None of these problems are insurmountable, but solving them requires time, attention, and in some cases, money.

The finding aids database was always intended to be for all collections in the region, not just for those institutions that participated in the Hidden Collections projects. The recent survey was sent to all PACSCL members as a way of capturing how, or if, the non-Hidden Collections institutions thought about the database. These institutions did not sit out the two PACSCL projects because of opposition or lack of interest, but rather because the projects had been based on the results of the earlier consortial survey that ranked the research values of unprocessed collections in the region. As a consequence, the institutions that processed their most important collections on their own did not have collections that qualified. A number of major institutions, such as the American Philosophical Society and Swarthmore College's Friends Historical Collections fell into this category. Because these institutions were not involved in the Hidden Collections grants, they did not give much consideration to the finding aids database, and in fact, several had not realized that contributing their findings aids was an option. Several reported that they would be interested, but would need advice on how to go about it. While these are only a handful of institutions, the responses indicate a continuing interest in using the database as a way of improving access to collections.

One of these institutions, the Hagley Museum and Library, on its own initiative contacted Penn about loading its finding aids on to the PACSCL site, and its finding aids now constitute nearly a quarter of the finding aids. The Hagley's decision came about as an offshoot of a larger project to create its own EAD finding aids database using the open source software XTF, developed by the California Digital Library. Once the finding aids were in their own database, it was a relatively straightforward matter to also make them available to the PACSCL site. At this point the benefits of having their finding aids in the PACSCL site aren't clear because they have only been there for a few months, but as Library Director Erik Rau reported, they thought that it was important for their finding aids to be findable as part of a larger regional collection and exposed to a wider population of users.

While there is interest among most PACSCL institutions in having the finding aids database succeed, the database is nonetheless a low priority for them. Most institutions have their own ability to mount finding aids so that they are findable through web searches, so adapting their practices to make their finding aids conform to the database standard and putting energy into addressing technical barriers are not seen to be good uses of limited staff and resources. It is hard to argue with this point of view because it is not clear what difference the database has made to the institutions whose finding aids are there, or indeed, if many potential researchers are using it. In part this may be an issue of critical mass; the richer the database becomes, the more essential it will be for researchers working in the Philadelphia region and consequently the more importance institutions will attach to participating. It also has to be actively promoted, both to people doing historical research and within PACSCL.

Which leads us to institutional structure. How does a voluntary organization develop, maintain, and promote an ongoing project that depends

upon a robust technical infrastructure? For this program, the answer has been that we rely on one of our members, the University of Pennsylvania Libraries, to handle nearly all of the work and all of the costs. Penn provides and maintains the hardware, manages the software and the design of the site, and assigns a regular staff member in Special Collections to troubleshoot problems and advise PACSCL members. The staff person, Holly Mengel, was also the director of the first PACSCL CLIR grant, so she knows the system well and is deeply committed to its success. Nonetheless, Penn, not PACSCL, is paying all of her salary. Neither PACSCL nor individual PACSCL institutions provide financial support to help Penn maintain the site, and there is not a formal PACSCL advisory group to help with policies, troubleshooting, and promotion for the site, although there are individuals within PACSCL on whom Holly has been able to call for assistance.

PACSCL has been remarkably successful over the last 30 years at raising money for projects with limited duration, and these projects have had an enormously positive impact on the condition of historical collections in the region. But it is no longer enough just to catalog and process collections, something at which we have excelled over the years. If the cataloging, processing and, now, digitization work is going to be useful for our publics, we also need to figure out how access is going to happen, and that means some form of technical infrastructure that draws the collections together and enables scholars and students to make connections across our holdings. Now, the institutions with enough resources are building those structures individually, and because of the needs of their institutions and users, this is entirely understandable. The problem remains that many of PACSCL's members are independent cultural organizations that struggle with managing an IT infrastructure that is both increasingly complex and increasingly essential for making their collections visible to the larger world. In the long run, if not immediately, making alliances with better-resourced institutions is becoming essential. For the non-PACSCL members—the 150 historical organizations whose collections are in the finding aids database because of the Historical Society of Pennsylvania project—their collections will be visible only if they can partner with bigger players. Even the wealthier institutions are not always able to draw on IT support in the way they desire. Institutions with large IT departments also have large IT needs that can absorb all of the available energy, leaving limited IT support for special collections work. Finding ways to leverage developments at any one institution so that many more can benefit, in the way that we have done with the finding aids database at Penn, looks like a sensible way to operate.

The challenge will be to find a sustainable mechanism for building a technical infrastructure in an environment of mostly private independent institutions, and without a publicly funded institution, like the California Digital Library, to provide continuity and leadership. Depending on the public service ethic of a few of the larger institutions, especially the University of Pennsylvania Library, has worked reasonably well, but it leaves all of the participants dependent upon Penn's decisions and priorities for the maintenance of a system that many of us are coming to depend upon. Finding ways of providing financial support to Penn seems only fair, but also risks putting Penn in the uncomfortable position of being a vendor. The more money involved, and the more reliance other institutions place in the

system, the more likely that our very informal, low-cost arrangements will need to be replaced by memoranda of understanding and service level agreements. These might be important improvements for the long-term health of PACSCL, but they will also be a major change from the handshake agreements that we have operated with in the past. It also only seems fair and wise for PACSCL institutions to find a way of taking a more active role in overseeing, troubleshooting, and promoting the finding aids database. The support structures that were put in place for the grant projects have now gone away, so reconstituting them in a sensible, sustainable way seems to be the next order of business.

Finally, and not to be forgotten, are our users. Philadelphia is an extraordinary cultural center, with rich collections spread among some of the country's oldest and most distinguished institutions. This wealth of cultural institutions also means that collections are dispersed across the region, making the task of locating all of the relevant sources on a topic a daunting one for researchers. The Philadelphia region is full of situations where family papers are split among several institutions, or records of organizations that deal with similar social, cultural, and political issues are housed in different repositories, or documentation of neighborhood and community life is found in numerous and often unexpected locations. It is true that much of this documentation could be found through web searches, but not as efficiently and not with the detail and nuance that a regional database can supply, particularly for researchers interested in social issues where proper name searches are less productive. For our local users, including the numerous college and increasingly engaged high school students, a regional database has the potential to encourage more ambitious and comprehensive research by revealing avenues to sources that would otherwise have remained obscure. The CLIR grants have given Philadelphia institutions the opportunity to create an exceptionally useful tool for scholars, students, and the public doing historical research in the area. Our challenge now is to find a way of erecting an organizational and financial structure to keep it going.

▶ Part II: The Institutional Host's Perspective

David McKnight, University of Pennsylvania

In the previous section, Eric has amply described PACSCL as a "coalition of the willing" and the issues related to managing a large collaborative project of participatory members of the consortium. What I would like to discuss in the second half of this paper are some of the issues related to the pros and cons of administering such a project from the perspective of the host institution—in this instance my own institution: The University of Pennsylvania Libraries.

I would like to preface my remarks by citing another example: The Digital Scriptorium. Unlike PACSCL, the Digital Scriptorium is a national consortium of institutional libraries that have chosen to contribute selected digital images from each institution's collection of Medieval manuscripts. There are currently 36 members of the consortium. What differentiates the Digital Scriptorium from PACSCL is that it does have an institutional host: The University of California Berkeley Library. Without going into great detail, the Digital Scriptorium (DS, hereafter) began its existence at Berkeley in the late 1990s; however, when the DS's executive director, Consuelo Dutschke, relocated to Columbia University the project moved with her including the digital assets and the DS database. In 2010, Columbia University Libraries made it clear that it was no

longer willing to host the DS. With the news that Columbia was abandoning the DS, there was shock and disappointment, especially among the members of the Executive Committee, of which I was a member. Although Columbia was bidding farewell to the housing and maintenance of this noble primary source digital collection, there may have been a sigh of institutional relief. But the library was willing to continue to underwrite Consuelo's salary on a part-time basis so that she could devote her time to her directorship of the DS.

After a two-year search, the DS Executive Committee reviewed several offers to host the DS project, its database, and digital images. The winner was UC-Berkeley. Thus the project that was launched at Berkeley with NEH funding ten years earlier was now returning to Berkeley. The primary challenge of the "new" technology host was to migrate the cumulative data that was housed at Columbia and "cross-walk" the data into a new framework: the California Digital Library. Thus the technology host was located on the west coast and the curatorial center of the project remains on the east coast. Like other long-term digital humanities projects, the DS is faced with ongoing sustainability and maintenance issues: uploading new data, web development, database issues, and data storage, among others. As technology host, Berkeley, like other institutions, has had to off-load the cost of maintaining the DS site by charging membership dues with the goal of raising $25,000 to pay for the annual maintenance fee. Based on a set of standard metrics, each member of the DS pays a prorated annual fee toward the upkeep of the DS database.

Although the DS has a technology host, the DS has found it difficult to establish itself as a 501(c)3 organization. This makes it nearly impossible for the DS to apply to accept donations or apply for grants. Thus the DS is in a state of limbo with respect to applying for external funding. (To compound the matter, grants would be administered by UC-Berkeley and subject to California law, which might pose problems for both parties.) PACSCL's situation is not as complex as the Digital Scriptorium's, and it has the advantage of having 501(c)3 status. Nonetheless, PACSCL is still dependent on member institutions to serve as the administrators of major grant projects because it is such a small organization that a large grant would complicate its legal status and overwhelm its ability to manage the project. So if you look at the awards for the 2008 and 2012 PACSCL CLIR grants, the awardee is not PACSCL but rather the University of Pennsylvania. Penn Libraries was a founding member of the consortium and is, if I may say, modestly, among the larger and better-resourced institutions in PACSCL. With that said, at the time that PACSCL was planning to submit its initial application, I volunteered to inquire if Penn might serve as institutional host for the grant itself and also provide the infrastructure for supporting the project team for the duration of the project. Penn Libraries' Vice-Provost and Director of Libraries H. Carton Rogers endorsed the notion of Penn Libraries serving in both capacities to support the first CLIR grant in 2008 and the second grant in 2012. What does institutional support actually mean?

First it meant that Penn Libraries would serve as the 501(c)3 institution that submitted the grant application on behalf of the consortium. During the application authoring process, Penn's role as host institution was clearly defined according the following criteria:

- Project oversight
- Administrative support
- Human resources

- Infrastructure
- Technology
- Long-term sustainability

The project managers for both grants were hired as University of Pennsylvania employees (as were the graduate student processors), but were also expected to report regularly to the PACSCL board. In terms of project oversight, in the case of both CLIR grants there were two principal investigators, one representing PACSCL and the other the University of Pennsylvania Libraries. PACSCL was very fortunate in hiring two exemplary project managers. There was no need to remind the managers that the reputations of both PACSCL and Penn were dependent upon the successful completion of their respective projects.

Institutional support meant more than just providing office space, lights, heat, an Internet connection, a photocopier, and office supplies. These were included in the cost-sharing portion of the project budget. What was not included was the time of human resources staff, who assisted with the posting of jobs and payroll issues, and the library's business office, which managed the grant funds and generated the financial reports for inclusion in our interim and final reports.

Finding offices in a building with few free office spaces for staff proved challenging but not insurmountable; the one area in which Penn played a leading and essential role was technology. Apart from buying project laptops and setting up or moving phones and computers (during CLIR I the project staff were moved to another location in library because of construction on the sixth floor of Van Pelt Library), technology proved to be the most demanding aspect of these projects.

Penn had implemented its digital library application to search, retrieve, and display XML-encoded data, and through it, the library's technology unit had implemented a Penn EAD repository. Extending the model to the PACSCL project with 20 participants as opposed to Penn's three separate units was perceived by the Library Information Technology group as an opportunity to develop the library's digital library technology staff. It should be noted that in both CLIR budget requests we included a modest five percent of the total grant for technology support. The actual cost in terms of coding labor has not been quantified to date. As Penn finessed the EAD repository, the project managers took care of site technology problems: setting up local instances of Archivists' Toolkit, creating web folders, training staff, and troubleshooting. It was part of their job description.

Although the two CLIR PACSCL Hidden Collections processing grants have been completed, the question of whether Penn will continue to serve as host to the EAD PACSCL finding aids database has not been resolved. The data are secure. The finding aids are searchable and accessible. But what is the cost of long-term maintenance? I refer not to the actual data itself, but the tweaks to code, code errors, improvements, and possible migration of data to a new platform. These represent challenges for both PACSCL and the University of Pennsylvania to resolve over time.

Conclusion

This is a time of transition for both PACSCL and historical collections management as a whole. For many years our emphasis has been on reducing backlogs of uncataloged and unprocessed collections and placing the catalog records and finding aids into both our

institutions' library systems and the handful of national systems, particularly WorldCat. For the last 25 years, PACSCL has been a model of how a private group of cultural institutions can work together effectively to raise money to address backlogs and professional standards and capabilities in a major metropolitan area. The success in Philadelphia and nationwide in reducing backlogs, coupled with the growing amount of digitization, has now changed the landscape for cultural organizations, for with increasing amounts of collections metadata and digital images come new opportunities for scholars and students to access collections and undertake research in ways that were difficult in the past. This new research will depend upon easy access to data, and preferably access that draws collection data from across institutions in easily searchable databases.

We are already seeing numerous structures being created that take advantage of collections data from multiple institutions. In addition to the PACSCL Finding Aids Database, Penn supports the Provenance Online Project and the Early Novels Database, and has just brought up the OPenn site that will host the digital collections in the PACSCL diaries project currently under way, and the proposed collections of medieval manuscripts housed in the region. Outside of the Philadelphia region, Bryn Mawr has led an initiative among the schools formerly known as the Seven Sisters College to create a mechanism for searching the digitized collections of student letters, diaries, and scrapbooks, and there are many more multi-institution projects under way in other regions. At this time of excitement over digital technologies and their potential for transforming the ways in which scholars and students interact with cultural collections, it is understandable and laudable that libraries and archives are experimenting with developing new structures for providing access to their holdings. Finding a means of sustaining the most important of these structures will be the great challenge of the coming years.

References

College Women: www.collegewomen.org

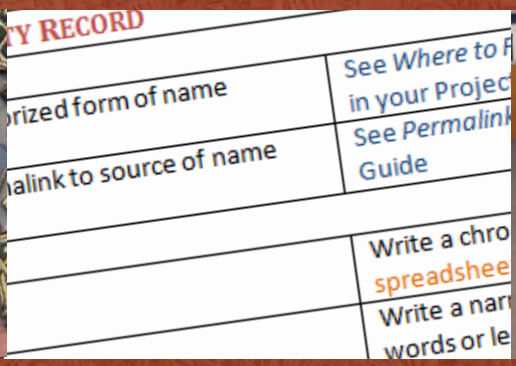
Digital Scriptorium: http://bancroft.berkeley.edu/digitalscriptorium/

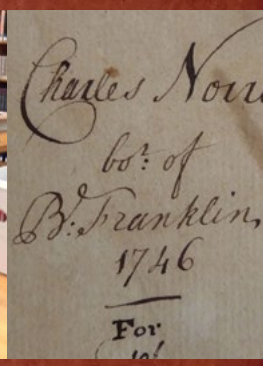
Early Novels Database: http://earlynovels.org/

OPenn site: http://openn.library.upenn.edu/

Provenance Online Project: https://provenanceonlineproject.wordpress.com/

Student and Faculty Involvement





Maximizing Partnerships: Faculty Buy-in, Service Learning, and Hidden Collections

Christopher Harter, Director of Library and Reference Services, Amistad Research Center

Elisabeth McMahon, Department of History, Tulane University

Abstract

Beginning in 2008, the Amistad Research Center (ARC), an independent, nonprofit special collections library and archives housed on the campus of Tulane University in New Orleans, initiated a plan to expand access to its collections. This effort has been aided by two Cataloging Hidden Special Collections and Archives grants from the Council on Library and Information Resources (CLIR). In 2009, ARC began its role as a community partner in Tulane University's service learning program, a post-Hurricane Katrina addition to the university's undergraduate curriculum. These developments have led to a collaborative partnership between ARC and faculty and students of Tulane University that not only provides faculty members with a willing, content-rich community partner that supports their course aims, but also affords undergraduate students a direct, hands-on experience working in a library and archival setting and brings to ARC additional assistance in exposing its hidden collections.

"I think that until I had actually worked in an archives, it was really difficult to get a sense of what that actually would mean. The class readings have given us an introduction to archival procedures and the work that archivists do, and it is really interesting to be able to put this into practice. Working with these documents is such a rewarding experience because I feel like there is so much to be learned, a whole side of history that I could not grasp in this way from just reading a book. It is also rewarding to feel like I can really make a difference in helping to shape the way historians for years to come will find access to this material. It is interesting to think about the fact that, if there were not people to organize and to create finding aids for these resources, it would be virtually impossible for historians to know what was out there or to gain access to it."

—Excerpt from Tulane University student's service learning reflection paper, 2009

Interesting, *rewarding*, and *access* are three words that stand out in the student reflection paper quoted at left. *Discovery* is perhaps a good way to describe what this student encountered. The term is certainly fitting to describe the outcome of the grant-funded initiatives at the Amistad Research Center (ARC) over the past seven years and ARC's efforts to partner with faculty and students. That partnership aims not only to expose both parties to the rich sources within the collections, but also to transform them into stakeholders in order to turn hidden collections into bountiful discoveries.

The Amistad Research Center and Efforts to Expand Access

The Amistad Research Center is the nation's oldest and largest independent archives and special collections library that chronicles race relations, civil rights, and ethnic history in the United States. Founded in 1966 as an outgrowth of the United

Church Board for Homeland Ministries' Race Relations Department at Fisk University, an early civil rights training ground, ARC incorporated as a 501(c)3 nonprofit in 1969 and today resides on the campus of Tulane University in New Orleans. Originally founded as a repository for the archives of the American Missionary Association—a Christian abolitionist organization that later worked to found schools for the Freedmen following the Civil War—ARC expanded its collections into the areas of race relations, the Civil Rights Movement, the Harlem Renaissance, and other aspects of the history and culture of ethnic communities in the United States.

This expansion was based in part on an aggressive acquisitions policy that focused on collecting materials documenting underrepresented peoples at a time when such a strategy was not widely emphasized by cultural institutions. However, budgeting and staffing for this nonprofit organization did not always allow for the arrangement and description of these collections according to professional archival standards. Over time, ARC became a well-established research center on the topic of racial and ethnic history. However, a significant portion of its collections remained underused by scholars and the general public because they had not been processed or cataloged, which limited access to holdings. Beginning in 2008, staff sought to change this by embracing the idea of expanding access through the use of increased technology and the pursuit of grants and partnerships.

The Amistad Research Center received the first of two CLIR Cataloging Hidden Special Collections and Archives grants in 2008. The initial grant allowed for processing and cataloging specific civil rights-related collections within its holdings, but the major success of that grant was that it aided in re-envisioning how access could be interpreted and demonstrated in a variety of ways to ensure that ARC fulfilled its basic mission. This was most notably seen through the implementation of collections management software and the development of a comprehensive manual that outlined policies on access, collection development, exhibitions, security, and other areas. The manual also detailed procedures in preservation and handling, processing and cataloging, and data entry for the collections management software. In short, the manual now serves as the major form of documentation for staff, administration, and trustees, as well as a comprehensive training manual for new staff, students, and volunteers.

The completion and development of professional policies and procedures as deliverables for the first CLIR grant helped in planning for the 2011 grant, which called for processing records generated by the American Committee on Africa (ACOA) and its sister organization, the Africa Fund. These related organizations were based out of New York and founded during the 1950s and 1960s, respectively, with the goal of educating the U.S. populace and policy makers on anti-colonial and anti-apartheid movements in Africa. Totaling almost 600 linear feet of organizational records, as well as publications and ephemera collected from groups and individuals throughout the African continent, these collections contained a wealth of information and resources for anyone studying U.S.-Africa relations on the macro level and individuals and groups working to end colonialism and apartheid within individual countries. The additional impetus for targeting these collections for processing and cataloging was a growing interest among researchers in accessing these collections, based on the results of a 2009 partnership with a faculty member and students at Tulane University.

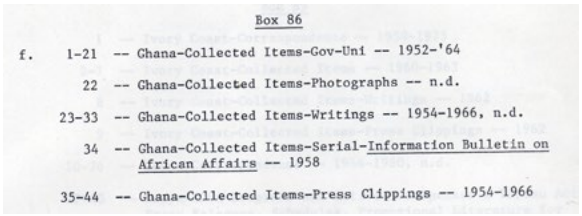
Service Learning and the Development of the Archiving Africa Course

In the spring semester of 2009, ARC partnered with Elisabeth McMahon of Tulane University's department of history on a course called Archiving Africa. Dr. McMahon's goal for this course was to expose her students to the rich Africana holdings found within ARC's archival collections as a means of developing an appreciation of historical research methods and to introduce them to the benefits and difficulties of researching and interpreting primary sources. This course was a research-level seminar as well as a service learning course in which students would each commit 20 hours of volunteer time to ARC during the semester.

Service learning as an educational concept emerged in the 1970s. First viewed as experimental, it has gained in popularity on college and university campuses. This concept is defined in various ways, but to quote the National Service Learning Clearinghouse, it is "a teaching and learning strategy that integrates meaningful community service with instruction and reflection to enrich the learning experience, teach civic responsibility, and strengthen communities." Tulane University's Center for Public Service defines service learning as "an educational experience based upon a collaborative partnership between the university and the community. 'Learning by doing' enables students to apply academic knowledge and critical thinking skills to meet genuine community needs. Through reflection and assessment, students gain deeper understanding of course content and the importance of civic engagement."

Tulane University's requirement that all incoming undergraduate students complete one semester of service learning coursework and an additional semester of independent service learning work with a community partner began in 2006. Elisabeth McMahon joined the history department the following year. As a professor teaching African history, she sought to locate all of the Africana-related collections in New Orleans. Although there were a number of collections available in the city for students to work with, ARC offered the largest collection and one with the most coherent link with the African continent. In designing her class, she wanted to find an archives to consistently work with over time. As a faculty member, she knew that working with an archives in this way would require a considerable amount of time and work, thus it was ideal that ARC had enough materials for her students to use year after year. Ease of location was also a factor. That ARC qualified as a community partner and nonprofit institution, yet was located on Tulane's campus, made it much easier to access geographically than other collections around the city.

McMahon and ARC have partnered three times for the Archiving Africa class. The first two classes, held in 2009 and 2012, focused on the ACOA/Africa Fund records. Although a portion of the ACOA records (about 146 linear feet) had been previously arranged in 1983, the resulting finding aid was not sufficient to provide access to the myriad letters, speeches, brochures, periodicals, and other documents that formed the collection. Lack of deeper access to the collection had long been an issue for staff and researchers.



Box 86

f. 1-21 -- Ghana-Collected Items-Gov-Uni -- 1952-'64

22 -- Ghana-Collected Items-Photographs -- n.d.

23-33 -- Ghana-Collected Items-Writings -- 1954-1966, n.d.

34 -- Ghana-Collected Items-Serial-Information Bulletin on African Affairs -- 1958

35-44 -- Ghana-Collected Items-Press Clippings -- 1954-1966

Fig. 1: Example of initial finding aid for the ACOA records, illustrating lack of description.

Of particular interest to researchers were the files for ACOA's activities within various African countries, which number more than 70 cubic feet. These include an extraordinary trove of correspondence between activists, U.S. government officials, African political leaders, organizers, students, and others. The ACOA records provided an ideal trial project for a course-based undergraduate service learning experience for several reasons. First, staff had identified the need to index and inventory portions of the collection as a project that could not be undertaken without staff help. Second, students would work on a large but singular collection, which would provide opportunities to discuss common issues and methods during their work. Third, the collection was diverse enough that students could use it for seminar papers and the professor could use documents for a number of class discussions.

Over the course of the semester, students indexed correspondence and inventoried ephemera in more than 25 boxes, far exceeding Amistad's expectations. (For the purposes of the initial service learning class, students merely listed sender and recipient, date of letters, and their location within the collection as an aid to locating letters by particular individuals. They did not index or create transcripts for the content of the letters.)

The second Archiving Africa class, as well as ARC's second CLIR grant, were the direct results of this initial foray into a service learning partnership. The 2011 grant called for the processing of multiple, later addenda to the ACOA records (140 linear feet) and those of the Africa Fund (350 linear feet). Additionally, 135 boxes of books, pamphlets, newspapers, and other publications collected over the years by the two organizations were to be cataloged within ARC's library holdings. As part of the work plan for the second grant, Dr. McMahon served as a subject-area specialist, helping to identify individuals, organizations, events, and related topics within the collections, which provided more authoritative description during the organization of the records. Students in the course focused their service learning on helping

Sender	Recipient	Date of Letter	Box	Folder
Nyerere, Julius K.	Houser, George M.	1958, 3-25	142	32
Kawawa, Rashidi M.	Jonas, Gilbert	1958, 7-1	142	32
Houser, George M.	Kawawa, Rashidi M.	1958, 7-15	142	32
Munanka, Bhoke	Houser, George M.	1958, 7-18	142	32
Houser, George M.	Munanka, Bhoke	1958, 10-7	142	32
Munanka, Bhoke	Houser, George M.	1958, 10-17	142	32
Chamshama, S.S.	Houser, George M.	1958, 11-6	142	32
Houser, George M.	Chamshama, S.S.	1958, 11-24	142	32
Chamshama, S.S.	Houser, George M.	1959, 1-12	142	32
Chamshama, S.S.	Houser, George M.	1959, 1-28	142	32
Chamshama, S.S.	Houser, George M.	1959, 2-3	142	32

Fig. 2: Example of correspondence index spreadsheet for 2009 service learning class

staff with the initial collection survey of the records, publications, and audiovisual materials, and with compiling the container lists that were then used to create processing plans for both sets of records. (See Appendix for a short project description.)

Partnering Hidden Collections and Service Learning

The Amistad Research Center's integration of a service learning collaboration with Dr. McMahon's Archiving Africa course and the adoption of a hidden collections focus have yielded great results for the students, the faculty member, ARC, and the community of Africanist scholars. As seen in the definitions above, service learning as a concept entails both educational and community outreach components. Part of the collaboration's success is based on the belief by both the faculty member and community partner that service learning is successful when it fully and equally embraces both service and learning. This shared belief is fundamental and has been communicated to students. Students who most effectively embraced their service were those who considered their work in light of ARC's mission and goals, and who thought about the roles of archives and libraries within their communities and society in general. This is a fundamental goal of Dr. McMahon's course. Following are some of the objectives for her Archiving Africa course:

- Introduce students to the practice of history through work in archives, archival methods, and archival research
- Introduce students to the methodology of historical practice and, in particular, the methods of African history
- Allow students to consider the methods necessary to preserve the history of ordinary individuals and non-elite institutions
- Help students see the value of archives for building community identity
- Consider best practices for archives, especially related to collection development and preservation

In addition, during their initial introduction and training at ARC, students were introduced to its history and mission, and they learned how expanding access and having the support of granting agencies such as CLIR aid in fulfilling that mission.

Working with students as part of a service learning course is very different from overseeing student internships, putting on show and tell visits, or conducting inquiry-driven collaborations between librarians/archivists and faculty. Service learning requires that librarians and archivists simultaneously supervise multiple students over the course of a semester. It also requires librarians and archivists to be co-teachers with the faculty member by integrating service learning into course content. ARC staff have had to adjust to different levels of student interest, motivation, and abilities. Although staff encountered a few students who required Dr. McMahon's intervention to spur them to complete their service hours, most students took the service learning to heart and worked diligently at ARC.

Student reactions to the service learning and their work at ARC are seen in their (required) reflection journals and course evaluations. At least one student listed service learning hours as "kind of boring," but most had a more positive experience:

- "The service learning was directly related to a potential path in history, so it was cool to work in an archive."
- "[The service learning] tied in very well with

the course and I felt like I was doing something useful."

- "All of the experience I have been gaining through the readings we have done and the work and research I have begun at ARC has really brought new aspects to my understanding of archives and how they function."

Not only did the students find the service learning of interest, they found the opportunity to engage in extended archival research rewarding as well. As Dr. McMahon reported one semester, "In over ten years of teaching, I have never seen a class embrace and take ownership in a research project as this class did with the ACOA materials."

Fig. 3: Service learning students from Dr. McMahon's class at work.

Perhaps the most rewarding outcome of the service learning classes was to see students return to ARC in various ways. Students became more knowledgeable of ARC as a valuable on-campus resource, knowledge they then spread to their peers. In addition, some students returned to ARC to complete individual internships or work as part-time student employees. One student from the first class was hired as a student assistant and worked as part of the second CLIR grant because of her experience and interest in continuing to work with the ACOA records.

Dr. McMahon's exposure to the ACOA records, both as teacher and scholar, has added to the content of her Archiving Africa course as well as to other courses. For the service learning course, students' experience at ARC has often influenced class discussions. Dr. McMahon encourages students to describe their own work and to share discoveries of relevant resources and documents that may be useful to their classmates as they develop seminar papers. Dr. McMahon finds that students approach primary sources in very different ways when they are conducting research, as opposed to when they are inventorying or describing the materials. While the first often entails students narrowly searching for specific information or documents, the second allows them the leisure of exploring the collections more fully. She found that "By working in the collections first, they have an opportunity to explore and have a better sense of the sources before they decide on a paper topic. It gives them a better sense of what historians really do. Rather than simply searching databases, the service learning allows students to do research in a way that is more holistic to the practice of the profession. They get the chance to work with uncataloged materials and to discover what is hidden within."

Most of the students in the classes were history majors. As the Archiving Africa course incorporates a strong methodology component both in class discussions and during the service learning, Dr. McMahon has found that students gain a better understanding of what historians do in terms of working with documents. They also gain a better understanding of professions that are allied to the field of history, such as those of archivists and librarians. In course evaluations, students repeatedly felt they were making a difference through the work completed at ARC,

viewing the outcome of their service learning as something tangible, with long-term impacts.

During the most recent Archiving Africa class, in 2014, Dr. McMahon incorporated scholarship that resulted from the increased access to the ACOA records provided by the first two classes and funding from CLIR. This increased access to archival collections led to new scholarship, which serves as the basis of classroom discussion for future students. That new scholarship is a fitting legacy of the partnership. This collaboration, as well as the receipt of support from CLIR in expanding access to ARC's collections, resulted in a sustainable and adaptable partnership that focuses on the merger of service and learning to the benefit of all involved—faculty partner, students, ARC, and its global constituency of scholars and researchers. The collaboration has allowed us all to become stakeholders in turning hidden collections into bountiful discoveries.

References

Center for Public Service, Tulane University: http://tulane.edu/cps/students/servicelearning.cfm.

National Service Learning Clearinghouse: http://www.servicelearning.org.

Appendix

Short Version of Project Description for Spring 2012 Service Learning Class

Project Description

The Amistad Research Center will serve as the community partner for the service learning component in HISB497. The Center has received funding from the Council on Library and Information Resources to process the remaining portion of the ACOA/The Africa Fund records. Students from HISB497 will assist the ARC by conducting preliminary inventories and surveys, physically shifting the collections, and conducting preliminary preservation work on the materials. A series of tasks will be developed and students will be allowed to sign up for tasks based on their class schedules. Initial tasks will concentrate on these areas:

- Inventory audiovisual materials within the collection to provide information on format, date, title, etc. In addition, students may index VHS tapes containing episodes of "South Africa Now," a weekly news program produced by The Africa Fund from 1988-1991.
- Inventory books, pamphlets, and periodicals produced by ACOA/The African Fund or collected by the organizations. Many of these were produced by organizations within various African nations and are not widely held by library institutions within the United States.
- Shift and separate records of ACOA and The Africa Fund at Amistad's offsite facility. Will require lifting boxes of 40 lbs.
- Create container lists summarizing content, formats, and dates of records.
- Begin preliminary preservation work and organization of file units within the records.

Students will be given detailed instructions for this project and will partake in an initial introduction to the Amistad Research Center and

training session. Amistad staff will discuss issues of access and preservation in archives and how these relate to the project at hand.

For the inventorying of audiovisual and printed items, using an Excel spreadsheet, students will enter information according to set criteria. Students will also be given composition books with which to record questions, organizational information, or related documentation.

Goal

The goal of this project is to enhance access to the records of the American Committee on Africa and The Africa Fund for researchers and scholars worldwide. Student assistance will help the ARC to jumpstart its work on this grant-funded project and aid staff in later phases of the project. The project will also introduce students to resources that may assist them in their coursework for HISB 497.

Collaboration and Education: Engaging High School Students with EAC-CPF Research*

Valerie Addonizio and Christopher Case, Sheridan Libraries, Johns Hopkins University

Abstract

The Special Collections Research Center at Johns Hopkins University received a CLIR Hidden Collections grant for processing the archives of the Roland Park Company from March 2013 to March 2014. The grant included partnerships between the university's Archives and its Technical Services department to create best practices for EAC-CPF records, as well as a partnership with a local high school history class to complete the required research. Rather than simplify ISAAR (CPF) and EAC-CPF for the students, project staff distilled them into discrete, easy-to-understand tasks that allowed for the production of controlled data in a high school environment.

Project Background and Grant Requirements

The Ferdinand Hamburger Archives is the official archival repository for Johns Hopkins University's Homewood campus divisions, the School of Education, the Carey Business School, and the Paul M. Nitze School of Advanced International Studies. Archival holdings include the business records of the university as well as a substantial body of manuscript collections documenting a variety of research areas including the history of science, literature, higher education, politics, and regional planning and development. In 2010 the archives received the Roland Park Company Records as a transfer from Cornell University. This collection, which focuses on the development of several important Baltimore neighborhoods, together with the Martin L. Millspaugh Archives, another prominent Baltimore collection gifted to us, was the impetus to write a CLIR Hidden Collections grant in 2011.

The project outlined here is a result of that grant, which included the development of a set of EAC-CPF (Encoded Archival Context-Corporate bodies, Persons, and Families) best practices through collaboration with the university's Archives and its Technical Services department; collaboration with a local high school for research and identification of biographical and related archival holdings information; and the adaptation of that information to create EAC-CPF records.

The success of this project was measured in three ways: first, against the terms laid out in the grant; second, in terms of the students' experience; and third, in terms of whether the quality of the records hewed to the best practices developed by the project team.

Formulating Local Best Practices

Because the vast majority of the EAC-CPF standard consists of optional elements, the team (consisting of the project archivist and the content management librarian representing technical services) knew that it would benefit all parties, and potentially the wider archival community, to

* This paper is an abridged version of Addonizio and Case 2014.

create a series of local best practices for implementing EAC-CPF. Background research included evaluating best practices and draft records in a number of EAC-CPF instances, including those created by Harvard and Yale (Diaz and Pyzynski 2012), Tufts (DAT-040), and the SNAC (Social Networks and Archival Context) project. For a full treatment of this process, see Addonizio and Case 2014.

Collaborating

After creating best practices, we began the collaboration between Johns Hopkins and the local high school. This involved the Johns Hopkins University archivist, the Roland Park Company records project archivist, the class's history teacher, and two school librarians.

The idea for this collaboration came out of a simpler question of whether a senior high school history class could come to campus for a project involving primary sources. Then we came up with the idea of doing EAC-CPF research. In conversation with the educators, we learned that they hoped to contextualize the assignment to be as much about research using archival material as about a real-life application of the work. In other words, the educators were interested in treating the assignment like a client-based work order with the client, Johns Hopkins, having a real-life request and the students having to follow strict instructions to deliver a usable product. This expectation was significant, because there were a number of times when Johns Hopkins and the school staff questioned whether the requirements of the project would overreach the normal workload or comfort zone of high school students.

Distilling

Because the educators were undaunted and, in fact, enthusiastic about the project's complexity, we moved forward with trying to distill our best practices and the encoding realities of EAC-CPF into a project that could capture the complexity of ISAAR (CPF)/EAC-CPF and be managed outside our immediate supervision.

The initial idea was to provide the students with a single spreadsheet to fill out for each entity. Yet the reality was that we asked for more than the biographical and relationship information suggested by ISAAR (CPF). We also asked for controlled data required for encoding that reflected our best practices.

An example using <placeRole> helps demonstrate the challenge. For the actual encoding of EAC-CPF, our best practices limit the value of <placeRole> for a person to one of birth, residence, education, marriage, occupation, travel, death, or burial. We could not ask for only a list of predominant places added to a spreadsheet. We had to limit the places to those prescribed by our best practices, and the students needed to know that. This is the simplest example of the complicated interplay of content and requirements that we had to anticipate.

Given this complexity, oversimplification was a risk. Therefore, we decided not to simplify the requirements (i.e., fill out this single spreadsheet), but the method by which they could be fulfilled. As a result, we provided students with a suite of four documents, each of which helped explain what we were looking for (content) and how we needed to see it (controlled data, including authority work). The documents use the ISAAR (CPF) section numbering system like a primary key, allowing a student to cross-reference the documents. Our aim was that, following a general introduction to the project, we would only need to explain how to use the documents, then the documents would stand on their own. What follows is a description of each document, followed by how students used it.

The ISAAR Roadmap. This is a blank ISAAR (CPF) record repurposed as a roadmap that lists every requirement that needs to be met, and then points students to what that requirement means (content), and where and how to fill it in (controlled data). The fields are ISAAR (CPF), but the instructions combine our best practices and the requirements of EAC-CPF. The roadmap's secondary purpose is to help delineate between two major concepts—gathering information about corporate bodies, persons, or families (ISAAR 5); and gathering information about related resources on corporate bodies, persons, or families (ISAAR 6). For instance, it helps to point out that the bibliography the students write for the <biogHist> is different from the list of related resources. We created two roadmaps—one for persons and one for corporate bodies.

By referring to the ISAAR roadmap, the first thing the student sees is the ISAAR (CPF) section "5.1 Identity" (Figure 1). In that section the first task is to find the "5.1.2 Authorized form of name." That has two components—the authorized form of name, and a permalink. What does that mean to the students? Further to the right, the roadmap points to two sets of instructions—"Where to Find Names" and "Permalinks" in another document (the project guide). After students read those instructions in the project guide, the roadmap shows that the data are to be entered on tab 5.1 on the project spreadsheet.

The Project Guide. At 10 pages, the guide is weighty, but it was written as a reference document rather than a long exposition. It provides definitions and context to certain usage, detailed instructions on how to get authorized names, places to look for related resources, and ways to determine permalinks and list relationships. Each of the sections is referenced from the ISAAR roadmap. We purposely used informal language and we aimed for relatable examples (see Figure 2).

The Spreadsheet. Ultimately, we provided a spreadsheet for all the data (except the narrative biography/history), but with multiple tabs, each of which represented a different ISAAR (CPF) section. Using a spreadsheet has two benefits: first, spreadsheet fields are easy to export

ISAAR 5. Elements of an Authority Record		
5.1 Identity		
5.1.2 Authorized form of name Use spreadsheet tab 5.1	Authorized form of name	See *Where to Find Names (CPF)* in your Project Guide.
	Permalink to source of name	See *Permalinks* in your Project Guide
5.2 Description Area		
5.2.2 History		Write a chronology using spreadsheet tab 5.2.2
		Write a narrative biography (800 words or less) and submit it using this link.
5.2.3 Places Use spreadsheet tab 5.2.3		See *Where to Find Place Names* in your Project Guide
5.2.5 Occupations Use spreadsheet tab 5.2.5		See *Where to Find Occupations* in your Project Guide

Fig. 1: ISAAR roadmap

Associative

This is the broadest relationship and the one you will use the most. Basically, if the relationship you're describing isn't one of the others in this list, use Associative. The important part is really in the description, which allows you to briefly state how the two entities are related.

Examples:

Mutual relationships If the relationship is mutual, like the ones listed below, the description does not need to imply the "direction" of the relationship.	**Directional relationships** Here are examples of when the direction of the relationship is described in the description.
Main topic: John Lennon **Associative relationship:** Paul McCartney **Description:** Band mates in The Beatles.	**Main topic:** Steve Jobs **Associative relationship:** Apple Computer, Inc. **Description:** Company founded by Steve Jobs.
Main topic: Ben Cohen **Associative relationship:** Jerry Greenfield **Description:** Co-founders of Ben & Jerry's Ice Cream.	**Main topic:** Pew Charitable Trust **Associative relationship:** National Public Radio **Description:** Receives financial support through charitable donations.

Fig. 2: Sample from the project guide

and manipulate, and would be advantageous in post-production; and second, drop-down fields allowed us to control vocabulary dictated by our best practices. We created two spreadsheets—one for persons and one for corporate bodies.

The narrative continues using the <placeRole> example from above. The roadmap indicates that the information for "5.2.3 Places" be added to tab "5.2.3 Places" on the spreadsheet. The cells in that tab limit students to the list of values for that element, allowing our best practices to guide student data entry (Figure 3).

The actual ISAAR (CPF) standard was used for conceptual context. We provided it just in case it helped the educators or students to pin down what we were looking for. Importantly, the numerical sections provide the vital framework for enabling all the documents to enter into a relative relationship.

Acquainting Students with the Project

After writing and testing the suite of documents in August 2013, we introduced students to the project in September. We developed two presentations. The first dealt with the nature of archival material and, significant to the project, examples of real finding aids. The second was an introduction to ISAAR (CPF), and consisted almost entirely of reading the ISAAR roadmap document together as a class. Students opened all four documents on their laptops and went from document to document, as the roadmap instructed.

We also discussed the difference between ISAAR 5 (the section that defines biographical information, familiar to students) and ISAAR 6 (the section that defines related collections, and was a new concept). At the end of the session, one student said, "But you can't have an ISAAR 5 without an ISAAR 6," meaning you can't have information about an entity without sources about

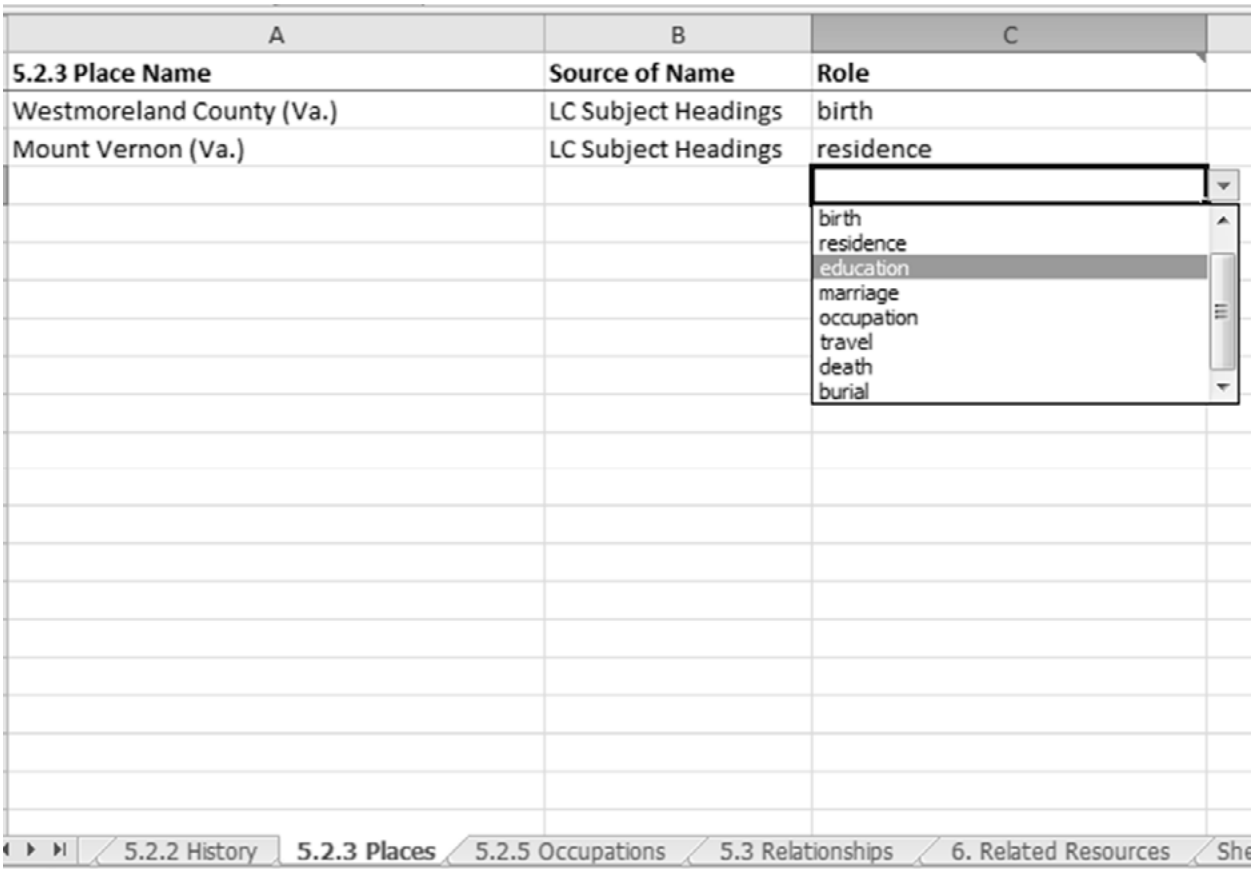


Fig. 3: Sample from the person spreadsheet

that entity. This statement demonstrated a fundamental understanding about the relationship between biographical content and the archival collections that contain it, and was a cause for team celebration.

Over the next few weeks, students e-mailed us some of their questions. Two examples reflect an increasing awareness of the complexities in creating archival authority records, suggesting that the students were participating in a valuable learning experience:

"When you have a person who has the same occupation for overlapping or several years, how do you differentiate? Example - Kessler was an urban planner, landscape architect, for several jobs, overlapping in time."

"I had a question regarding my ISAAR 6 on Charles Grasty. I have found some archival materials that are individual and not members of larger archives [*i.e., catalogued individually in Worldcat*]. They therefore don't have finding aids. Should I instead provide the Worldcat permalink of the source?"

Results

We reviewed the data twice before to the end of the project, in November and in late December 2013. In both cases, students provided sample records, which we reviewed in detail and found were quite solid. The students returned their data in January 2014, when the project archivist set to work adapting the information into EAC-CPF.

We immediately observed that the results

required careful proofreading before the data, especially the biographical or historical notes, could be adapted. Students used subjective language throughout. The only expectation was that the records be "Wikipedia-worthy," but nearly every note still needed our attention. Also, because of the complexity of the instructions, some students entered information in the wrong fields, requiring minor corrections and some additional research. Furthermore, we anticipated that almost all dates would need to be normalized, but mistakenly expected that we could automate that process.

Upon reflection, we realize that date normalization is something that can be easily explained to students in a future version of this project. Other challenges that arose with dates include incorrect dates, the recording of circa dates (something we had not anticipated and for which we take responsibility for not explaining), incomplete date ranges, and missing dates. Finally, in some cases there was missing and incomplete information, which ranged from not listing a person's occupation to not submitting any biographical or historical notes. We do not find the students primarily at fault for these issues. All of these challenges can be better addressed in the future with careful instruction and ongoing evaluation by the archives team.

Analysis

As stated above, project success was measured in three ways. First, in relation to the terms laid out in the grant, we found that 15 EAC-CPF records were created in accord with EAC-CPF emerging best practices. Archives and Technical Services collaborated on a set of best practices. We were successful in translating the complexity of EAC-CPF requirements to high school students. And virtually all of the students demonstrated a fundamental understanding of what was required.

Second, we measured project success by responses from students and educators. Everyone at the school agreed that this was an exceptionally unique and meaningful project for students and a

Names	Chronology	Names	Sources	Bioghist	Occupations/ Functions	Places	Relationships	Related Resources
Atterbury, Grosvenor	X	X	N/A	X	X	X	X	X
Baltimore Country Club	X	X	X	X	X	X	X	X
Baltimore Sun	N/A	N/A	N/A	X	N/A	N/A	N/A	N/A
Bouton, Edward J.	X	X	N/A	X	X	X	X	N/A*
Fowler, Lawrence Hall	X	X	N/A	X	X	X	X	X
Grasty, Charles	X	X	N/A	X	X	X	X	X
Kessler, George Edward	X	X	X	X	X	X	X	X
LaMotte, Frank Linton	X	X	N/A	N/A	X	X	X	N/A
Maryland Transit Authority	X	X	N/A	X	X	X	X	X
Mowbray, John McConkey	X	X	X	X	X	X	X	N/A*
Olmsted Brothers	X	X	N/A	X	X	X	X	X
Olmsted, Frederick Law (Junior)	X	X	N/A	X	X	X	X	X
Palmer, Edward Livingston	X	X	X	X	X	X	X	X
Russell Sage Foundation	X	X	N/A	N/A	X	X	X	N/A
Urban Land Institute	X	X	N/A	N/A	X	X	X	N/A

*-Some missing information was discussed with the Project Archivist before the records were returned, and it was determined that this information truly did not exist and the students were not accountable for it.

Fig. 4: Analyzing the results

terrific example of experiential learning. They are eager to collaborate again. The Archives shares this sentiment.

The third measure is whether the quality of the EAC-CPF records hews to our best practices. Our assessment of this aspect of the project is mixed. Of the 15 EAC-CPF records created, only three were entirely complete according to Hopkins's best practices. Further, four could not be used because they would require more research to be complete. The remaining eight could be considered complete by a slightly less strenuous standard. If <sources> were not required, the number of complete records would increase to eleven.

We encountered a notable variation in record accuracy, quality, and comprehensiveness, because of the varied set of skills that each student brought to bear on the assignment. Figure 4 indicates which sections of the EAC-CPF record were provided. The extent to which these submissions meet our standards for publication was inconsistent. Over the course of the encoding, the project archivist found inaccurate dates, inconsistent dates, incorrectly formatted sections, and factual errors, suggesting that we should have had more direct collaboration with educators and students to better communicate our expectations. One of our most important lessons learned is that not every student should be expected to submit errorless work. In an actual work setting, professional staff do not just evaluate a product; supervisors review staff work and provide feedback until the product is worthy of submission. When partnering with students on a product, it should be the responsibility of the professional staff (here, the project archivist) to perform quality control and work closely with them to refine the product until it contains no errors.

Because very few of the records were entirely accurate or complete, we believe that in this third evaluation measure—the accuracy of EAC-CPF records vis-á-vis our best practices—the project was a complete success in terms of being a proof of concept, though the proof was in the negative. However, we also believe we can significantly address critical concerns in a future project.

These concerns fall into three areas. First, we believe that some of our quality control challenges raise interesting questions. For instance, does the quality have more to do with the age and experience of the students than with the method? Some of the work was excellent and complete, but was asking for this level of detail from high school students too much, even with educators' encouragement? It is possible that a similar project would be more successful with college-age students and a virtual environment made for populating EAC-CPF records.

Second, because this was a high school research project, students were not allowed to simply mine Wikipedia for the answers to their questions. We wonder if doing that for especially well-known entities such as, say, Frederick Law Olmsted Jr., would have been more successful than conducting a short-lived and limited-resourced research project on that individual. Allowing for repurposed information might have yielded fewer inaccuracies and gaps.

Third, we learned much about how and when to review data during a project. We conducted two reviews prior to final submission, and both times we deemed the data accurate. Unfortunately, those samples ended up being the most accurate records the students had produced. One recommendation is to review all the work being done, not only a few sample sets. Also, we would recommend there be two submission dates rather than having the project be the final assignment. This would allow for iterative grading,

something the educators might not usually do. Those students whose work did not meet our standards in the first submission would be given the chance to complete the records before final submission. We need to better account for the discrepancy between expectations within a learning environment and those within a professional environment.

As for the students, we believe they digested a seriously complex assignment, became familiar with archives and archival holdings, did some original historical research, tested the limits of their capacity for following instruction, and gained a greater appreciation for the real-life application of historical research. The variations in accuracy of their final data should not diminish the accomplishment of undertaking this complex assignment.

Conclusions and Observations

The success of this project depended heavily on the active participation of two school librarians. These individuals were essential to both the teacher and the students in interpreting and understanding the ISAAR (CPF) standard. One of the librarians became the principal contact after the project began, and asked questions that could only be articulated, understood, and applied by another information professional. Her fundamental understanding of how to apply the standard allowed her to be on the frontline of student questions, and helped assure us that the interpretation would be accurate. Without these two librarians, whose participation we accounted for in the project design, the work would likely not have gone so smoothly. The "do not try this at home" disclaimer can be qualified with "unless under the direct supervision of an information professional."

A second observation coalesced as we thought about the relationship between ISAAR 6 (relationships to resources) and the real-life advantages of EAC-CPF. In this project we tasked students with manually gathering biographical or historical information for an individual or corporate body, then seeking out and finding related collections. Clearly, this was a time-intensive approach. Like others interested in EAC-CPF, we contemplated the pros and cons of both manual input and automation models. At the same time, unrelated to that debate, we framed a lot of project discussion around the differences in the need for accuracy between the information required for ISAAR 5 and that required for ISAAR 6. We found that we preferred that the accuracy lie on the ISAAR 5 side of the record (accurate birth dates, lists of related corporate bodies, persons and families, occupations), but we could never expect an exhaustive list of every related archival collection for every entity.

Finally, we hope that this case study of collaboration and EAC-CPF will be useful as the profession moves forward with the standard. Although the future of EAC-CPF will likely include automated data harvesting such as that done by SNAC, this project demonstrates that well-designed and monitored cooperation may also play a significant role. Although automation offers the prospect of generating mass numbers of records, the reality remains that many older EAD finding aids do not contain the level of detail and nuance that EAC-CPF allows for, and archivists' workloads can only benefit from well-organized collaboration. We also hope that the outreach component of this collaboration provided a meaningful experience to the students and faculty that extends beyond the project outcomes produced.

We anticipate that all related project material, including the best practices and three student documents, will be available on the CLIR Hidden

Collections website at http://www.clir.org/hiddencollections/registry/hc.0880.

References

Addonizio, Valerie, and Christopher Case. 2014. "Collaboration and Education: Engaging High School Students with EAC-CPF Research." *Journal of Archival Organization* 12, 1-2.

DAT-040 DCA Best Practices for EAC. Available at http://sites.tufts.edu/dca/files/2011/03/DAT-040DCABestPracticesforEAC.pdf.

Diaz, Enrique A. and Susan C. Pyzynski. 2012. Connecting the Dots: Using EAC-CPF to Reunite Samuel Johnson and His Circle, Harvard University and Yale University. Available at https://wiki.harvard.edu/confluence/display/connectingdots/.

SNAC Prototype, The Social Networks and Archival Context Project. Available at http://socialarchive.iath.virginia.edu/xtf/search.

Engaging Students in Complex Description: Two CLIR Hidden Collections Projects

Lois Fischer Black, Curator of Special Collections, Lehigh University

Ilhan Citak, Archives and Special Collections Librarian, Lehigh University

Gregory A. Edwards, Project Archivist, Lehigh University

Andrew Stahlhut, Project Cataloger/PhD candidate, Lehigh University

Abstract

Lehigh University's Special Collections received two CLIR Cataloging Hidden Special Collections and Archives grants. In 2009, Lehigh was awarded a grant to partner with the Moravian Archives (in Bethlehem, PA) on the cataloging project The Moravian Community in the New World: The First 100 Years. In 2013, Lehigh was awarded a grant to process six civil engineering collections. In both projects we used a student workforce. Here we summarize efforts to meet the goals of both projects. We provide background on both grants, and include two student essays about the experience.

Introduction

Lehigh University received two CLIR Cataloging Hidden Special Collections and Archives grants. The first one, in 2009, supported the project The Moravian Community in the New World: The First 100 Years, and was conducted with the Moravian Archives in Bethlehem, Pennsylvania. Both Lehigh and the Moravian Archives hold collections documenting the material culture, religious values, and cultural diversity of the Moravian community of Bethlehem from its founding in 1741 until the opening of the community to non-Moravians in 1844 and the subsequent incorporation of Bethlehem in 1851.

The records reflect the multifaceted life of this transatlantic community in its interaction with other cultures. Because of church control over every aspect of life within Bethlehem, matters that went unrecorded in other communities were recorded in order to be reported to church leadership. The collections reflect the breadth and depth of life in Bethlehem during this period. We processed personal papers of artists, tradesmen, missionaries, and sailors, along with business records comprising ledgers and inventories detailing operations of grist and saw mills, tailors, weavers, dye works, soap factories, taverns, tanneries, and lumberyards. We cataloged books from the Bethlehem Congregational Library (est. 1751) that were considered helpful in building a settlement in the New World. Finally, we included in this project about 800 maps and architectural drawings of buildings constructed in Moravian communities, revealing the earliest documentation of European settlement.

In 2013 Lehigh received its second CLIR grant, for the project Bridge and Building Forensics: Civil Engineering Archives at Lehigh University. The twentieth century saw many advancements in civil engineering technology. These collections include the personal and corporate papers of prominent civil engineers and influential societies including Blair Birdsall, John Fisher, Willis Slater,

and the Council on Tall Buildings and Urban Habitat. Both Fisher and Birdsall made significant contributions to bridge engineering and research. Fisher is best known for his work on fatigue and the cracking of steel bridges around the world, and Birdsall was an expert in cable-stayed and suspension bridges. Bridges represented in these collections include the Tappan Zee, Verrazano, Golden Gate, Brooklyn, and the Akashi (in Japan). Slater was a pioneering educator brought to Lehigh in the 1920s to direct the activities of the innovative Fritz Engineering Laboratory. The project also includes a collection of some 200 postcards featuring American bridges. The field of transportation studies has been gaining momentum in recent years, as evidenced by the number of researchers contacting our special collections and others. Once processed, these archives will provide access to correspondence, reports, subject files, court records, images, and engineering data, among other materials.

A significant component of these two projects was Lehigh's dedicated and often innovative use of student staff. Lehigh is committed to creating paid internships for students and providing training that will benefit not only the student, but the project as well.

The Moravian Archives in the New World

Upon notification of receipt of the CLIR grant, Lehigh University and the Moravian Archives sent press releases to the local media; to library, archives, and museum outlets; and to scholarly publications to spread the word about the wealth of resources that would be revealed through the grant. By networking with faculty, students, and lifelong learners, the project staff has promoted collections, student internships, and the potential for scholarly research. Project participants presented various aspects of the project at a wide range of venues: the Mid-Atlantic Regional Archives Conference (October 2011), the Rare Books and Manuscripts Section committee meetings at American Library Association conferences, and the Philobiblon Club (a Philadelphia-based bibliophiles society). With the help of Lehigh University's CLIR postdoctoral fellow, staff contributed to a blog documenting some of their more unusual discoveries and experiences on the project (http://hiddencollectionmoravianarchives.blogspot.com/).

As anticipated, recruiting students with German-language expertise to process archival collections proved to be challenging. Nominations and recommendations for potential candidates from Lehigh University faculty were invaluable. Through correspondence with faculty in a number of academic disciplines and many print and electronic advertisements, we developed a pool of candidates from which we hired a diverse group of students. During the interview, we asked students to translate a passage from a nineteenth-century German text and gave them a tour of the facility. We quickly learned that excellent English writing skills and a reading knowledge of German were required to work efficiently. We found that students who were not native speakers of English translated German texts first into their native languages and then into English, thereby losing too much in translation. Through the same process, we identified student project staff with a general familiarity with foreign languages to work on the book collection. These students were able to identify bibliographic elements on title pages (e.g., author, title, publisher) and search them in the OCLC database.

The first students hired to join the project staff represented a wide range of academic disciplines, including history, civil engineering (with a

minor in German), and political science. Two worked on archival collections, while the other two cataloged printed books. Two graduated during the first year of the project, but we found that the quantity of work they produced in this short period more than offset any investment in training them. As would be typical for most library and archives projects, we had anticipated recruiting students from the English and history departments. But we found that the diversity of this student group easily met project needs. After two students graduated, we hired a supplemental rare book project cataloger to work over the summer. He stayed with the project through the fall and spring semesters.

By the end of the project, we had hired 15 undergraduates and one graduate student. They represented a wide range of academic disciplines, with majors ranging from engineering to the humanities to the social sciences. Two of these undergraduates continued to work on the project as graduate students, and the lone graduate student hired in the first year remained through project completion.

We continued to modify training models established at the beginning of the project to suit needs as new staff were hired. Among the training sessions offered to project staff were German script, printed Fraktur, archival processing, and rare book cataloging. We also presented a Society of American Archivists EAD (Encoded Archival Description) course to all project staff.

Project managers offered specialized training based on the format of the materials. Training project staff to create both original and copy cataloging records was a collaborative effort. Two of Lehigh's catalogers and Lehigh's CLIR fellow coordinated the initial training and creation of the book cataloging portion of the project, while primary responsibility for daily supervision of project staff remained with a retired special collections curator volunteering at the Moravian Archives. Lehigh staff prepared a schedule to ensure that a supervisor was available to students during their first few weeks of employment.

All students cataloging books were trained to search the OCLC database, and were then divided into searchers and catalogers. We found it was more efficient to have one student vet viable records in OCLC, and to have a second student use the preferred records as a foundation on which to create local records in Mandarin Oasis, the catalog used by the Moravian Archives. Students formulated local notes containing provenance information, as well as the original number of each volume in the nineteenth-century Congregational Library catalog. (A search of the Moravian Archives book catalog using the term "conglib" will reveal records for titles cataloged through this Hidden Collections project.)

Paul Peucker, director of the Moravian Archives and project manager of the Hidden Collections project, introduced students working on archival collections to best and local practices. He trained project staff in archival principles and in using a template created to input records into Augias, the German system in use to process archival collections. Peucker reviewed and revised the work of all project staff handling personal papers and business records. The Moravian Archives staff trained newly hired students in reading German script.

Dr. Peucker also trained the student tasked with cataloging maps and architectural drawings in this process, drawing on his experience managing the cataloging of similar collections of historical documents. Special attention was paid to the diversity of measurement units, as

data describing each drawing was entered into a custom-designed template in Augias. Cataloging of the 1,200 drawings was completed early in the second year of the project.

Because of the diversity of the materials and formats associated with this project, permanent staff at both Lehigh University and the Moravian Archives spent more time on the project than initially anticipated. In the second year, Lehigh assigned its CLIR postdoctoral fellow to the Hidden Collections project to help supervise students and revise student work. This added contribution proved beneficial to all.

As planned, project staff worked with Lehigh's library technology team to develop a mechanism for presenting encoded finding aids in CONTENTdm. The Augias system served as an EAD template capable of producing XML documents for import into CONTENTdm, which is presently being used as a digital repository system at Lehigh (http://digital.lib.lehigh.edu/hidden/).

Project staff, including the student processing archivists, scholars, and faculty members identified by the principal investigator and project manager, hosted a visit by the CLIR scholarly engagement survey team. This interaction led to further promotion of the resources the catalogers were bringing to light, as one of the participants, Katherine Faull, professor of German and humanities at Bucknell University, later gave a presentation at Lehigh on her work with early Moravian maps to an audience of faculty, students, and community members.

The Bridge and Building Forensics Project

At time of writing, Lehigh is nearly one year into its second CLIR grant—Bridge and Building Forensics: Civil Engineering Collections at Lehigh University. While we hope to build on the experience gained from the Moravian Archives and other CLIR Hidden Collections projects, this project presents some unique challenges.

As noted above, Lehigh's Special Collections is able to partner with faculty to identify and recruit project staff with particular academic strengths. We received recommendations from faculty in both the history and civil engineering departments, and hired five students. We found that the two in civil engineering had a strong foundation in terminology acquired through the course of their studies.

To ensure that students understood the complexity of the structural technology and the basic concepts and terminology of civil engineering in general, Lehigh's engineering librarian trained the processing staff on Engineering Village (a web-based discovery platform) and on Compendex (an engineering literature database). In the Q&A and show-and-tell training sessions, the librarian demonstrated how to navigate the sites and use the search functions to find the most proper terms.

The processors' hands-on basic processing training covered the "more product, less process" philosophy, archival arrangement, preservation, and description basics. Training in provenance, in understanding personal papers and business records, in archival classification, and in identifying various audiovisual formats occurred as questions arose during processing.

One of the largest collections in the project consists solely of civil engineering testing photographs and negatives. To better understand the context and concepts, the project team took a field trip to Lehigh University's Fritz Engineering Laboratory to observe individual testing setups

and equipment, and ask the testing professionals questions. Visiting the actual scene where the tests took place was indispensable.

Project Staff Perspectives

One of the key elements of these two projects was the use of student project staff. Lehigh University perceived the significant educational opportunity contained within these two grants. By using a student workforce, we enriched the education of multiple students, at the same time introducing fresh faces to the world of archives and special collections. This benefited the institution and the students. With such an emphasis on student work, it is only fitting that we include student perspectives. Below are essays by Andrew Stahlhut, a doctoral candidate at Lehigh who was a project cataloger for the Moravian Archives project, and by Gregory Edwards, a recent Lehigh history graduate who was a project archivist for the civil engineering project.

▸ The Moravian Archives: Student Impressions

Andrew Stahlhut, doctoral candidate, Lehigh University; project cataloger, Moravian Archives

My perceptions of the project are shaped by my identity as a doctoral candidate in colonial American history and my personal interest in historical book and print culture. My motivations to participate stemmed from professional curiosity about what kinds of books a sect of eighteenth-century pietists in colonial America kept in their congregational library, and a personal desire to work with and organize old books. I'm very much a bibliophile in my private life and the opportunity to help organize and catalog a library collection like this for the first time was something I couldn't pass up. I cataloged the roughly 1,700 books accumulated in the Moravian Congregational Library from the time the Moravian community settled Bethlehem, Pennsylvania, in the early 1740s.

My project with the congregational library was only one of several under the larger project. I find it interesting that all of the students involved with the Hidden Collections project brought their personal and professional experiences and interests to their work. We were not just library science or archival students performing simple sorting tasks. We accomplished what we did because of our diverse skills and disciplines. For example, in sorting and cataloging architectural diagrams and blueprints, an undergraduate engineering major worked more efficiently and with more appreciation for the task than a humanities student such as me might have. Another graduate student with German fluency and a background in political science and international studies read, sorted, and cataloged eighteenth-century manuscripts written by Moravians in Bethlehem. Her fluency in German made this task possible and she performed a task that someone such as myself, without knowledge of the German language, simply could not do.

The project presented both obstacles and valuable lessons. The biggest obstacle was that I do not speak or read German. For the books I was cataloging, all I could do was match words from the title pages to works in online databases because they looked the same. It was more like comparing symbols than words, because of the German Fraktur printing that graced so many of the title pages. Often I had to resort to other ways of identifying a book such as by the imprint or pagination. This initial obstacle led to new ways of problem solving, so I suppose it counts as a lesson learned and an opportunity to hone a new skill.

The work led to more lessons and an appreciation of the nuances of the congregational library. I was awed at how diverse a library the

mid-eighteenth century German colonists had. Admittedly, most of the library comprised religious texts. However, it also contained a number of books on mathematics, which included works that explained numerical and geometric theories behind shapes and simple machines and applied them to everything from pulleys to large military fortifications. Another book was purely academic, teaching mathematic proofs. The library also contained a large, hand-numbered atlas composed of dozens of printed, hand-colored maps of Europe, Africa, and Asia. The collection even held books on unexpected religious topics such as vampires. Although the Moravians settled in America's proverbial wilderness, they maintained links to the larger world of knowledge.

It is this last point that especially makes the project valuable to researchers. Of course my work has aided historians of the Moravian Church, or religion generally, by helping make the existence of these works known so people can visit and use them. However, the collection's diversity of topics means that researchers seeking to access books on other topics, such as mathematics or navigation, could visit the Archives and access those works as well. The diversity of the collection thus diversifies the archives' role beyond that of a repository for texts related to the Moravian Church and perhaps more accurately reflects the links the eighteenth-century Moravians had to the larger Atlantic world around them.

My participation in the project has fundamentally changed how I perceive the behind-the-scenes activity of libraries and archives. Like most academics, I used to take catalogs for granted. Now I realize just how much work goes into an accurate catalog record. For any specific book I had to check the title and subtitle very carefully. I checked all information in the imprint, the pagination, and to make sure our catalog records represented our unique, specific copy of the book I made notations for inscriptions, signatures, marginalia, or any other unique markings. In one case I found an old pressed flower in an eighteenth-century book. Sometimes, several smaller books would be bound together and I'd have to make sure each was cataloged accurately but keeping the same base call number for all of them. It was a lot of work, and I loved every minute of it—and I now appreciate archivists and catalogers that much more.

In sum, the project was a great experience in my life and I wouldn't have traded it for anything. I call it an experience, and not a job, because it was more than a job. It was something I loved to do. I often joked that it was something I would have done for free. Being trusted with access to the two vaults of books was an important milestone in my bibliophilic life. Being on the other side of the line between the public and archives insiders was exciting. I will now pursue my academic career with a greater respect for the amount of work that happens on the other side of that line.

▶ Bridge and Building Forensics: A Student's Perspective

Gregory Edwards, Lehigh University graduate, history/public history; project archivist, civil engineering project

The Council on Tall Buildings and Urban Habitat (CTBUH) collection, the largest collection in the grant, comprises a significant amount of both visual media and paper materials, and presents many unique challenges in processing. Foremost among such hurdles is original order, and disorganization. Created by Lynn S. Beedle, one of the founders of CTBUH and later head of Fritz Laboratory at Lehigh University, this collection, as found, uses a custom numerical indexing system, contains much duplication, and separates related materials. An index does exist, but is incomplete.

These issues are prominent in the roughly 7,500 slides in the visual portion of the collection, which are accompanied by related notebooks, prints, and cases of thousands of duplicates. To complicate things, the duplicate cases occasionally contain slides absent from either the binders or the index, while the index lacks entries for many slides and contains entries for many slides that are missing. What results is a complicated situation in which the processor must take painstaking measures to attempt to maintain the original order of the slides while reconciling the differences between the organization of the materials and the index while searching item by item through thousands of duplicates to check for interspersed originals.

These organizational difficulties carry over into the Blair Birdsall Papers, a collection containing materials created by Birdsall, Steinman, and other engineers involved with the Roebling firm and, later, the Steinman firm. The papers are a conglomeration of material from different authors, many of whom used separate and distinct indexing systems, so organization is difficult. It is hard to determine the author of many papers, and the indexes make little sense. Attentiveness to physical elements such as folder color and paper type proved useful in reorganizing the papers in a way that recaptured the original order.

A challenge with the John Fisher Papers is that John Fisher, still employed at Lehigh University, remains actively engaged in using his papers. He has agreed to keep them at Lehigh, but they are being transferred in sections for processing. Benefits and difficulties come from such a situation. Fisher is available for limited contact and consultation during processing, but his use of the papers also creates a break in the context of the collection. It is difficult to create series and subseries when not all of the material is immediately available. Chronologically, the materials are largely in order, allowing for logical series in that manner, but organization based on topic and format suffers.

One collections-wide challenge that especially applies to the Fritz Laboratory Negatives and Photographs is the specificity of content. These four collections contain items exploring specific and detailed fields of civil engineering that can be very difficult for a non-engineer processor to understand. For example, the Fritz Lab negatives contain many close-up images of step-by-step testing that are nearly impossible to identify without proper context. This difficulty translates to many aspects of processing, making it challenging to accurately select subject headings or group material into series. The best solution was a trip to Fritz Lab, on campus, where we explored testing apparatus. With the help of lab staff and faculty, we were able to identify individual testing setups, terminology, and locations.

The personnel acquired for this grant proved to be an interesting variable. Considering this grant is composed of six civil engineering collections, in an ideal situation processing staff would simply be plucked from the readily available supply of engineers that Lehigh University has to offer. Unfortunately, it seems that the engineering students were not of the same mind. Thus far, only one civil engineer has joined the project staff. Her background has proved very valuable in her work on the John Fisher collection, allowing her to more quickly understand and identify key concepts. Enlisting campus resources such as Sharon Siegler, the senior engineering librarian, proved effective in coping with the lack of staff with engineering experience. Sharon trained us in navigating and using engineering databases, which helped us identify appropriate vocabulary for collections description. Importantly, patience and attention to detail were immensely helpful in processing these civil engineering collections, and we will continue to learn more as the project continues.

The "Deceased" Preaches His Own Eulogy: Training Students to Provide Access Points on Discovery-Level Records

Felicia Piscitelli, Associate Professor, Special Collections Cataloger and Italian Resources Librarian, Texas A&M University Libraries

Lisa Furubotten, Associate Professor, Serials Cataloging. Texas A&M University Libraries

Anton duPlessis, Clinical Assistant Professor, Curator of the Mexican Colonial Collection; Director, Primeros Libros de las Américas, Texas A&M University Libraries

Alma Beatriz Rivera-Aguilera, Coordinadora de Automatización, Biblioteca Francisco Xavier Clavigero, Universidad Iberoamericana, Mexico

Ángel Villalba-Roldán, Jefe del Departamento do Catalogación, UNAM-IIB-Hemeroteca Nacional de México

Abstract

This paper is an outgrowth of Discovering a New World: Cataloging Old and Rare Imprints from Colonial and Early Independent Mexico, a project funded by the Council on Library and Information Resources (CLIR). The grant developed a methodology by which Hispanophone student workers, with minimal training, can create discovery-level records for collections by inputting bibliographic metadata into a web-based template.

In addition to descriptive elements, access points for corporate and personal names and some type of subject access are needed to truly make this collection discoverable by users. We asked: Can students with rudimentary training provide such access points? If so, are they limited to inputting uncontrolled terms, or can they be trained to use controlled vocabularies? We experimented with four input methods to evaluate which ones would effectively enable students to populate the access point elements:

- **Uncontrolled vocabulary:** for example, name headings entered in a [surname, forename] format, keywords, and summaries.
- **Controlled vocabulary lists embedded in the metadata collection tool:** for example, genre terms taken from the Rare Books and Manuscripts Section (RBMS) Controlled Vocabularies, and geographic subject headings taken from Library of Congress Subject Headings (LCSH).
- **Dynamic feedback list:** for example, an embedded list of controlled name headings for printers from the National Authority File (NAF), with a feedback mechanism permitting a cataloger update the list as needed.
- **Use of authorized thesaurus:** for example, the Virtual International Authority File (VIAF), from which students can select authorized name headings.

In this paper, we discuss the pros and cons of each method. We also note that while providing lists of controlled headings might be useful for small or homogeneous collections, this method proves unwieldy for a large heterogeneous collection, and it is worth training data inputters to use authority files.

The Mexican Colonial Collection at Cushing Memorial Library and Archives

The Mexican Colonial Collection at the Texas A&M University Cushing Memorial Library and Archives comprises more than 4,000 books, manuscripts, broadsides, newspapers, pamphlets, and religious works portraying practically all aspects of life in New Spain and Mexico from the sixteenth century to the second *Imperio Mexicano* (1867). This collection encompasses the fields of history, anthropology, linguistics, religion, political science, philosophy, civil and canon law, military science, the history of the book, the sciences, and arts. There are legal documents including items circulated by the *Santa Inquisición,* theological and religious works (many of which contain beautiful woodcuts and engravings), newspapers, discourses concerning Mexico's independence movement, and even cookbooks. Most items are in Spanish, but some pieces are in Latin, Nahuatl and other indigenous languages, French, English, or Italian. Many are unique as Texas A&M is the only U.S. institution owning copies.

Project Overview

Several hundred bound volumes containing multiple works, and more than 600 single-sheet items required cataloging to make them discoverable. In a 2012-2014 CLIR project, the investigators—three from Texas A&M and two from Mexican libraries—developed a methodology whereby Hispanophone students input bibliographic metadata into an easy-entry web-based template developed in SharePoint.[1] We hired 13 detail-oriented students with robust Spanish language skills—all the students were native or second generation Hispanophones; half were Latin American international students or permanent residents.

The challenge was to identify strategies that would enable students to produce records similar to those created by professional catalogers, and which would integrate well with records already in the bibliographic catalog.

We divided the work into stages to determine how students would handle increasing degrees of difficulty, and to calculate the cumulative cost of generating more complex records. We prepared an instruction manual with illustrations and examples and conducted one-on-one student training sessions.

Basic Descriptive Elements		Additional Elements
Title	Author	Illustrations
Variant Title	Author 2	Illustration note
Subtitle	Author: Corporate Body	Ornaments
Variant Subtitle	Place: Country	Font Color
Statement of Responsibility	Place: City	Printer's Device
Publication: Place	Printer: Authorized Form	Binding
Printer	Printer: Corporate Body	Marca de Fuego
Date	Date: Date Form	Handwriting
Language		Bookplate/Ex Libris
Pages		Bookstamp

Fig. 1: Sample elements

These stages were as follows:

- **Basic descriptive elements.** The students were asked to input into the template values for the simple descriptive elements shown in Figure 1.
- **Keywords.** Students tagged the items, using natural language terms they thought appropriate.
- **Additional descriptive elements for rare books.** In this phase, additional descriptive element fields particular to the features of rare books were added to the template.

1 SharePoint is a web application framework and platform developed by Microsoft that facilitates collaborative projects. https://products.office.com/en-us/SharePoint/collaboration

- **Single-sheet items.** Because these formats require some unique elements, we designed a second template for the single-sheet and broadside items.
- **Authority file use.** Students were instructed in the use of Virtual International Authority File (VIAF) to test whether they could effectively select appropriate name headings for the item at hand.

Helping Students Populate the Record Elements

To enhance discoverability, besides descriptive elements, we wished to supply access points for persons and corporate bodies associated with an item, and some sort of subject access. We wondered if it were possible for students with only rudimentary training to provide such access. We assumed that our students could furnish uncontrolled access such as keywords, free-form notes (including summaries), corporate names in natural language order, and personal names in a simple [surname, forename(s)] format. We also inferred that the students' knowledge of Spanish and of naming practices in the Hispanophone world would help them correctly formulate names and deduce the substance of the items. Another question then arose: can students be trained to deploy controlled vocabularies such as the RBMS genre term list or Library of Congress Subject Headings, or copy headings from an authority file such as VIAF?

Because the records were destined for the bibliographic catalog, we wanted them to conform as much as possible to established bibliographic norms. At the same time, traditional cataloging requires mastering many complex standards such as rules for description, headings, and subject strings. This paper focuses on developing approaches that enable minimally trained students to populate bibliographic fields. We enjoyed the challenge of determining if we could enable students to produce records that would approximate those produced by trained catalogers, experimenting with the following strategies.

1. Uncontrolled Values

Some elements were satisfied with unrestricted uncontrolled values that students judged appropriate. Such values cannot later be indexed and are only useful for keyword searching of the records. These values provide imprecise inconsistent retrieval, faceting, and collation.

- **Subject keywords.** Students were instructed to consider terms *they* would use if searching for the item at hand, and to input those into an open-ended keyword field. They reported being comfortable with this task, and that information found on the item preliminaries[2] often adequately indicated the item's nature and contents. In other cases, they demonstrated resourceful ingenuity in searching the Internet for information about the item itself or the entities named on the item, to glean sufficient historical context to choose appropriate keywords.
- **Summaries.** For the proclamations and broadsides, students input summaries describing the nature of the item. Occasionally summaries could be copied verbatim from a vendor-supplied description.
- **Uncontrolled titles.** Students were asked to input the titles as found on the item. They were not required to identify the original title if the item was a translation or reprint; therefore, the resulting records will lack any required uniform titles and linking notes for

2 Title page, cover, colophon, etc.

the original publication, unless subsequently supplied by a cataloger. We considered these elements too difficult to require from students.

- **Uncontrolled author names.** In the first stage of record construction, students were asked to input the author's name in the Author field in [surname, forename(s)] format. Elsewhere in the record (e.g., the statement of responsibility, summaries), the students were to record names as found on the item. We thought this would be an easy task, but it was found to be difficult, as discussed below.

2. Controlled Values from Drop-down Lists

One of the reasons SharePoint was selected as a tool for hosting our record database is that it permits us to embed drop-down lists of preselected values in the fields. We used terminology from standard controlled vocabulary lists; in the case of a name, we used the authorized form found in the authority file. Database administrators may also implement rules controlling whether a selection is required, or if multiple values are permitted.

- **Static lists.** Lists of controlled values were embedded in these element fields: country, language, illustrations, ornaments, font color, binding types, author role, and title source. Students were given instructions and definitions to enable them to select appropriate values.
- **Static lists: genre.** A list of collection-appropriate genre terms selected from the RBMS Controlled Vocabularies was embedded in the genre element field.
- **Static lists: LCSH subjects.** Broad subject headings selected from the *Library of Congress Subject Headings* were used in the subject element field.

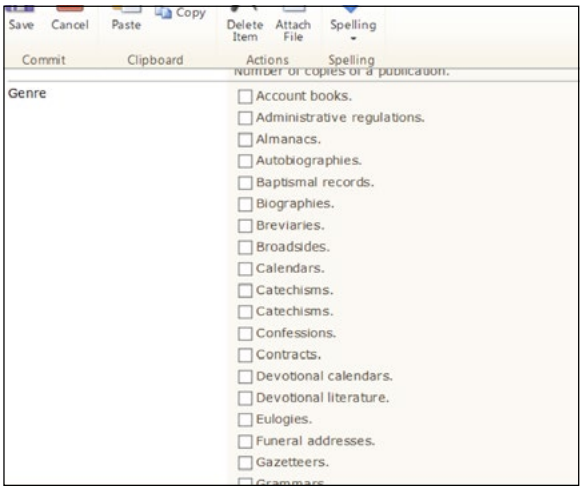


Fig. 2: Genre element with embedded list

Fig. 3: Subject element with embedded list

Controlling the possible values ensures compatibility with values used in existing cataloger-produced records. In turn, this promotes indexing and faceting when the new records are integrated into the catalog. The controlled vocabulary lists also reduced typographical errors.

Drop-down lists are most useful when working with smaller homogeneous collections with a limited number of applicable terms or headings. Conversely, a broad collection like the Mexican Colonial Collection requires more terms, yet this creates a problem—a larger list of specific terms that becomes unwieldy when embedded in a template field, and which requires more extensive student training. The only way to revert the list to manageable size is to use broader terms.

Furthermore, if the cataloger must examine a large number of items in a heterogeneous collection to develop a large list of precise terms, this is so labor intensive that he/she might as well assign the subject headings in the traditional manner.

To overcome these problems we attempted to modify this strategy and experimented with making lists dynamic, expanding them as needed through feedback provided by the students.

- **Dynamic lists. Example—City of publication:** For the city of publication element, we developed a list of city names in the form authorized by the Name Authority Cooperative Program (NACO) authority file, with cities likely to appear on the items in the collection. But if the students encountered a city of publication not on the embedded list, they could select the *Specify your own value* option and input the new city name as found on the piece. A cataloger would periodically review the database to find these new city names, look up the authorized form of the city, add the authorized form to the list as a value likely to occur on items in this collection, and correct that record by removing the *Specify your own value* choice, replacing that choice with the authorized form of the city now on the list.
- **Dynamic lists. Example—Authorized name of printers:** Similarly, we developed a list of authorized names of printers likely to be associated with this collection. Using the same process described above, new printer names could be added to the dynamic list as needed.

Using feedback generated by the student's examination of the items permitted us to provide larger numbers of controlled terms. Still, this strategy is more useful when working with a more homogeneous collection such as the *Primeros Libros,* which is limited to sixteenth-century imprints and thus associated with a limited number of printers. For the larger Colonial collection, because the number of printers grew exponentially throughout the eighteenth and nineteenth centuries, it became increasingly difficult to modify the list in a timely manner.

Fig. 4: Drop-down list for Publication: City field

3. Student Use of Online Authority File

To avoid the impracticality and labor of embedding an extensive list in fields, we attempted to test whether students could learn to effectively search name authority files and discover for themselves the authorized form of a corporate or personal name.

For a subset of records that did not already have name headings provided by one of the previous strategies, the students searched the VIAF. They copied the 1XX field containing the authorized form of the name into the SharePoint template fields. This strategy required investing in training students to navigate the authority file and choose the correct record corresponding to the needed name. There are benefits to such training. However, students sometimes had difficulty identifying when they had the correct authority record corresponding to a name found on the item.

What Catalogers Know About Spanish Name Headings

Making access points for personal names can be daunting, even to people familiar with the language of the items being cataloged. Catalogers have the National Authority File and the rules set forth in Resource Description and Access (RDA) and the Anglo-American Cataloging Rules (AACR2) to guide them, but non-catalogers would be unaware of these resources and would require complicated training to use them. This project involved Spanish-language materials. Spanish-speaking students were chosen with the hope that their knowledge of Spanish would permit them, with only basic instruction, to formulate names correctly. That did not always happen.

Some Spanish names consist solely of forename(s)-surname. As is the case with names in English, a basic [surname, forename(s)] heading would suffice, as long as the design of a project does not require that each name is unique.

But most Spanish names are not so simple. Most people in the Hispanophone world use the surnames of both parents. The father's surname comes first and, thus, serves as the beginning of the access point. For example, in the case of Juan Antonio Pérez Vásquez, whose father's family was Pérez and whose mother's was Vásquez, his heading would be Pérez Vásquez, Juan Antonio. Prepositions such as *de*, articles (*el, la, los, las*), the two in combination (*del, de la, de las, de los*), or conjunctives such as *y* may occur as part of a name. Often when a woman marries, she appends her husband's first surname to her family's surname(s), with *de*. Thus, if Isabel María Ortíz y Pino marries Juan Antonio Pérez Vásquez, her full name would be Isabel María Ortíz y Pino de Pérez. Her surname begins with Ortíz.

According to the rules for establishing Spanish names in both AACR2 and RDA, the heading should begin with the first element of the surname, excluding prepositions like *de*.

Before the seventeenth century, some people simply appended the name of a place or an attribute to their forenames. An example of this is Pedro de Alcántara, a sixteenth-century Spanish saint.

It is customary among Catholics to name their children after saints, so sometimes people are named with phrase-like names such as Pedro de Alcántara Martínez. It is easy to think that "de Alcántara" is part of this man's surname, but here it is part of his first name. He is named after the sixteenth-century saint. The correct input for his name is *Martínez, Pedro de Alcántara.* Inputting such names requires some expertise.

Variations in Spanish form or spelling often occur during the time period covered in this collection, for instance, Yturriaga vs. Iturriaga, Leyva vs. Leyba, Florez vs. Flores, Gerónimo or Gerónymo or Jerónimo, Xavier vs. Javier, and so on. The popular first name José was usually rendered Joseph or Josep as late as 1800. Another potential difficulty is the versions of a name appearing in other languages. Pedro Martínez wrote in Latin, using the form Petrus instead of Pedro on the items. "San Juan Nepomuceno" is the Hispanicized form of a Bohemian-born saint, John of Nepomuk.

Titles of nobility add yet another layer of complexity, particularly if a person tended to use his or her title rather than their name on documents. The Colonial Mexican collection contains a document by General Francisco Xavier Venegas addressed to "Excmo. Sr. Duque del Infantado." The Duke's name does not appear in or on the piece, and the student entering the metadata only gave Venegas' name. It required research to reveal that the Duke's full name was Pedro de Alcántara Álvarez de Toledo y Salm-Salm.

Other categories of names, such as those represented by initials only, and pseudonyms, present problems in any language. For example, one would need to look in the authority file to know that Ignacio Tomay (also spelled Tomai or Thomay) was the pseudonym of José María Genovese, a Sicilian-born Jesuit living in New Spain.

Discussion

What Was Easy?

Students found the title element easy, because the students recorded information as it came on the item in hand. They also had no trouble with elements for which we were able to provide drop-down lists. For the investigators, the dynamic lists were easy to manage until the amount of feedback became overwhelming.

What Was Difficult?

Occasionally students had difficulty identifying the author(s). When several names were mentioned on an item, students were sometimes misled by the wording into selecting the wrong individual as the author. For example, in Spanish the word *por* can mean *for* or *by*, depending on the context. Usually the name of an author is introduced by *por*, as in *Don Quijote de la Mancha* por [by] Miguel de Cervantes. However, in funeral sermons, where the title statement appears in the form *Oraciones funerales por* [name of the deceased], students automatically assumed the individual named after the word *por* was the author. The word *del* can also be misunderstood in the same way, as in *"Solemnes exequias del illmo. señor dr. d. Jose Gregorio Alonso de Hortigosa, obispo que fue de la ciudad de Antequera …"* for which a student entered Hortigosa in the record as the author, when he is actually the deceased. This mistake happened often enough to provoke the humor in our paper's title.

Other errors occurred when students formulated name headings without help from the VIAF file. The following types of mistakes appeared frequently.

Type of Error	Entered	Correct
Incorrectly selecting an element as the first element of the surname	Antonio de Oviedo, Juan Jose De Urdanivia, Cárlos Joseph Diez	Oviedo, Juan José Antonio de Díaz de Urdanivia, Carlos Jose
Incorrectly including abbreviations of titles as part of the name	De Castro, D. Juan Francisco Br. D. Juan Gabriel de Contréras	Castro, Juan Francisco de Contréras, Juan Gabriel de
Beginning a name heading with *de*	de Contreras, Juan Martin de Jesus Soria, Rafael	Contreras, Juan Martín de Soria, Rafael de Jesús
Incorrect use or lack of diacritics	Perez Quixano, Agustin	Pérez Quixano, Agustín
Not recognizing when a person's forenames include the name of a saint	Nepomuceno Rivas, Juan	Rivas, Juan Nepomuceno (*Nepomuceno* is really part of the man's forename, as he was named after St. John of Nepomuk)
Not recognizing non-Spanish forms of names; i.e., names in Latin, Italian, or French, or even archaic spellings of Spanish	José Gregorio Alonso de Hortigosa Joseph Gregorii Alfonsi de Ortigosa (Latinized form of the name)	Alonso de Ortigosa, José Gregorio (heading from Biblioteca Nacional de España)
Not recognizing when an individual had only a forename (despite training)	Transfiguracion, Francisco de la Sales, Francisco de	Francisco de la Transfiguración Francis de Sales

Table 1: Most common errors, with examples and corrected forms.

What Was Impossible?

Three challenging areas were typically beyond students' abilities.

- **Works by corporate officials are entered under their official titles rather than their personal names.** This is not something we would expect a non-cataloger to know. For example, a document issued under the authority of the viceroy Martín de Mayorga should have as the author the form *New Spain. Viceroy (1779-1783: Mayorga),* not *Mayorga, Martín de.* We accepted that records for official documents would incorrectly list personal name forms in the element fields.
- **Translations.** Without substantial training and extensive searching in OCLC, it would be difficult for students to determine when a publication was a translation or what its original title would be. We accepted that these records would not have translation uniform titles and would not collate with, or have linking notes to, the original editions.
- **Pseudonyms.** Although sometimes—with the help of the VIAF files or through an Internet search—the students were able to find the real name of an author appearing on the item with a pseudonym or initialism, this was a difficult task we did not require.

Conclusion

The purpose of this paper is not to argue for the strict use of controlled vocabularies and headings in bibliographic descriptions, either in traditional catalogs or in other databases using non-MARC based records. We understand that it is acceptable for record standards to vary depending on the type of access needed for a particular collection or project. And it is clear that there is a lack of trained personnel to create records at the highest level for the backlog of items that need to be processed. It was not our intention to demonstrate the value of the work of the Library of Congress, OCLC (via VIAF), Dublin Core, ALA, and others. in promoting the sharing of authority records and of controlled vocabularies outside the traditional cataloging community. Yet, unexpectedly as a related issue, this project has reiterated to us the importance of these organizations' efforts as they facilitate the participation in projects (especially digital) of non-catalogers.

Such efforts allow project managers to shift focus away from the difficult, expensive, and time-consuming task of training non-catalogers (usually temporary personnel) to construct headings or use vocabularies according to standards. Instead, we can focus on the lesser task of training staff to navigate the authority record databases and select the correct record showing the authorized form of the topical or name heading needed. Users of the authority file records can see the correct authorized form of a heading, including diacritics, order of the elements of a name, correct spelling, and so on, so they do not need to know how to construct a heading. Professional catalogers—trained in the rules and standards for constructing headings and vocabularies—can dedicate themselves to building the authority records.

Using the online authority and vocabulary resources frees us from training non-catalogers how to construct a heading, but we still need to train them to identify the appropriate record. Experienced catalogers can identify non-obvious bits of information in authority records (including in the citations), which help them determine that a record corresponds to the individual named on the item at hand. For example, one student selected an individual as an author for an eighteenth-century item, while a cataloger would have known that, since the physics text cited

in the record's 670 *Source Data Found* field was published in the twentieth century, the individual represented by the authority record could not be the same individual as author on the older religious-themed item in hand. Non-catalogers, who lack the training and experience, cannot see these clues. The new RDA attributes[3] permit a non-cataloger to better identify which record/individual is the correct one, allowing staff to focus on inputting the proper information rather than debating among several possible records. As another unexpected outcome, this project demonstrated to us the merit of RDA authority record attributes in helping non-catalogers select correct individuals/corporate bodies and the authorized form of name for use in name and subject access.

References

Joint Steering Committee for Revision of AACR2. 2002. *Anglo-American Cataloguing Rules,* 2nd ed., 2002 rev. Chicago: American Library Association.

Joint Steering Committee for Development of RDA. 2010. *RDA: Resource Description & Access*. Chicago: American Library Association.

Name Authority Cooperative Program (NACO): http://www.loc.gov/aba/pcc/naco/.

Virtual International Authority File (VIAF): http://viaf.org/.

3 Associated Dates, Associated Place, Address, Field of Activity, Associated Group, Occupation, Gender, Family Information, Associated Language, Fuller Form of Personal Name.

Additional Reading

Almuzara, Leticia Barrionuevo, M. Luisa Alvite Díez, and Blanca Rodrígue Bravo. 2012. A Study of Authority Control in Spanish University Repositories. *Knowledge Organization* 39 (2): 95-103.

DeZelar-Tiedman, Christine. 2011. Exploring User-Contributed Metadata's Potential to Enhance Access to Literary Works: Social Tagging in Academic Library Catalogs. *Library Resources & Technical Services* 55 (4): 221-233.

Lu, Caimei, Jung-ran Park, and Xiaohua Hu. 2010. User Tags Versus Expert-Assigned Subject Terms: A Comparison of LibraryThing Tags and Library of Congress Subject Headings. *Journal of Information Science* 36 (6): 763-779.

Maron, Deborah, et al. 2014. Lo-Fi to Hi-Fi: A New Way of Conceptualizing Metadata in Underserved Areas with eGranary Digital Library. *Proceedings of International Conference on Dublin Core and Metadata Applications,* Austin, Texas, October 8-11, 2014, 37-42. Available at dcevents.dublincore.org/IntConf/dc-2014/paper/download/259/228.

Yusuf, Felicia, and Christopher Nkiko. 2010. Involvement of Non-Professionals in Cataloguing: A Survey of Practices in Three Academic Libraries in Nigeria. *Malaysian Journal of Library & Information Science* 15 (1): 83-90.

Cataloging




Obstacles and Solutions in Establishing Cataloging Standards for Fine Print Collections

Katharine Malcolm, Project Coordinator, Georgetown University Library
Christen Runge, Art Collection Assistant Curator, Georgetown University Library

Abstract

In 2011, Georgetown University was awarded a Cataloging Hidden Special Collections and Archives grant from the Council on Library and Information Resources (CLIR) to research and catalog the collections of six significant twentieth-century American printmakers. The artists include renowned wood engraver and book illustrator Lynd Ward (1905–1985), the preeminent printmaker John DePol (1913–2004), and four highly talented, but underappreciated, women printmakers of the same period: Louise Miller Boyer (1890–1976), Helen King Boyer (1919–2012), Marguerite Kumm (1902–1992), and Kathleen Spagnolo (b. 1919). More than 5,400 objects were cataloged for the Undiscovered Printmakers: Hidden Treasures in Georgetown University's Library project, which was successfully completed in 2014.

This paper highlights some of the complex issues involved with establishing standards for cataloging the unique physical nature and characteristics of fine prints. There is not always a clearly defined, logical, or consistent set of rules for recording object information relating to these artworks. The process of creating fine prints is very personal and experimental in nature, producing one-of-a-kind objects that exhibit extremely subtle differences between each printed impression.

Customizing the Cataloging Workspace

The Undiscovered Printmakers project has been a case study for establishing policies and practices for cataloging fine print collections at Georgetown University Library, where it served as a pilot project in the implementation of a new-to-us collections management database, EmbARK. Curators and project staff customized many aspects of EmbARK from setup to interface, anticipating that future catalogers are likely to range from experienced art curator to student volunteer and thus will have varying levels of art knowledge. The database has proven to be an excellent and flexible tool, enabling staff to catalog objects confidently in a coherent and consistent manner that facilitates access and scholarly research.

The EmbARK collections management system is

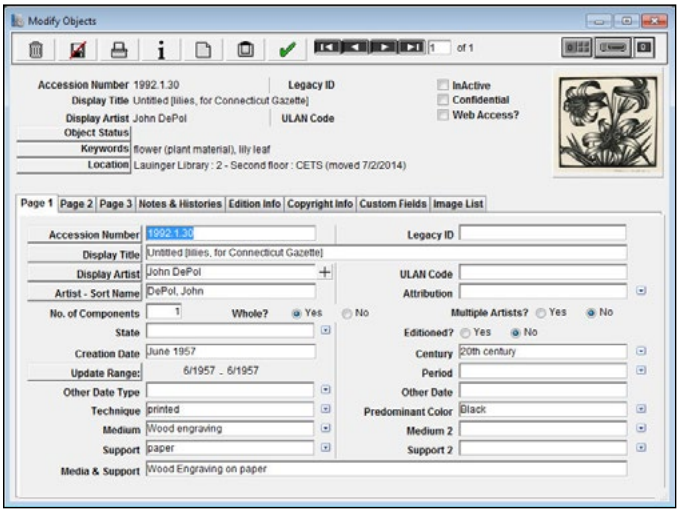


Fig. 1: EmbARK default cataloging view

designed specifically for the management of visual collections and has a vast variety of tools for recording art object information: 17 interrelated tables record everything from multiple sets of measurements to publications that mention the artwork. Many of these tables do not apply directly to the cataloging process, but the crucial objects table alone contains 145 unique fields spread across 8 tabs (Figure 1).

To facilitate the cataloging process, curators developed a customized cataloging view that displays a streamlined selection of only the essential fields for catalogers (Figure 2). This alteration reduces the eight tabs to just one screen and pares down the number of fields by over 60 percent. As a result, the cataloging process is more efficient, and it is easier to ensure that important fields are not skipped or ignored.

An important aim of the project was to introduce and develop a standardized vocabulary of specific printmaking terms within EmbARK. Drop-down lists were created for many fields, offering catalogers a select list of terms and descriptions to choose from when describing the characteristics of a print (Figure 3). The drop-down lists virtually eliminate the chance for human error involved in free text entry and greatly reduce the amount of time curators spend in checking catalogers' work. They also guarantee that records can be searched more confidently, without the worry that misspelled words may cause objects not to be found.

During the course of the project, the curators and project coordinator edited the lists, as the scope and breadth of material became evident. These lists were retained in

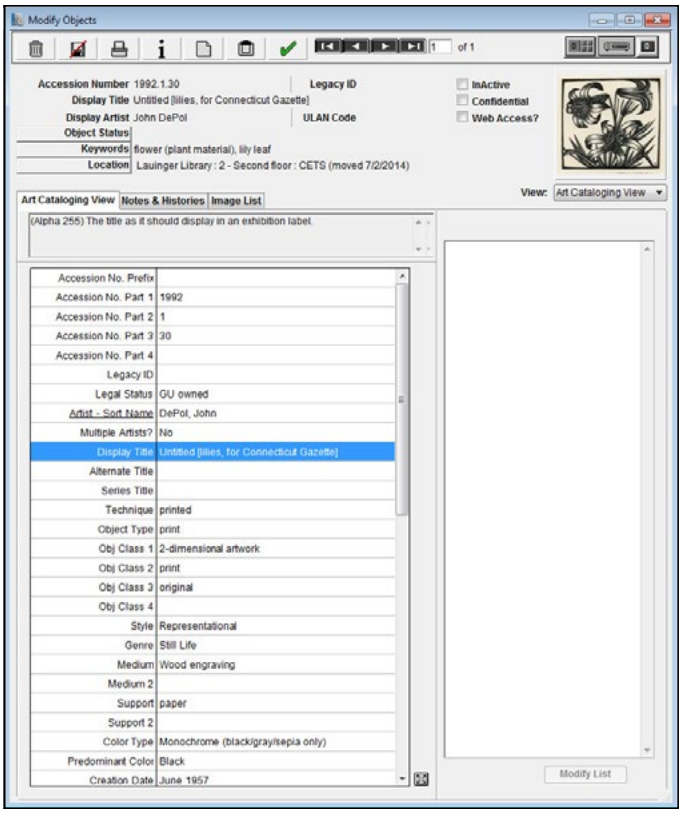


Fig. 2: Customized cataloging view

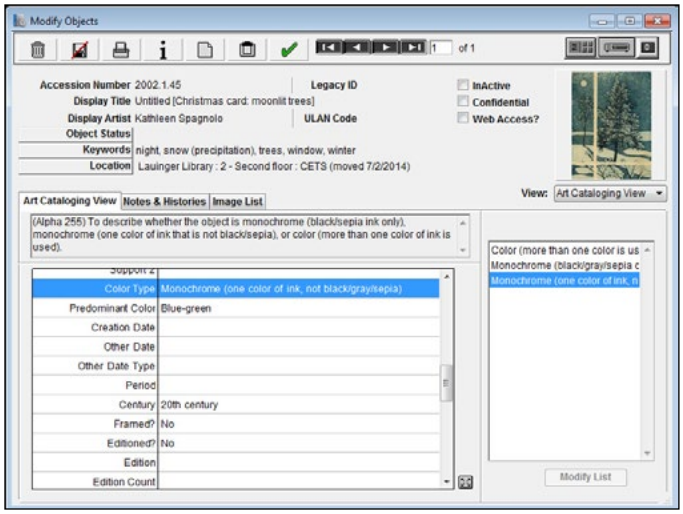


Fig. 3: Drop-down lists that eliminate typos and spelling errors from many fields

EmbARK and are now being augmented by the curators to accommodate the many other types of objects in the broader collection. The original drop-down lists form the core of the vocabulary, however, as the bulk of the collection consists of fine prints.

The following selections from this diverse group of objects illustrate our own cataloging processes as they developed. The questions they presented required research and discussion to resolve how to structure and format particular object information in the customized EmbARK template to best suit project needs and resources. They demonstrate how and why we determined and established particular cataloging policies during the project.

Describing Complex Mediums

Georgetown University Library holds a uniquely personal collection of 70 artworks by Kathleen Spagnolo. She initially worked as an illustrator and eventually specialized in etching, using a special process called multilevel viscosity printing (Figure 4). This method of printing separate layers of colored inks was developed by Stanley Hayter, Krishna Reddy, and others in Paris from the 1920s onward. Spagnolo studied etching at American University in Washington, D.C., with artist-teacher Robert Gates and color etching with Krishna Reddy of Atelier 17, Paris, when he came to the university to teach viscosity printing in 1964. Spagnolo used this printmaking technique to produce her most innovative and outstanding work. By studying a number of the artist's working proofs in the collection, we discovered Spagnolo's experimental use of vibrant color using the viscosity printing technique. She created rich textures through which her bold, three-dimensional forms truly came alive.

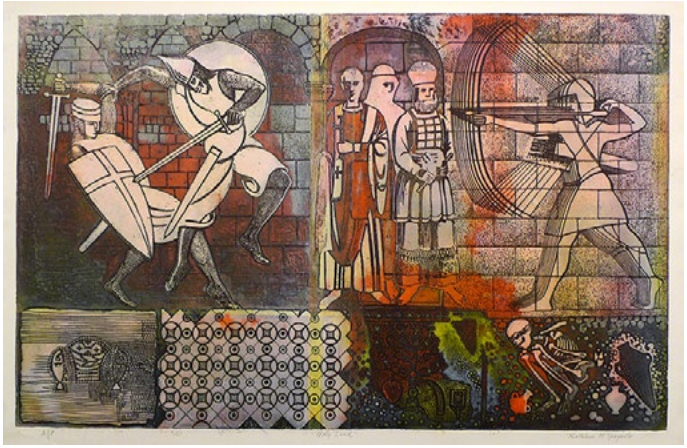


Fig. 4: Kathleen Spagnolo, Holy Land *(1967)*

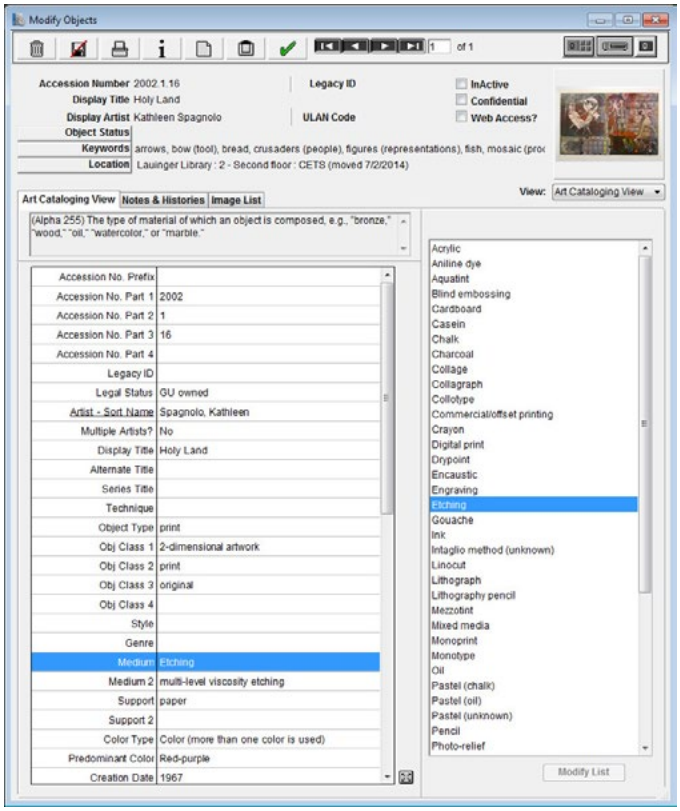


Fig. 5: Object record for Holy Land *showing choice of medium as etching*

When we started cataloging these prints, we wanted to include as much information as possible about the intricate printing technique involved. Drop-down lists proved to be an excellent means of clearly describing the multiple techniques involved. Looking at the object record for Spagnolo's *Holy Land* (Figure 5), we see how the catalogers first selected "etching" from the drop-down list in the Medium field. The choice of etching for the medium determined the choices that then became available in the Medium 2 field (Figure 6), which in our example is a specific type. For other multimedia objects, the Medium 2 field can describe a second, additional medium used (e.g., an etching "with pencil") rather than a narrower description of one medium.

Describing Multifaceted Objects

Lynd Ward is the most renowned and acclaimed artist of the six artists studied in this project. He was a pioneer of the graphic novel and a well-known illustrator of many adult and children's books, and the Ward collection at Georgetown University Library has many key works from his career, including 203 finished original illustrations in ink, crayon, gouache, watercolor, pastel, and scratchboard; 767 sketches; 245 original illustrations for book mock-ups (or artist "dummies"); and 481 prints. Additionally, there are wood blocks used to print Ward's wordless novels: a complete set of 155 blocks for *Gods' Man* (1929), a complete set of 124 blocks for *Madman's Drum* (1930), and an incomplete set of 23 blocks for *Song Without Words* (1936). This original illustration (Figure 7) by Ward for his wife May McNeer's children's book *Go Tim Go!* is an example of an object that required classification in more than one way.

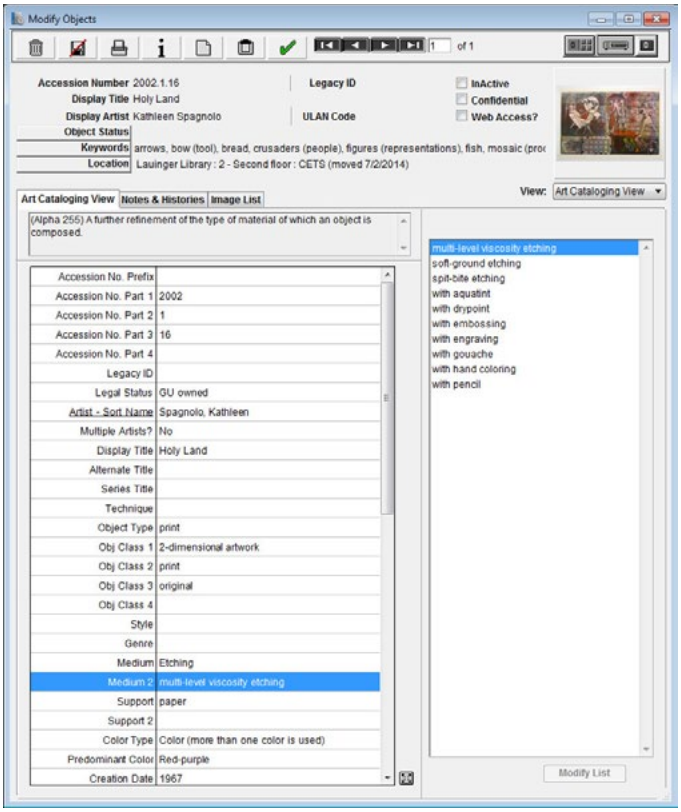


Fig. 6: Object record for Holy Land *showing Medium 2 as a refinement of Medium*

Fig. 7: Lynd Ward, original illustration in a book mock-up of May McNeer's Go Tim Go! *(ca. 1967)*

The artwork was defined as an original drawing under the object classification field, but was also differentiated as a book mock-up within the object type field in EmbARK (Figure 8). Both definitions are important for characterizing the nature of this object (a drawing that is one piece of the larger book mock-up), and anyone searching the catalog would find the object under one or the other term. In addition, all the individually cataloged drawings that are part of the same mock-up are connected via the Related Objects field (Figure 9) so that with one click a researcher can see a list of all parts of the whole. This necessitates a rather finicky process of connecting all related objects after they are individually cataloged, but it is a very important piece of information to have about a work that is part of a larger whole. This information will also be needed for artworks taken from a portfolio or those published in a series.

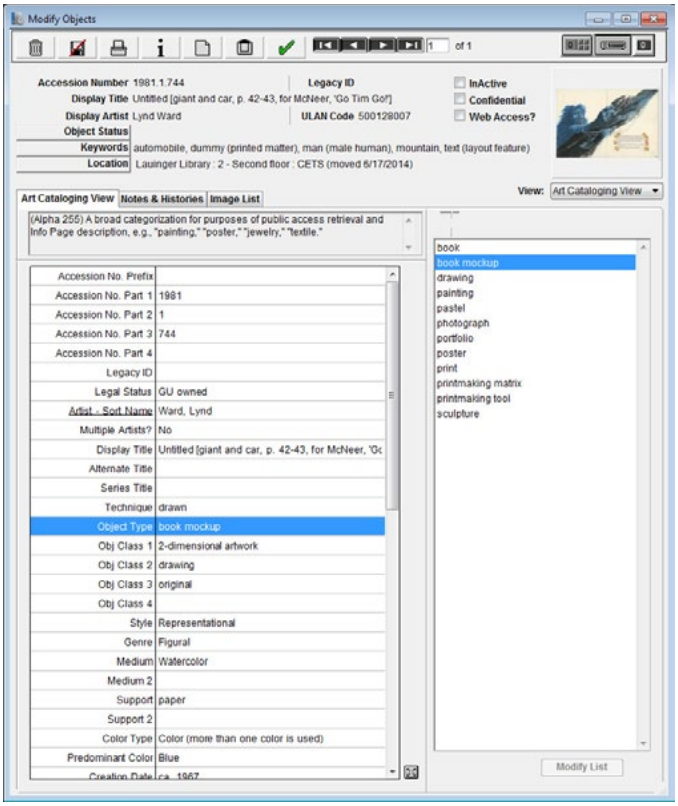


Fig. 8: Ward drawing that is both an original artwork and part of a book mock-up

Recording an Artwork's Evolution

The art collections of Louise Miller Boyer (1890–1976) and her daughter, Helen King Boyer (1919–2012), held at Georgetown University Library are the largest publicly held collections of their work. In the process of cataloging Helen Boyer's artworks, we noted the variety of artistic techniques and media that she experimented with and developed in her varying career as an artist. Research into the collection of this underappreciated artist revealed that Helen Boyer's talents went beyond merely the drypoint printing technique with which she and her mother are most associated. Throughout the course of cataloging, we discovered that some of the Boyer prints had multiple dates associated with them. Helen

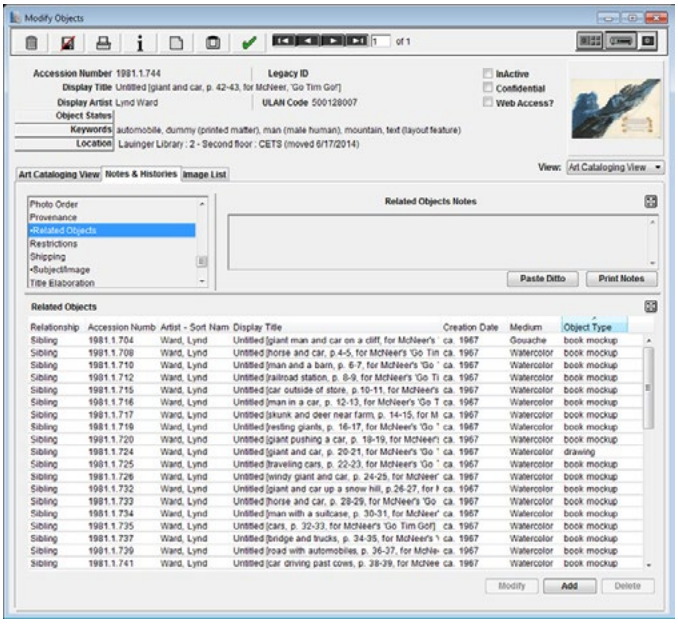


Fig. 9: Related Objects view showing all other components of the book mock-up

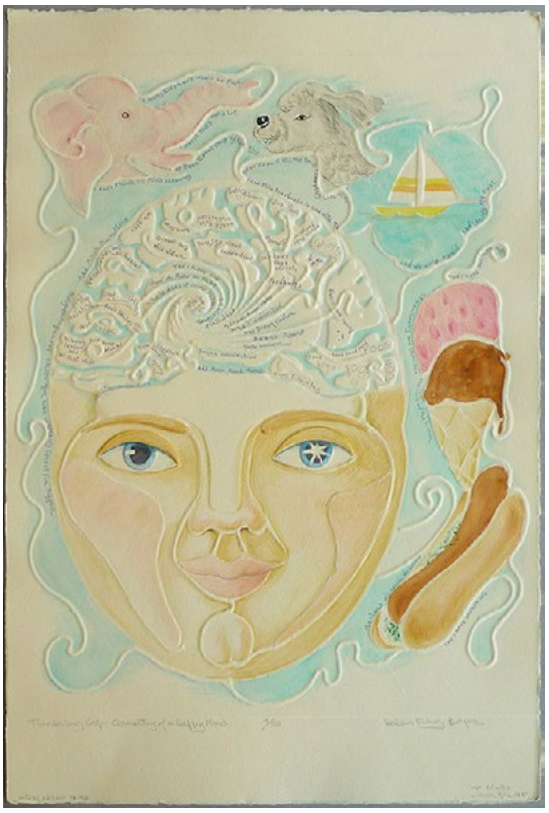


Fig. 10: Helen King Boyer, Thinking Cap—Circuitry of a Happy Mind *(1983)*

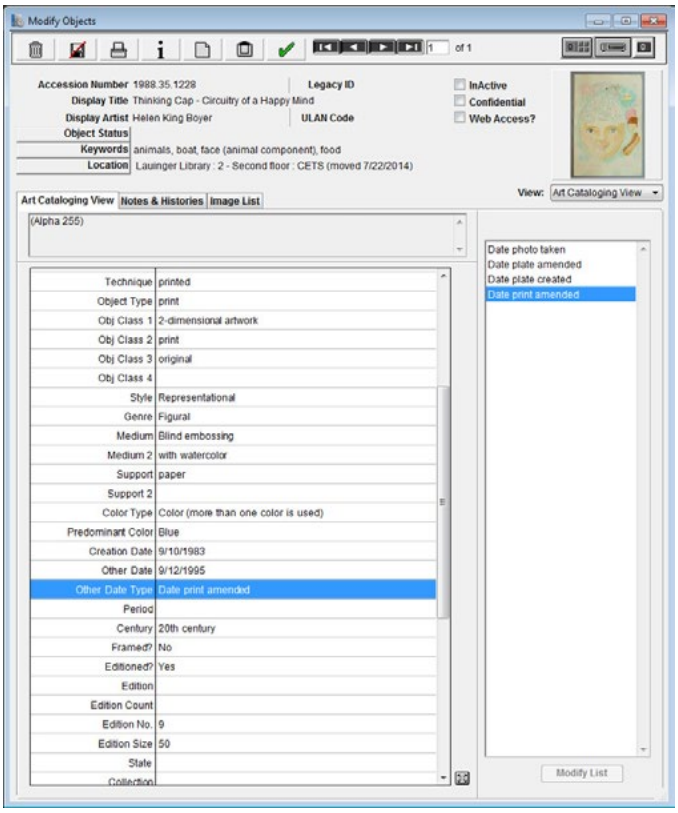


Fig. 11: Two different date fields allows for complexity in an object's creative history

Boyer's *Thinking Cap—Circuitry of a Happy Mind* (Figure 10) is a blind embossing with watercolor. Embossing is a technique in which a raised, relief image is created on the paper, using a dye that is passed through a printing press with the paper. When this is done without the use of any ink, it is described as blind embossing.

Helen Boyer kept detailed records about much of her work. This print includes a print mark stating that she made the print impression (blind embossing) in 1983, but she did not add watercolor to the print until 1995. We wanted to list both these dates for the object to provide as much information about the story of this print as possible. Therefore, we entered "09/19/1983" under the Creation Date field and "09/12/1995" under the Other Date field with the Other Date Type of "date print amended" (Figure 11).

The Other Date and Other Date Type fields were originally intended to record dates from other calendars, such as the Islamic calendar, or the Chinese calendar or dynasty name. Because such calendars are not used to date any objects in Georgetown's art collection, we decided to repurpose those fields. Setting up this ability to record and describe two dates for one object anticipated another common scenario for objects in the larger collection: that of a photograph, where the picture was taken on one date but the print itself was made later—sometimes decades later.

Georgetown University Library holds the only significant public collection of artwork by

Fig. 12: Marguerite Kumm, Christ Church, Alexandria, Virginia *(1942–1943)*

Marguerite Kumm. The collection includes 595 drawings (both preparatory drawings for prints and stand-alone drawings), 398 prints of varying techniques, and 123 printing matrices (i.e., the plates or blocks used to produce a print). Describing her style as Realist, Kumm depicted people in everyday situations, at work and play, in her prints. She was influenced by the work of John Sloan among others.

This collection of Kumm's work includes numerous preliminary works that accompany many of the final prints. These include multiple impressions of printed proofs in a variety of experimental states that led to the final artwork. In printmaking, a state is a different form of a print, caused by a deliberate and permanent change to a matrix. In these three images (Figure 12), we see several stages of development for Kumm's print *Christ Church, Alexandria, Virginia*. First, she made a preliminary pencil sketch of figures outside the church. The next image shows a sixth state proof of a similar image. This is the earliest proof that we have of this print, but the artist would have made five preceding states of this composition before reaching the version that we see in the sixth state. We can see that between the sketch drawing and the sixth state proof, the artist has revised the design. Lastly, we see the twelfth and final state of the print. The design is still very similar to that seen in the sixth state proof. However, Kumm has added more tone and detail, giving the image improved definition and perspective. This sequence of artworks demonstrates the detail-oriented nature of this artist's work and the intricacies of her working methods.

The EmbARK State field proved extremely useful for cataloging these types of preparatory print proofs (Figure 13). Although the artist indicated the state on the paper in some cases, in many other cases, the catalogers had to study all the available impressions of the same image, to identify whether impressions were earlier or later states of the print, based on slight

amendments in the composition. We originally supplied the State choice list with up to five states, but when we started cataloging Kumm's work we found the list needed to accommodate up to twelve. We also added "early," "late," and "final" to the list, for cases in which no state numbers were indicated in any of the proofs of a particular print. Recording the print state in each record allowed quick identification of the extent of different versions of each print that we had in the collection when viewing a list of artworks of the same title (Figure 14).

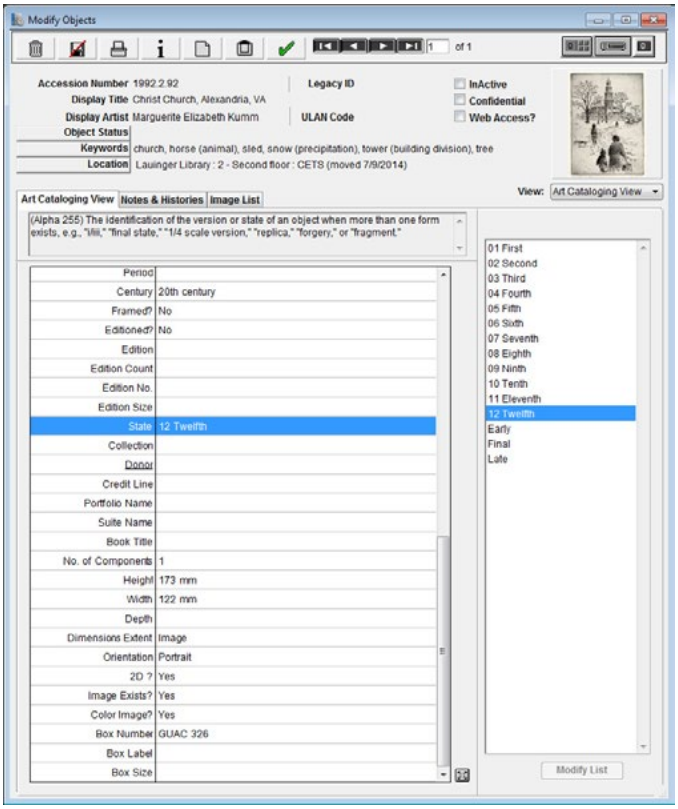


Fig. 13: Well-developed list of states

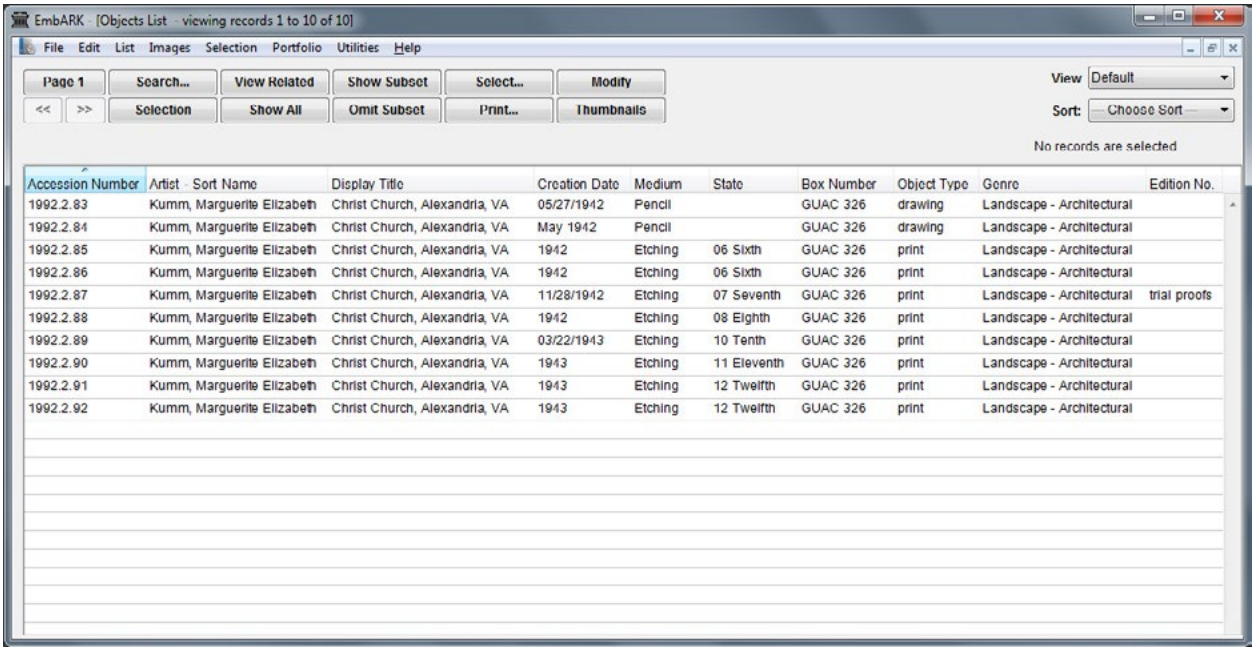


Fig. 14: Preparatory sketches and many states of Christ Church, Alexandria, Virginia

Describing Annotations as Key Elements

John DePol is recognized as one of America's finest wood engravers. A self-taught artist, DePol was devoted to his craft and produced an incredible output of work during his lifetime. His work ranged from commercial pieces with national financial printing firms to more personal illustrations for which he freely volunteered his time, working with associates in the private printing press world. The DePol collection includes many images depicting the subject of printing.

One of the most important aspects of cataloging prints has to be recording print marks (broadly, anything communicative on the paper other than the image). Because marks can reveal a great deal about an artwork and can distinguish between two almost identical prints, they are absolutely crucial elements of cataloging for scholarship. For this project, the marks that appeared on the prints in the collection ranged from the usual titles, signatures, and dates to the more esoteric marks, such as edition information, impression numbers, notes about who printed the artwork (if not the artist), and artists' technical notes.

The most common place to find an artist's signature is just below the bottom right corner of the image, but the artist's initials or signature mark might appear almost anywhere—including within the printed image itself (in the plate). In DePol's wood engraving, *The Kelmscott/Goudy Press* (Figure 15), the artist incorporated a single letter D into the design (Figure 16), serving as a signature. DePol always included his first initial within his printed designs and often included a signature in pencil beneath the image as well.

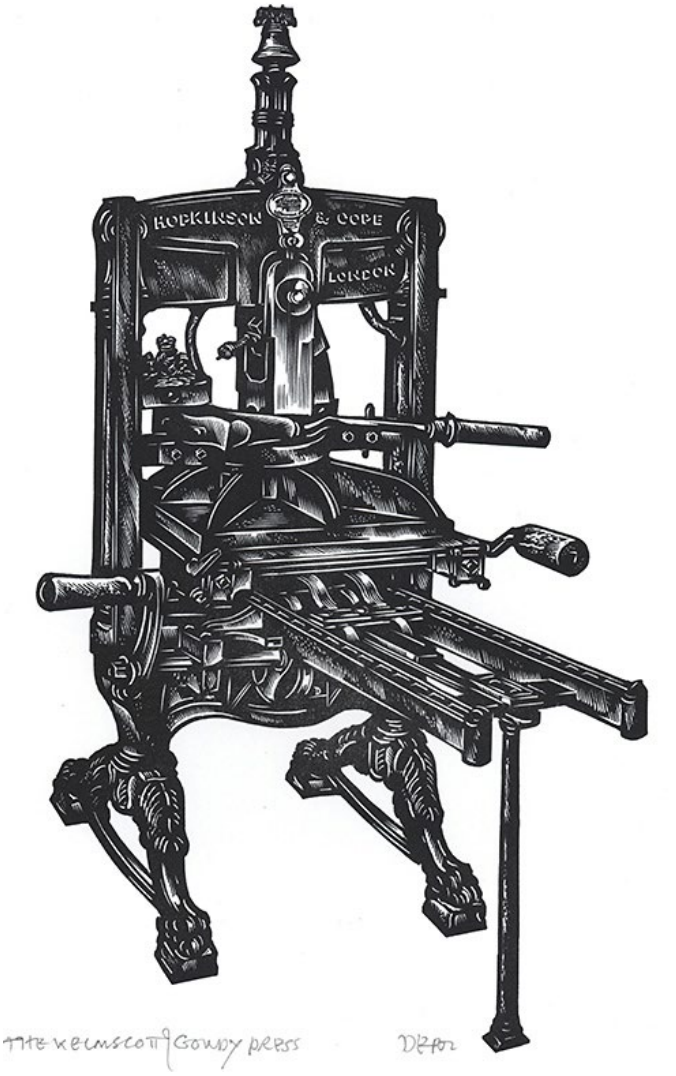


Fig. 15: John DePol, The Kelmscott/Goudy Press *(1984)*

Fig. 16: Artist's signature in the plate

DePol's initial would be difficult for anyone unacquainted with his technique to find, but it is a crucial point of identification for his work.

Sometimes the lack of a signature or other mark is the salient point. For prints created prior to the mid-1800s (of which the larger art collection has at least 1,000), signatures are far less common than they are on later prints; in the case of many old master printmakers, the presence of a signature is a reason to doubt the print's authenticity. Thus, it is essential to know that our cataloging of marks is comprehensive and detailed (i.e., to know that an absence of recorded marks is deliberate and not simply an omission).

EmbARK provides an extensive set of fields concerning print marks, which we were again able to tailor with drop-down lists to allow a breadth of information to be accurately recorded for each individual mark. We even created a comprehensive 25-item list of marks' possible locations (Figure 17). Recognizing that future catalogers might not have terms such as recto, verso, and margin on quick recall, the curators also created a visual key for easy reference (Figure 18).

For an important collection of prints such as the DePol collection, this level of detail is not excessive. Additionally, the specificity of the drop-down lists and their matching keys (we also have keys for taking measurements and describing colors) allows less experienced cataloguers to familiarize themselves with print morphology and empowers them to make these determinations on their own and to work more independently.

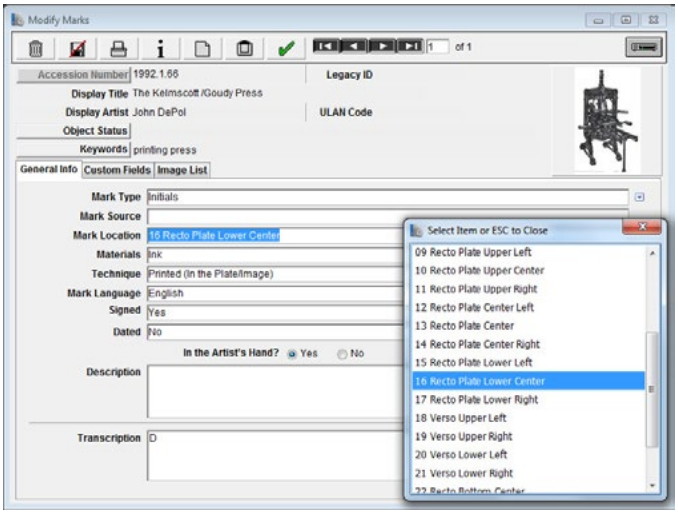


Fig. 17: Choices for recording a mark's location

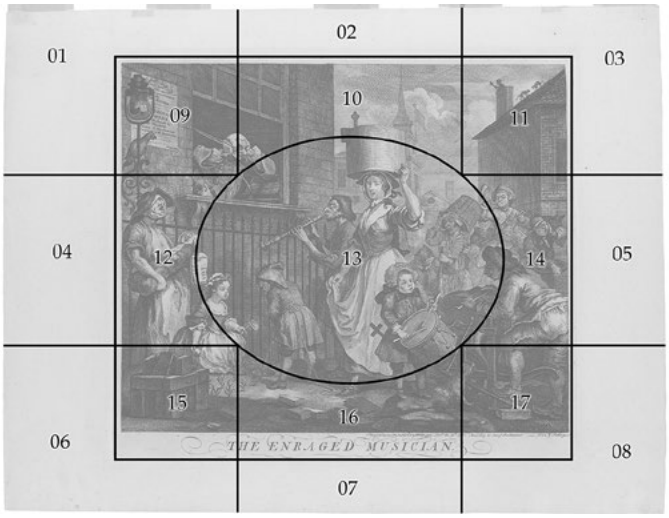


Fig. 18: Key for identifying areas of a print (William Hogarth, The Enraged Musician*, 1741)*

Fig. 19: Helen Boyer, The Spectator *(1979)*

In the early stages of processing the Helen Boyer collection, we noticed some unusual marks that appeared regularly in the bottom left corner of a number of Boyer's prints (Figure 19; detail, Figure 20). At first, we presumed that the dashes represented 1's, but the numeral in combination with various quantities of dots was a puzzle. What, if anything, was being counted?

It was not until later in the course of cataloging, after studying multiple impressions of many different prints, that it became clear. Boyer used a tallying system of dots and dashes to record the sequence in which each impression was printed in each edition she made. She marked a dot for each impression number up to five, then a line and dot system for numbers five onward, so here we determine that this is the eighth impression she made of this print. EmbARK enabled us to record and fully describe this mark (Figure 21).

Fig. 20: The Spectator, *detail*

We see the details for each mark, including the type, location, materials, and technique that were used to make it. At the bottom of the record, catalogers can add a transcription of the mark (i.e., exactly what it says), and above, they can add a description of what the mark means. Not visible here, but also present, is a field to record a translation of a mark in a foreign language. Even some non-Roman alphabets are accommodated; we have successfully pasted Arabic text into the Transcription field.

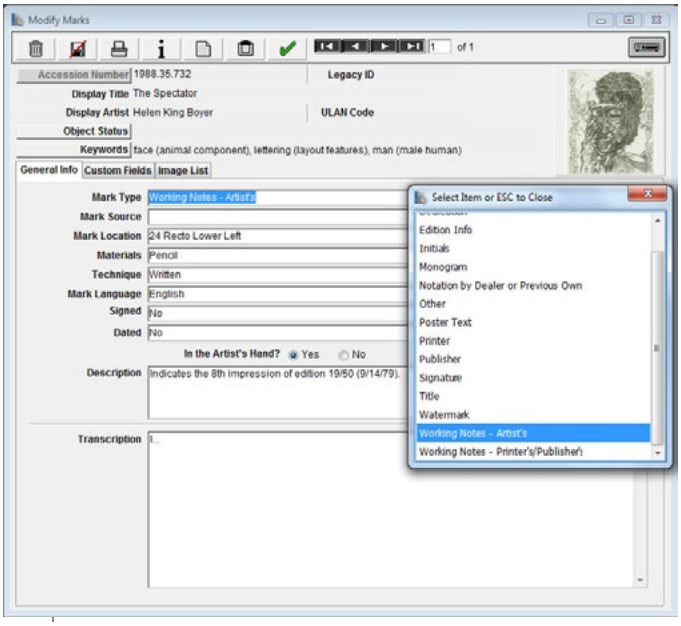


Fig. 21: Mark record for Boyer's impression tally

Documenting the Process

EmbARK's flexibility enabled us to make many customizations, resulting in a database specifically suited to our kinds of collections. Because we made changes from the defaults, it was crucial that we document the process in as much detail as possible. The discussions and deliberations involved in developing cataloging policies for this project were recorded as the project progressed; reviewing our thinking on previous questions often made it easier to answer subsequent ones. The cataloging decisions established were also summarized in reports at the end of the project. Catalogers referred to how-to guides for particular aspects of cataloging, including the basic techniques as well as those we have described. As a result, we have a reference guide for future print cataloging within Georgetown University Library's art collection, and the foundation of a comprehensive cataloging manual for the collection in the future.

Working through a fairly homogeneous collection as we familiarized ourselves with EmbARK and its capabilities was a good plan. The greatest difficulty—and it is a very minor one—was remembering not to tailor the database too narrowly for the Undiscovered Printmakers objects. Project staff had worked only with these collections and had many good ideas for customizations; however, some of these ideas would have been counterproductive to cataloging the collection at large. Having maintained both proper perspective and meticulous records, we are pleased to have produced an excellent collection catalog and a database distinctly suited to our needs.

The Churchill Weavers Collection: An American Treasure Uncovered

Jennifer Spence, Churchill Weavers Project Coordinator, Kentucky Historical Society

Abstract

In 2007, the Kentucky Historical Society (KHS) acquired the fabric archive from one of America's finest handweaving studios—Churchill Weavers. Following preliminary efforts to catalog the fabric archive at the item level, KHS staff determined that a hybrid of archival and artifact cataloging techniques would work best for this collection. Staff have since cataloged the archive at the box level, using the Archives module in PastPerfect to retain crucial information about weave structures, patterns, products, fiber content, and colors of more than 34,000 textiles. This paper looks at the challenges, successes, and innovative work that occurred in cataloging the textile collection. Presented are the unique promotional efforts making this important fabric archive known.

In 2012, the Council on Library and Information Resources (CLIR) awarded the Kentucky Historical Society (KHS) a grant under its Cataloging Hidden Special Collections and Archives program for the project, The Churchill Weavers Collection—40,000 Textiles Uncovered. This financial support enabled KHS to catalog its single largest acquisition, and one of the most important textile collections for Appalachia. The 24-month project ended on March 31, 2015. It made thousands of textiles by American handweaving legend Churchill Weavers available online and in person. Now that the collection is no longer hidden, it has the potential to make a profound impact on scholarship.

Churchill Weavers was a cherished Kentucky handcraft business and a leader in the American textile industry. KHS has dedicated significant resources to preserve the company's legacy and its place in Kentucky history. The Churchill Weavers Collection at KHS is vast; it includes the fabric archive, a comprehensive collection of three-dimensional artifacts, and business records documenting the company's productive history. It is rare to find a handweaving collection as complete as this.

As the largest component of the collection, the fabric archive needed the most work by KHS staff to make it accessible. KHS had hired a project assistant to evaluate the collection and design a cataloging approach in 2010. Yet it was not until 2012, upon receiving the CLIR grant, that staff made real progress with the collection. This paper looks at the cataloging methods developed for this specialized collection and the creative work behind it.

Churchill Weavers: An American Handweaving Legend

Churchill Weavers was a handweaving business that operated in Berea, Kentucky, from 1922 to 2007 (Figure 1). David Carroll Churchill and his wife, Eleanor Franzen, founded the company and ran it for more than 50 years. In 1973, the Churchill family sold the company to business partners (and husband and wife duo) Richard and Lila Bellando. The company changed ownership once again when the Bellandos sold it to Crown Crafts in 1996.

Fig. 1: Churchill Weavers loomhouse, ca. 1960

Churchill Weavers set the national standard for luxury handwoven goods. By the 1940s, the business had both an expansive network of salespeople across the United States and standalone retail shops in major U.S. cities. Fine department stores, such as Saks Fifth Avenue and Lord & Taylor, and small independent stores carried its products. Baby blankets, couch throws, and ladies' accessories were Churchill Weavers's mainstays, but the company produced everything from swing capes and bow ties to experimental cloth for NASA's first spacesuits.

The company used traditional handweaving methods to make fashionable textile pieces for an upscale market. Churchill Weavers's main consumers were middle- and upper-class women, though fashion designers, celebrities, and even European aristocrats sought out its styles. But for Kentuckians, Churchill Weavers was a handcraft business rooted in Appalachia; for many, giving a new mother a Churchill handwoven baby blanket was a longstanding tradition.

As a record of production, Churchill Weavers maintained a fabric archive that they called the Master Sample Collection (Figure 2). It contained a sample of every design, product, and experimental piece produced, as well as textiles in various stages of testing and production. If a customer questioned a product's design or color, employees could find its master sample to verify its quality and appearance.

Fig. 2: Master Sample Collection in the Churchill Weavers loomhouse basement, 2007

Churchill Weavers archived the textiles according to inventory numbers called style and cloth numbers, and stored the fabrics in cardboard shipping boxes. The boxes were the size of two conventional shoeboxes and printed with the Churchill Weavers slogan, "America's Finest and Largest Handweavers." Although the boxes mitigated damage from moisture, pests, and dirt, storing textiles in cardboard boxes is not ideal; thus it is remarkable that the samples had survived into the twenty-first century. Staff wrote the inventory numbers on the boxes and developed a card file system for finding items in the archive. Churchill Weavers hired an archivist in 2002 to write a basic finding aid; staff continued to update the fabric archive until 2006.

Churchill Weavers at the Kentucky Historical Society

After 85 years of operation, Churchill Weavers could no longer compete with less expensive foreign imports or overcome the organizational problems of its parent company, Crown Crafts. In 2007, Crown Crafts sold the Churchill Weavers

name and equipment to a handweaving business in Indiana and auctioned off the remaining items. Factory operations in Berea came to an end.

Lila Bellando recognized the importance of preserving the company's long history. After Churchill Weavers closed, Lila purchased the company's business and weaving records, fabric archive, and other artifacts and sought a permanent home for them. She approached the KHS about acquiring the collection. Elated at the opportunity, KHS secured private funding to purchase it from her.

The Churchill Weavers collection is KHS's single largest acquisition; the fabric archive consists of 2,369 boxes of textiles (Figure 3). Churchill Weavers and KHS staff did not have an official item count at the time of acquisition, although some estimates were as high as 100,000 samples.

Fig. 3: Fabric archive on pallets in KHS's museum storage, 2007

Besides the fabric archive, the collection includes business, marketing, and design records; swatch books, weft write-ups, draft files, and pattern books; photographs, audiovisual materials, and oral histories; and looms, signs, and tools. It is truly a comprehensive collection, both in breadth and scope.

Early Cataloging Efforts

Cataloging of the fabric archive began in May 2008. KHS created a small working exhibition to celebrate the acquisition and to share it with museum visitors. The exhibition, titled "Magic in the Weaving: The Churchill Weavers Collection Revealed," showed collection highlights and staff at work cataloging it. KHS's registrar, assistant registrar, archivists, and curators worked behind an area of the gallery walled off with plexiglas. This arrangement gave museum visitors a "behind-the-scenes" look at how KHS preserves archival materials and artifacts, and the opportunity to ask staff questions about caring for them (Figure 4).

Fig. 4: Cataloging stations in Magic in the Weaving, 2008

Collection staff chose to catalog the fabric archive at the item level, which is standard practice for documenting museum artifacts. It is a process in which each artifact is given an individual catalog number and collection record, and each item is marked with its catalog number. For textiles, the number is often written on a fabric label that is sewn onto the artifact. Catalogers took a photograph of one sample per box of textiles cataloged and completed a catalog worksheet by hand for each item. The worksheets had blank spaces to record information such as object names, measurements, weave structures, materials, and condition issues. By the time the

exhibition closed in September 2008, staff had cataloged 1,629 textile samples.

Two years later, KHS hired a part-time project assistant to reassess the fabric archive. She evaluated earlier cataloging work and determined that item-level cataloging was ineffective for this collection. Churchill Weavers and KHS staff estimated that there were 40,000 to 100,000 textiles in the archive. If staff continued to catalog the textiles item by item, she projected it would take them 10 or more years to catalog 40,000 pieces. KHS could no longer dedicate a team of staff members to the project as it had in 2008. Besides, item-level cataloging would generate thousands of database records, overwhelming researchers and staff using the collection.

A New Cataloging Approach

The project assistant worked with KHS's special collections administrator, registrar, director of museum collections and exhibitions, and the director of special collections and library to develop a new cataloging strategy. After several months of re-evaluating the collection and looking at several cataloging methodologies, they agreed on a hybrid approach of item-level and box-level cataloging.

In evaluating the collection, the project assistant found that the archive's inherent structure would lend itself well to a hybrid approach. The contract archivist hired by Churchill Weavers arranged the collection into archival series and subseries based on product types. When creating the arrangement and developing the finding aid, the archivist took into consideration the company's organizational system of keeping similar styles, fabrics, and patterns in the same box or series of boxes. She insisted on keeping true to the archive by retaining items in their original order. Although she used abbreviated terms to describe items and her resulting finding aid was skeletal, it became a helpful guide for cataloging the collection at KHS. It provided an overview of the collection, explained how Churchill Weavers had organized it, and took into account the inventory system's significance.

The project assistant determined that the new cataloging strategy should reflect the hybrid nature of the fabric archive—it is both an archive and a collection of individual artifacts. Churchill Weavers had done preliminary work for box-level cataloging; however, it was necessary to translate their records into a system that would work for KHS staff and for researchers. The project assistant's goal was to ensure that the new strategy would capture item details crucial to anyone studying handwoven textiles and present them in an effective, yet efficient format.

KHS uses PastPerfect Museum Software for documenting its collections. PastPerfect has separate catalogs for cataloging artifacts and special collections. These modules are simply called Objects and Archives. The museum collections and exhibitions team, who manage KHS's artifact collections, including the Churchill Weavers fabric archive, normally uses the Objects catalog. The team decided that the Archives catalog would work best for this new approach, however, because it would facilitate both item-level and box-level cataloging in ways that the Objects catalog cannot (Figure 5).

The project assistant began the process of cataloging the archive. In this new scheme, the boxes were cataloged one at a time, in original order, and a database record was created for each box. Detailed box information was captured in two crucial locations in the archives catalog record: (1) scope and content and (2) container list. The scope and content field was used to record the

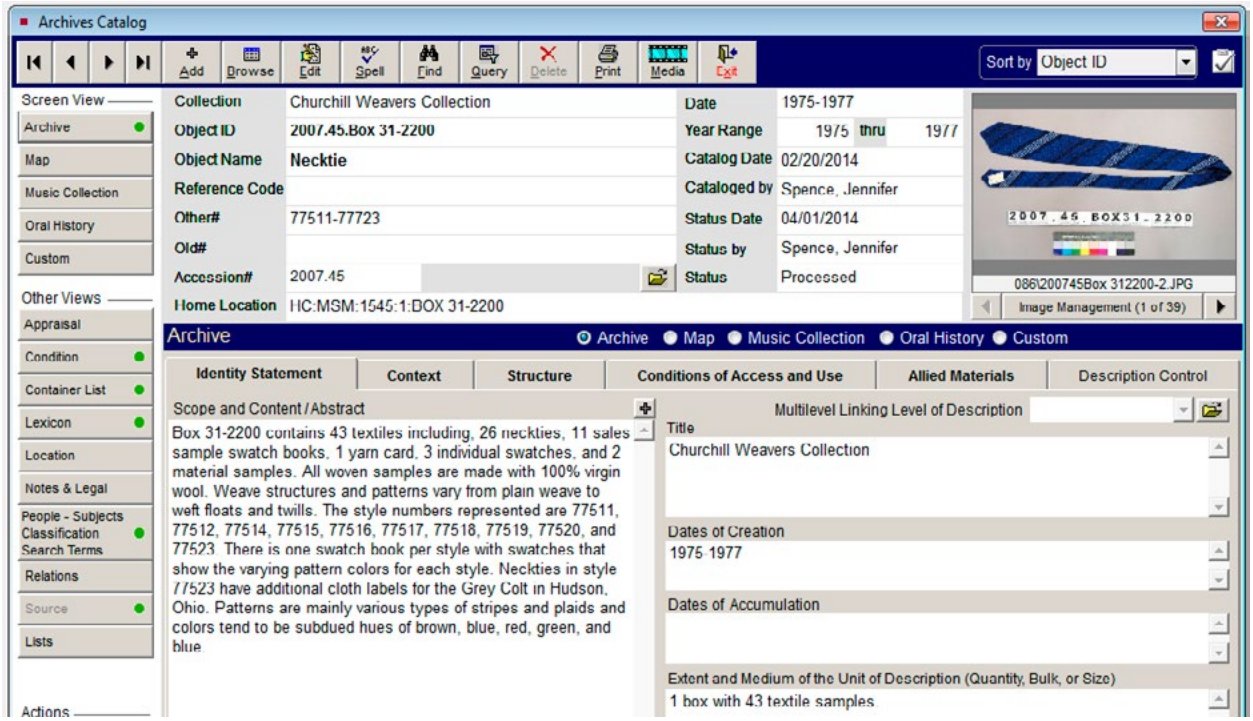


Fig. 5: Screen shot of an Archives record

overall description of objects in the box, such as numbers of items, product type (e.g., blanket, scarf), product format (e.g., swatch, final product), style numbers, weave structures, patterns and pattern numbers, colors, and fiber content.

Container lists were used with each box record to itemize textiles (Figure 6). This approach made it possible to isolate unusual and particularly interesting artifacts with more detailed cataloging that is not covered by the scope and content. Many boxes had dozens or even hundreds of swatches. A container list record may contain information for two or more samples that are duplicates or very similar in design and style instead of creating hundreds of records. The container lists were used to group like items together, isolate them, and catalog them separately from other unrelated items in a box. Itemization details included attached notes identifying specific dates of creation or the name of the weaver who created that piece. The ability of both staff and researchers to access this detailed information is critical for this collection. Staff anticipate that future researchers will add more data to this expandable structure.

Other main fields used in the Archives record were collection, object name(s), object ID number, date, other numbers, administrative/biographical history, creator, other creators, system of arrangement, location, and condition report. Built into PastPerfect is Nomenclature 3.0 for Museum Cataloging, which is the standard cataloging tool for man-made objects. PastPerfect has fields for up to three object names based on the built-in lexicon. This feature worked well for this type of box-level cataloging, as there was often more than one type of textile object per box—a box may have a swatch of baby blanket fabric and a finished blanket.

One of the first challenges encountered was devising a numbering system for this hybrid system. The accession number for the fabric archive

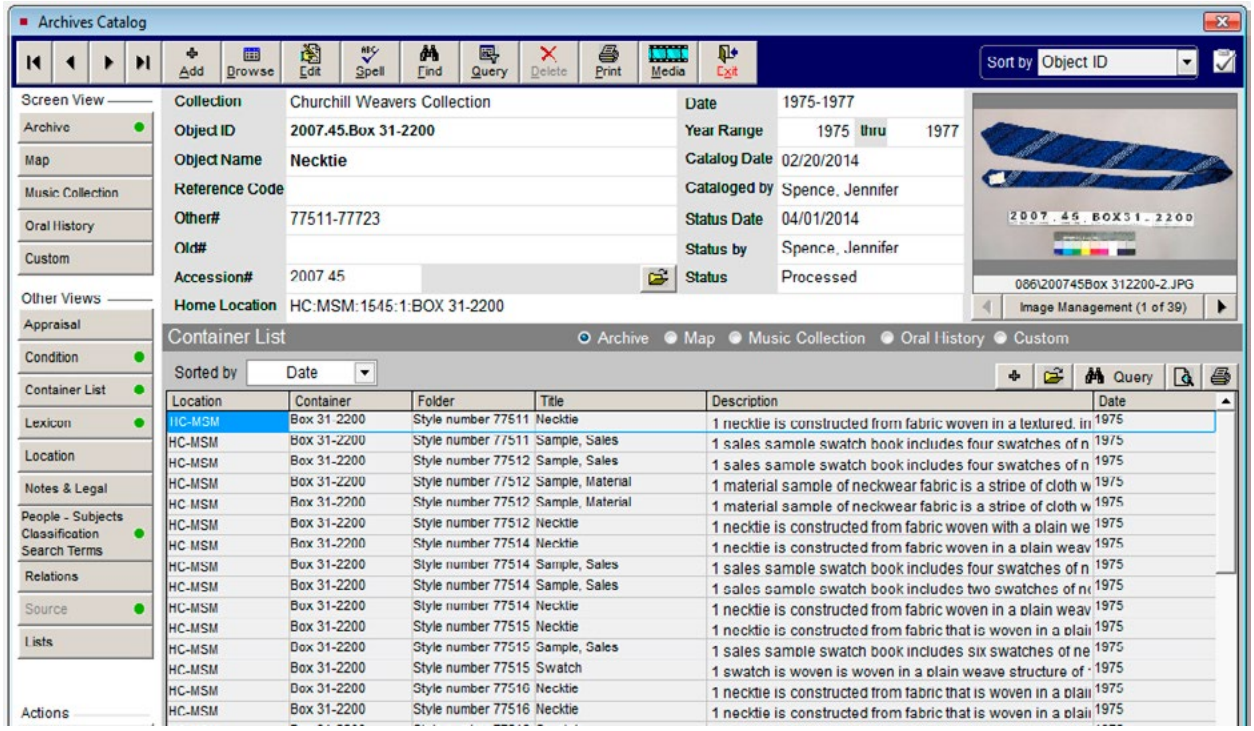


Fig. 6: Screen shot of an Archives record's container list

is 2007.45, and catalog numbers had been assigned to 1,629 samples in 2008. In this new system, staff assigned an object ID number to each box of textiles, but not to each textile. Staff devised a three-part, alphanumeric numbering system. In the number 2007.45.Box 31-114, for example, 2007.45 is the collection number, and Box 31-114 delineates the box number.[1] The textiles are not assigned catalog numbers. If someone removes a textile from its box, a fabric label with the box's catalog number is sewn on or pinned to the fabric.

From 2010 to 2012, the Churchill Weavers project team was the project assistant and a cadre of volunteers. They photographed the fabric archive in "box shots," where textiles of the same box were pictured together in groups of two or more. They photographed large finished products (e.g., baby blankets, throws), unusual artifacts, and items that cannot fit into a group shot individually (Figure 7.) An unlimited number of images can be attached to a single catalog in PastPerfect. Each shot was linked to the collection record with image information and uploaded to the online collection database. The project team believed it was important to provide a visual image for each box, both as an aid to researchers and as a conservation tool. The images taken allowed staff to record the condition of each box at the time of processing.

Fig. 7: Box record shot

1 Box locations for artifacts managed by the museum collections and exhibitions staff include the box's length, thus allowing for better management of the various boxes and their room and shelf locations. Staff store fabric archive textiles in 31-inch long newspaper boxes.

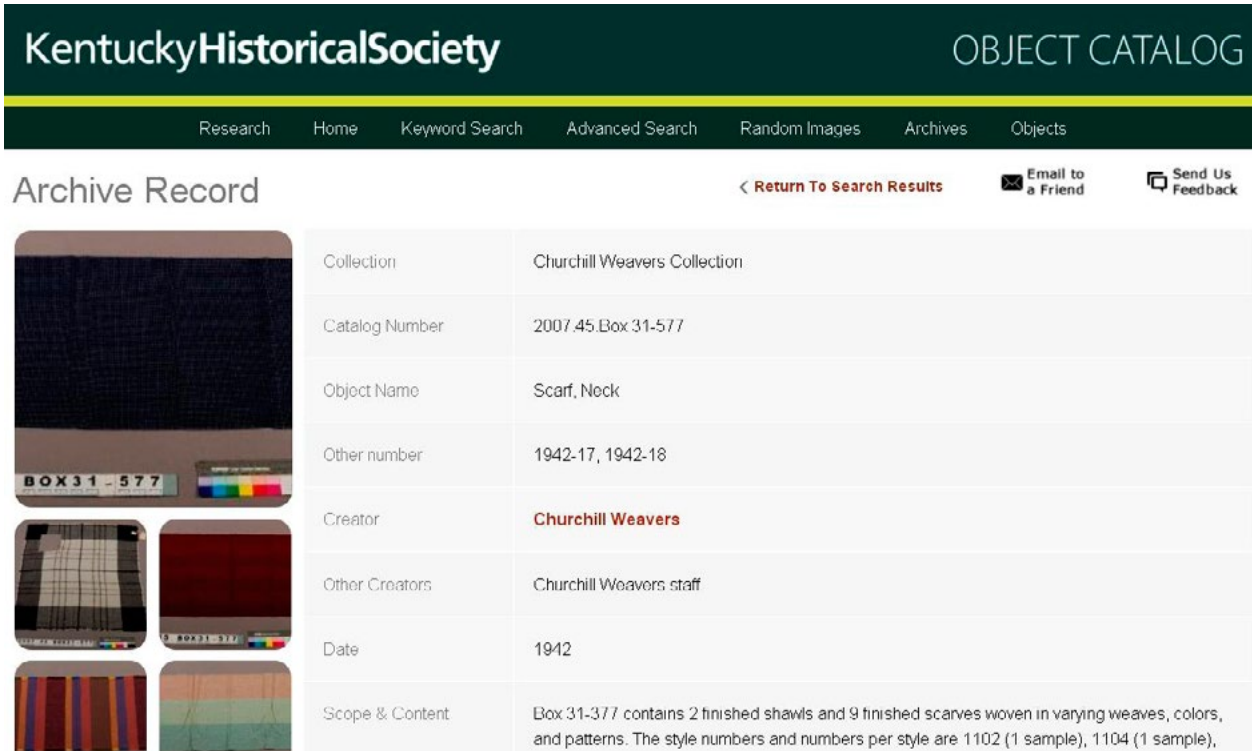


Fig. 8: Screen shot from a PastPerfect Online record

KHS uses PastPerfect Online as a record hosting platform and search tool. This is an easy and inexpensive service that uses MWeb technology and allows for easy Google indexing of records and customization of data displays. The online portion pulls directly from the PastPerfect database, and putting a record online was as easy as clicking a small "Include in Web Export" dialog box in the Archives record. Once the images were attached and the record verified for accuracy, the record was ready to be included in the queue for web export to PastPerfect Online. The export occurred weekly, and the data were transferred to an external server.

A select number of fields from the catalog record were exported to PastPerfect Online: collection, catalog number, object name, other number, creator, other creators, date, scope and content, system of arrangement, and any available images (Figure 8). Researchers can access the online Objects catalog through KHS's website, as well as at pastperfect-online.com, where artifacts from hundreds of museums, archives, and libraries can be searched. The online catalog can be found at http://kyhistory.pastperfectonline.com/.

The project team incorporated textile rehousing into the cataloging workflow. The fabric archive had been packed in small, brown acidic boxes. Once cataloged and photographed, all textiles were removed from these boxes, rehoused into longer and wider archival boxes, and supported with unbuffered acid-free tissue paper (Figure 9). Textiles found to be damaged by pests, mold, or mildew were treated accordingly. General textile cleaning was done if fabrics were found with dirt and accretions caused by previous storage conditions at Churchill Weavers.

The project assistant and volunteers were making progress with the collection, but they needed more help to complete the project. In 2012,

Fig. 9: Before rehousing (left) and after rehousing (right)

KHS received the Cataloging Hidden Special Collections and Archives grant from CLIR. In early 2013, the project assistant became a full-time project coordinator, and KHS hired a full-time assistant project specialist. KHS was now able to fully implement the hybrid strategy and complete collections cataloging.

The cataloging goal in the team's project plan was 25 to 30 boxes per week, and the team consistently met this goal. On average, the team cataloged 30 boxes per week, or 119 boxes per month. The hybrid cataloging approach was effective, although a small quantity of materials, such as bolts of yardage and sample books, needed item-level cataloging. After all of this, the team had an accurate item count of the archive. They revised the estimated item count from 40,000 to 100,000 items to 34,000. As of February 2015, 100 percent of the Churchill Weavers fabric archive is accessible, both online and in person.

Churchill Weavers Volunteers and Interns

Volunteers and interns were integral to helping the project coordinator and assistant meet project goals. More than 20 volunteers and 8 interns served on the project. Their responsibilities included everything from rehousing and photographing the textiles to marketing and promoting the collection (Figure 10). The project team would not have accomplished digitization initiatives without their help, especially because CLIR stipulated that no grant funds could be used for digitization.

Before KHS received the grant, the project assistant worked to train volunteers and interns in digital photography and photo editing. Project staff used experienced volunteers to help train new volunteers and interns over the granting period. Staff created training materials to help keep work consistent, including documents for rehousing, cataloging, and using Photoshop to clean up images before adding them to the database.

Fig. 10: A Churchill Weavers's project volunteer rehousing textiles

The backgrounds of our volunteers and interns were diverse. The project team had interns from visual studies, public history, and history programs at the University of Kentucky; Loyola University, Chicago; and Middle Tennessee State University. The collection inspired several of them to pursue careers in public history and museum work. One intern wrote her master's thesis about the company's innovative business practices, and KHS is working to publish it in its scholarly publication, *The Register*. Our volunteers had backgrounds and experiences that were incredibly beneficial to the project. Several of them were weavers and fiber enthusiasts who helped to name weave structures and fibers, and two volunteers were professional photographers.

Promoting the Churchill Weavers Collection

The project team worked with KHS's marketing and communications department to develop strategies for reaching target audiences. They identified weavers, textile scholars, and other museum professionals as the people who would use this collection. In the first year of the grant, KHS created and implemented a communications plan to attract these key people. The plan included activities such as using social media (e.g., Pinterest and Facebook), presenting conference sessions and educational programming, and developing printed marketing materials.

In 2014, KHS created a Churchill Weavers Collection brochure (Figure 11) and distributed it to more than 300 institutions. Staff targeted museums, galleries, university programs, and weaving guilds in the United States. Institutions and individuals have responded positively to the brochure. The National Museum of the American Coverlet in Bedford, Pennsylvania, was the first museum to express interest in the collection. KHS's development team found the brochure helpful for connecting with present and future supporters. KHS anticipates more interest from targeted institutions as the brochure continues to circulate.

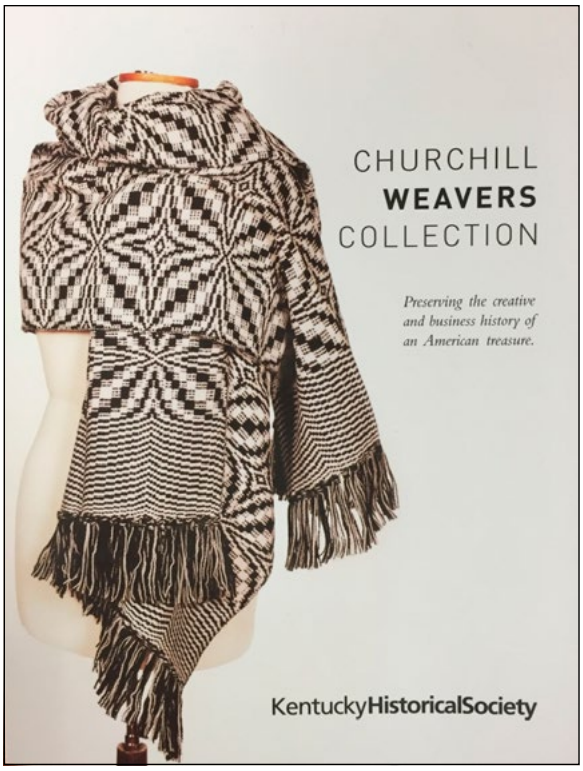


Fig. 11: Churchill Weavers Collection brochure

Staff developed a new line of Churchill Weavers–inspired products for the KHS gift shop (Figure 12). The items include a note card series, a Christmas ornament, magnets, a decorative box, and jigsaw puzzles. The products are a fun and creative way to promote the collection to museum visitors. Packaging on the merchandise includes links to the online catalog and collection information.

Fig. 12: Custom-made Churchill Weavers–inspired gift shop merchandise

The collection has had a strong social media presence. Project staff worked to promote it through Pinterest, Facebook, and the KHS blog, *Chronicle*; the staff also encouraged interns and volunteers to write blog and Facebook posts. The project coordinator and assistant wrote about the collection for several external blogs, including the Textile Society of America's *The Revivalist* and *Appalachian History*. They also presented several sessions at the Southeastern Museums Conference Annual Meeting in 2014, the Kentucky Council of Archives, and Kentucky Crafted: The Market. The project coordinator worked with KHS's education department to develop content for a fashion camp inspired by Churchill Weavers. Other outreach activities include work with Frankfort's Early Learning Village kindergarten school, where students are learning about weaving and the fiber arts.

KHS's Scholarly Research Fellowship Program developed a new fellowship for researchers who wish to study the fabric archive. The Churchill Weavers Fellowship is the first of its kind designed for a KHS artifact collection. KHS will use a private donation to fund the fellowship. The fellowship review panel selected two project proposals for 2015. The fellows' projects are to be completed by June 2016.

Conclusion

A weaving collection as complete as the Churchill Weavers fabric archive is extraordinary. Loomhouses in the Appalachian region rarely kept samples of their work. If they did, their collections often succumbed to pests, dirt, fire, and natural light. KHS is proud of the Churchill Weavers collection and the accomplishments in cataloging it. KHS is certain that the collection not only will inspire scholars and weavers, but also will enrich scholarship in many disciplines.

Cataloging the fabric archive was a daunting task, but the hybrid approach of item-level and box-level cataloging made the process more manageable. It is a new approach for KHS artifacts, and so far, it has worked well for this collection. Staff anticipate that future researchers will add to the data structure and make suggestions on areas needing improvement.

KHS is very thankful to CLIR and The Andrew W. Mellon Foundation for their financial support. This jewel of a collection would still be hidden without their help. KHS also thanks the volunteers and interns who have served on the project. To catalog, digitize, and rehouse more than 34,000 textiles is no small feat. Thank you, too, to philanthropists Joan Cralle Day and the Cralle Day Foundation, and Thomas P. Dupree for generous gifts that allowed KHS to acquire and preserve the Churchill Weavers collection.

Pennsylvania German Textile Cataloging

Candace Perry, Curator of Collections, Schwenkfelder Library and Heritage Center

Abstract

In 2011, the Goschenhoppen Historians, the Schwenkfelder Library & Heritage Center, and the Mennonite Heritage Center, all located in Montgomery County, Pennsylvania, received a Hidden Collections Grant from the Council on Library and Information Resources to catalog their collections of Pennsylvania Germans textiles. These collections represent a "common thread" of textile-making and use traditions based on the organizations' mutual Pennsylvania German heritage and insight into how an ethnically homogeneous community's material culture developed and evolved over time as a result of technology, new influences, religious beliefs, and availability of commercial goods. The chief objectives of the project were to: (1) catalog and photograph each organization's textile collections, and (2) provide access online through each organization's PastPerfect database. Secondary goals included a standardization of terms used for the various textile forms and the development of a lexicon of equivalent terms in the Pennsylvania German dialect.

The organizations hired a cataloger, Caitlin Harvey, to catalog and photograph their collections and enter the documentation into each organizations' online PastPerfect database. The outcome of this project was an extraordinary leap forward in the management, documentation, and accessibility of the collections that are essential both to our in-house activities and to research by others. The finished product greatly advances the standardization of terms and interpretation of Pennsylvania German textiles that will not only provide templates for the organizations' continuing documentation and preservation efforts, but will be a readily available resource for other collecting institutions and collectors.

Introduction

In 2011, the Goschenhoppen Historians, the Schwenkfelder Library & Heritage Center (SLHC), and the Mennonite Heritage Center (MHC) were the fortunate recipients of a Hidden Collections grant from the Council on Library and Information Resources (CLIR) to catalog their very significant, but relatively unknown, collections of Pennsylvania German textiles. This project was not only an innovative stride forward for these small organizations, but also it marked the first online accessibility of the Pennsylvania German textile collections as a cohesive, well documented entity.

The three organizations share common geography—all are located in Montgomery County, Pennsylvania—as well as the "common thread" of similar textile-making and use traditions based in their mutual Pennsylvania German heritage. Thus, a collaborative project that would focus on both the differences and the similarities of the collections was an obvious choice. The collections provide insight into how an ethnically homogeneous community's material culture developed into a distinctive culture and evolved over time as a result of technology, new influences, religious beliefs, and availability of commercial goods.

The chief objectives of the project were to: (1) catalog and photograph each organization's textile collections, and (2) provide access online through each organization's PastPerfect database. Secondary goals included a standardization of terms used for the various textile forms and the development of a lexicon of equivalent terms in the Pennsylvania German dialect. With the grant funding, the organizations hired a full-time cataloger, Caitlin Harvey, who worked with a small group of interns and volunteers to implement the project.

Background on the Three Organizations

Goschenhoppen Historians, Inc. Founded in 1964 as an educational organization, the Goschenhoppen Historians are nationally recognized for their expertise in the preservation and dissemination of the history of Pennsylvania German folk culture, specifically in the region known historically as the Goschenhoppen, which encompasses part of northwestern Montgomery County and northeastern Berks County in Pennsylvania. Their major projects include the Goschenhoppen Folklife Festival; the museum of Pennsylvania German folklife at Red Men's Hall, Green Lane, Pennsylvania; and the restoration of the eighteenth-century Henry Antes Plantation, a National Historic Landmark. The volunteer-based organization, registered as a 501(c)3 in 1967, is overseen by a board of directors.

A group of very dedicated volunteers, many of whom are original founding members, lead the organization. The main principal investigator on the project, Nancy Roan, is a recognized expert on eighteenth- and nineteenth-century quilts from the region and author of the book *Lest I Should be Forgotten* on local quilting traditions. Project consultant Alan Keyser, a member of the Historians, is an expert on weaving and woven textiles, and he is the author of several books on Pennsylvania German textiles, including *Forgotten Pennsylvania Textiles of the 18th and 19th Centuries.* A project priority was to capture knowledge about the provenance and the history of the textile collection from the original founding members while it is still available. Linda Szapacs provided assistance with organizing and inputting textile record information into digital form. Other Goschenhoppen volunteers assisted with cataloging, moving the collections, and other tasks.

Schwenkfelder Library & Heritage Center. Originally established as an informal collection in a private home, the SLHC was founded in 1884 to ensure the preservation of the cultural identity of the Schwenkfelders, an eighteenth-century German Protestant group. Today, the SLHC, a not-for-profit organization, is preserving, interpreting, and documenting the history of this group along with the local Pennsylvania German heritage and culture in their 15,000-square foot facility in Pennsburg, Pennsylvania. In addition to special collections of rare books, manuscripts, and photographs, the SLHC has significant collections of *fraktur* (the decorated folk art manuscripts and drawings of the Pennsylvania Germans), textiles, and other decorative arts. The organization employs five full-time and two part-time staff; adjunct staff serve as associate director of research and associate director of theology. The SLHC conducts symposiums, brown bag lunch series, children's programs, and workshops. Principal investigator Candace Perry, curator of collections, has been with the institution for 16 years and has published numerous articles and developed exhibits on Pennsylvania German culture and heritage.

Mennonite Heritage Center. The Mennonite Historians of Eastern Pennsylvania (MHEP) built the present-day Mennonite Heritage Center, a historical museum and library located in Harleysville, Pennsylvania, in 1990. The MHEP organization is a 501(c)3 and was incorporated in 1974. The Center's staff and board of trustees work to preserve and share information on more than three centuries of Mennonite faith and life in southeastern Pennsylvania. Archival collections include rare books and manuscripts, maps, broadsides, letters, genealogies, deeds, church records, and other printed materials. There are manuscript collections from many local Mennonite persons and families. The collection of 125 locally created and Mennonite-made fraktur is particularly significant. The artifact collection includes quilts, coverlets, samplers, clothing, furniture, farm implements, housewares, and musical instruments. Four full-time and three part-time staff members carry out a full schedule of exhibits, programs, and events. Principal investigator Sarah Heffner, director, initiated numerous programs and events relating to Pennsylvania German material culture as well as the annual Pennsylvania German Folk Art Sale that takes place in December. Joel Alderfer, collections manager, was the key MHC staff person on the cataloging project.

Collaborative Activities

The three organizations collaborate on programs, exhibits, and events. The Mennonite Heritage Center and the SLHC sponsored a workshop "Exploring the History and Artistry of Fraktur" in 2009, which was supported by a Pennsylvania Humanities Council grant. The workshop featured presentations by fraktur scholars Mary Jane Hershey, John Ruth, Allen Viehmeyer, Lisa Minardi, Candace Perry, Joel Alderfer, and Clarke Hess. Participants viewed historic fraktur from the collections of the SLHC and MHC, and toured historic meetinghouses and schools where the fraktur were created.

The organizations loan each other artifacts for changing exhibits. Staff at each organization frequently consult with each other and serve as guest curators for each other's exhibits. For example, the 2010 MHC exhibit "Comforts of Home" featured an early twentieth-century kitchen, parlor, and bedroom setting, and incorporated artifacts from all three organizations.

The Christmas Market tour is an annual event that takes place on the first weekend in December. It features exhibits and demonstrations at the Red Men's Hall of the Goschenhoppen Historians, the MHC, and the SLHC. Between 300 and 500 visitors attend the annual event.

Textile Project Personnel

Principal Investigator: Candace Perry, curator of collections, SLHC. Candace oversaw the textile cataloging at the Schwenkfelder Library and jointly managed the project with the other two principal investigators.

Principal Investigator: Nancy Roan, Goschenhoppen Historians. Nancy managed the project with the other principal investigators. She was in charge of organizing the textile collections at Red Men's Hall, the Goschenhoppen museum. Other volunteers from the Goschenhoppen Historians involved with the project included Bob Wood, who transported textiles to be cataloged and constructed shelving for the textile storage room; Sandi Karlson, who was responsible for maintaining spreadsheets of new acquisitions and their storage locations; Linda Szapacs, who worked with Nancy Roan to organize and direct the project; and volunteers Pat Gottshalk and Anne Grasberger, who assisted in labeling and proper storage.

Principal Investigator: Sarah Heffner, director, Mennonite Heritage Center. Sarah served as project administrator.

Project Advisor: Joel Alderfer, collections manager, Mennonite Heritage Center. Joel oversaw the textile cataloging at the MHC and managed the project along with the principal investigators.

Project Cataloger: Caitlin Harvey. Caitlin worked full-time from April 2012 to March 2014 to catalog the textiles of the Goschenhoppen Historians, the MHC, and the SLHC.

Financial Administrator for the Grant: Rose Moyer, assistant director, Mennonite Heritage Center. Rose maintained the financial records for the textile project.

Consultant: Linda Eaton, director of Winterthur Museum Collections. Linda provided guidance and direction to the project, particularly for developing a plan for the cataloging process.

Consultant: Alan Keyser, Pennsylvania German historian and textile authority. Alan provided guidance and direction, particularly on the woven textiles.

The Pennsylvania German Textile Collections

The Pennsylvania Germans produced their most distinctive material culture from the early eighteenth through the mid-nineteenth centuries. Among the most significant, but unexplored, parts of this culture were textiles. The collections of the Goschenhoppen Historians, the SLHC, and the MHC advance knowledge about the production, use, and traditions of Pennsylvania German textiles from this period. Included are early tablecloths, bed linens, coverlets, yardage, and grain sacks—domestic-use textiles that attest to home fiber production, an essential activity of rural daily life for the Pennsylvania Germans.

Needlework and quilts, usually the more decorative work found in the collections, were generally made by young girls and women. The making of samplers was an essential part of the domestic education of many women of German descent. By the nineteenth century, other types of decorative needlework created by young and older women replaced traditional samplers. Beginning in the mid-nineteenth century, patchwork and quilting became an activity enjoyed by Pennsylvania German women both for the provision of bedcoverings for the family and as a means of self-expression and creativity. The quilts in the collections record the patterns, colors, and textile preferences of the southeastern Pennsylvania German women as their acculturation with their neighbors and American life in general began to take root. The needlework and quilt provenance records will interest scholars in material culture and women's history, and will aid in understanding women's roles in the eighteenth- and nineteenth-century Pennsylvania German household.

The costume collections of all three organizations illustrate the early rural simple dress common to most Pennsylvania German groups in the eighteenth and nineteenth centuries. The distinctive Mennonite plain garb in the collections evolved in the early twentieth century to reflect the response to both the religious teachings of the Mennonites and the changing culture around them.

The Goschenhoppen Historians, the SLHC, and MHC textile collections are a significant resource for textile scholars because the collections were gathered and preserved in the context of the communities that produced them. Many of the artifacts in the three collections were donated by the original families, who had preserved them through the generations.

Compared with scholarship on furniture or ceramics, scholarship on textiles generally is still scant, and most of it centers on New England textiles. The cataloging project for the Pennsylvania German textile and clothing collections is contributing to the understanding of regional and ethnic similarities and differences in early America that continue to influence our lives today.

The Goschenhoppen Historians, the SLHC, and MHC textile collections were in various stages of cataloging at the beginning of the project. Volunteers for the Goschenhoppen Historians had completed basic inventories of much of their collection of 700 textile objects and worksheets for most of them. The SLHC collection of 900 textile objects had an extensive backlog of both accessioning and cataloging as a result of nearly a century of collecting. The MHC collection of 1,960 objects was 80 percent cataloged, but the information was in various formats and levels of description. The Center had an objects card catalog and worksheets, and the accessions of the last six years were in the Past Perfect database. Almost the entire MHC collection had been accessioned. Very little, if any, of the three collections had been photographed. The three organizations viewed the CLIR grant as an excellent opportunity to advance the care and stewardship of their collections.

All three organizations had a commitment to collection stewardship, but shared budgetary constraints. The SLHC and the MHC have full-time staff responsible for museum, library, and archival work, while the Goschenhoppen Historians operate with a dedicated and knowledgeable corps of volunteers. The SLHC undertook a capital campaign for building expansion in 2001 and at that time hired a curator of collections and an archivist. The MHC expanded its staff with the addition of a full-time archivist in 2006 to address the cataloging backlog, which persisted even though the collections manager had several summer college interns who had been assisting with cataloging projects before 2006. The SLHC and the MHC participated in the Advanced Stewardship Program of the Conservation Center for Art and Historic Artifacts, Philadelphia, and now have detailed preservation plans with annual goals. The Goschenhoppen Historians have volunteers and a collections committee who regularly dedicate time to working on the documentation of the collections, and a member of that committee holds the position of curator.

The Goschenhoppen Historians, MHC, and SLHC did not seek joint funding for their textile project from sources other than CLIR. The three organizations have been working steadily at cataloging with the goal of moving to digitization, but presently lack the staff and budget to handle the workload. Coming from small institutions committed to professional museum standards, we thought that the CLIR funding opportunity presented an excellent opportunity to have our textile collections fully cataloged and accessible. We appreciated the funding priority for hidden collections because, as small institutions that are competing with larger, more nationally recognized organizations, it is sometimes difficult to garner support or funding for projects.

Key Points

Project advisors played a key role early in the project in helping set the course. Winterthur Director of Collections Linda Eaton was very generous in sharing her cataloging expertise. The early meetings with her were influential in determining the course we took for cataloging; she also helped us understand that our project goals were ambitious and that we should probably reconsider them in light of the amount of time we had. She invited the project staff to visit with the

Winterthur cataloging team to discuss how to move forward efficiently with our project. In our discussions of categories and textile terminology, Linda and the Winterthur cataloging team advised cataloging by category instead of by each of the three collections as originally stated in the grant proposal. That advice proved very helpful in facilitating the workflow.

Linda also advised that we keep terminology simple and direct. In addition, she recommended that we record how much time was spent at each location and how many records were completed in that time so that we had a fair sharing of labor and progress among the three organizations and stayed on track according to priorities established. It was straightforward, practical advice, but very helpful. Winterthur's cataloging process should be a model for any museum collection where there is a cataloging backlog.

Selection of the right project cataloger was also important. The grant required a new hire for the project. We were very fortunate to find Caitlin Harvey, a young professional with a master's degree in the history and culture of fashion from the London College of Fashion, London, England. Caitlin had a good general background in textiles, specifically costume, but was new to Pennsylvania German textiles. Perhaps most importantly, she was flexible in working with our three similar, but distinctly individual, organizations and in juggling time with project staff who were also working with event and exhibit schedules. Caitlin brought a great deal of energy to the project, which we found was necessary for its timely completion, as well as for the physical demands of the work.

The project principal investigators knew each other well and were able to work through problems that emerged. Developing standardized textile description terms for the project was an involved process because of the specialized and distinctly Pennsylvania German nature of many of the textiles. The project staff had to balance the desire to be comprehensive with the need for brevity for the online catalog fields. We explored new ground with the database, both with the identification of the objects and the use of appropriate terminology. Pennsylvania German textiles share features in style and construction with other European folk and immigrant textile traditions, which can either aid or confuse identification and interpretation. As cataloger, Caitlin sorted through varied descriptions from all three institutions and worked toward standardization that we hoped would serve not only our researchers, but also other museums that own collections of Pennsylvania German textiles.

The project staff had Caitlin begin by photographing, measuring, and cataloging all the woven coverlets and quilts. The assistance of volunteers in hanging textiles and taking notes was key in this part of the project.

One of the main problems of the project was that the time needed to organize, photograph, measure, and record data for all the objects had been underestimated. Therefore, for the categories other than quilts and coverlets, we prioritized by provenance and rarity. For example, some categories of clothing at the Mennonite Heritage Center are repetitive, so we chose not to attempt cataloging all of the dozens of black wool shawls and net head coverings, but rather selected a sample group based on the provenance or the uniqueness of the objects.

Caitlin worked at both MHC and the SLHC, and the Goschenhoppen Historians brought their collections to the SLHC for cataloging. She shifted her work station every few months, which made it possible for her to work with staff of all three organizations.

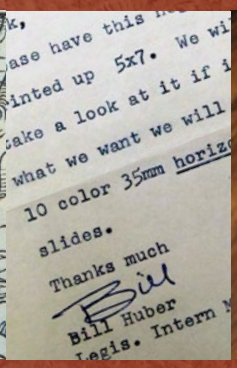
Regular meetings kept the project on track and allowed us to make collaborative decisions. Regular project meetings helped keep the lines of communication open and project expectations realistic. In the grant application, we ambitiously stated that we would have monthly meetings. We did not quite achieve that timetable, but instead met when beginning a new textile category or when Caitlin had questions to present to the group.

Conclusion

In early 2014, Caitlin completed her work on the project. The finished product is still evolving and probably will continue to evolve as researchers study the collections and staff are able to provide better provenance and other documentation. However, the CLIR grant allowed us to make an extraordinary leap forward in the management, documentation, and accessibility of the collections by creating records in the database that are essential both to our in-house activities and to research by others. It also pushed the Goschenhoppen Historians to implement new technology that gives them twenty-first century tools for cataloging and tracking their collections. Our finished product greatly advances our effort toward the standardization of terms and interpretation of Pennsylvania German textiles that we hope will serve other organizations as well as our own.

The rewards and pitfalls of this intensive collaborative effort among organizations that are very similar in many ways and quite different in others—and the group's ultimate development of a finished product that will serve the organizations and their audiences for years to come—is an excellent model for other small organizations that are embarking on a joint project.

Arrangement and Description



The Benefits of Planning: Cataloging the Vertical Files of the Anton Brees Carillon Library

Joy M. Banks, MSLS

Jaime L. Fogel, MLS

Abstract

In 2013, Bok Tower Gardens received a Council on Library and Information Resources Cataloging Hidden Special Collections and Archives grant to catalog the vertical files of the Anton Brees Carillon Library. This paper will discuss the key elements of laying the groundwork for this project and how the careful consideration of software selection, processing and cataloging standards, and future access to and dissemination of the information prior to the project's start has helped ensure a smoother, more successful project.

In 2013, Bok Tower Gardens (BTG) was awarded a Council on Library and Information Resources (CLIR) Cataloging Hidden Special Collections and Archives grant to catalog the vertical files of the Anton Brees Carillon Library—a unique collection encompassing more than 80 linear feet and 40 boxes of materials related to the carillon, an instrument of at least 23 bells in chromatic series, and the professionals that play them. This collection is only a portion of what is often considered one of the world's largest collections of carillon related materials (Figure 1). Receipt of the grant was the culmination of years of planning and preparing and was a significant step toward the successful processing of this hidden collection. The careful consideration of software selection, processing and cataloging standards, and future access to and dissemination of the information before starting the project has helped ensure a smoother, more successful project.

Fig. 1: Items from the Anton Brees Carillon Library

Pre-Grant

Although flexibility is essential in any project, especially when working with hidden collections, project planners and managers should take time before project implementation to prepare adequately. While not every question can be

addressed before a project starts, pre-planning serves as a valuable exercise to save time and resources later. Planning may involve thinking about the audience for the collection and how they will interact with the materials; imagining how the users will want to access the materials; considering time spent for physical arrangement versus time for thorough description; and realizing how much can be done in-house versus any outsourcing needs. Considering these factors in advance will help reduce interruptions to the project workflow and streamline processing even when unexpected challenges arise.

Planning for the Anton Brees Carillon Library (ABCL) project began even before the grant application was complete. For several years the librarian, the principal investigator (PI) for the grant, had been evaluating and arranging the library's collection in preparation for a large-scale cataloging project. On the basis of this initial evaluation, two collections were ultimately selected for inclusion in the CLIR grant: the vertical files of the ABCL and the archives of the Guild of Carillonneurs in North America (GCNA). Inclusion of the GCNA archives in the project was due in large part to an official agreement brokered by the librarian for housing the archives at BTG. Though the GCNA archives have been housed in the ABCL since 1993, no formal agreement existed until 2012. When that agreement was made, it provided only for minimal care and processing for the GCNA archives because the ABCL staffing at the time was .5 FTE. At the time of the agreement, a recommendation was made to the GCNA to streamline its holdings to focus exclusively on its history and story of the professional association. The scope of the existing GCNA collection overlapped many of the ABCL files, so this recommendation would allow for collection right-sizing, conserving space, and shifting redundant materials from the GCNA to either fill gaps in the ABCL collection or be disposed of appropriately.

Grant Writing Process

During the grant writing process, several decisions were made based on work done in previous years. The CLIR grant provided funding for the greatest collection need: people. The librarian's time with the collections clearly showed that the only way to undertake a successful project was through more hands helping with the collections. While volunteers offered support, the scale of processing the vertical files and GCNA archives required more dedicated staffing. The first request in the grant was to increase the librarian's hours to full time, allowing a 50 percent commitment to the grant, hiring a second full-time employee dedicated to the project, and hiring a series of interns over three summers.

The grant application also required certain details about software and processing plans for the project. ABCL had recently acquired a Quick Start version of OCLC's CONTENTdm through state access to OCLC's FirstSearch database, and it was determined that this software would be best suited to the project. Not only did CONTENTdm provide customization possibilities for the unique aspects of the collection, it also allowed for easy uploading to OCLC's WorldCat database, making the newly created records almost instantly available to an international audience. The hope was to provide users with one place to access all digital assets rather than parsing various holdings out to separate databases. CONTENTdm also offered the hosting services that the ABCL needed. Because the Quick Start version only allowed limited entries, the request was made to upgrade the subscription, a significant cost, and slowly have the Gardens take over the annual cost of the product.

The collections grow annually, so it was necessary to find a method that could adequately describe what existed while allowing for future growth. Because CLIR emphasizes interoperability with other collections, we chose to implement EAD finding aids. This would allow a resource to be created for the GCNA archives, each of the individual carillons documented in the ABCL, the various bell foundries, and any individuals of significance. Items that fell outside the scope of these broad categories (such as audio recordings or other ephemera) would be cataloged individually.

After Grant Award Notice

The grant application process took several months to complete. While the librarian was hoping that the project would be accepted, work continued as usual in the library. When the award letter arrived in December, the librarian set the project in motion by hiring the library special projects assistant (PA).

First, we conducted a pre-inventory to get a broad sense of the materials to establish hierarchies for the finding aids. These hierarchies would also guide the physical arrangement as processing began. Alhough physical arrangement is often part of the structure of the collection, the ABCL and GCNA collections had no such limitations. Plans were made to maximize space and consolidate holdings, and processing space was limited. Bearing in mind the condition of the materials, that they are on two different floors, and that the spaces are often used for tours, a physical workflow was created to use a lesser-used archive space. The pre-inventory also determined the order of handling the materials to minimize reprocessing.

The ABCL had no tools for creating EAD finding aids, and we did research on various no- or low-cost options by reaching out to various listservs and forums, including CLIR, Society of American Archivists, Connecting to Collections, Gold Coast Archives, and the Florida Opening Archives program. We received numerous suggestions and further researched options including Archon, Archivists Toolkit, ArchivesSpace, PastPerfect, Excel, and Access.

It is worth mentioning that as this project started and research was conducted, EAD 3 had not yet been released. The PA worked with John Nemmers, descriptive and technical services archivist at the University of Florida and co-chair of the Florida Opening Archives Program, to determine the best course of action regarding this update. Its pending release meant that there would not be as many tools available using EAD 3, much less ones that would suit the needs and limitations of this project. Ultimately, we decided to forego EAD 3 and proceed with creating finding aids using EAD 2002, with the idea that they could be migrated in the future.

Many of the most popular tools for creating EAD finding aids are not feasible options for smaller collecting institutions. Open source software databases such as Archon, Archivists' Toolkit, and ArchivesSpace require knowledge of computer programming, investment of funds, or server storage space and support. We investigated more readily accessible options to implement EAD finding aids in our collections.

The PI and PA decided that using Microsoft Access had much merit. Access is pre-installed on nearly every computer with a Windows operating system, including those at Bok Tower Gardens, and is thus available at no additional cost. When installed on a shared server, an Access database allows multiple users access simultaneously. The software also includes the ability to create forms,

making the input of information easier than using an Excel spreadsheet, and information can be exported into various formats including Excel spreadsheets, XML, and PDF, as well as text and Word documents. The team was also excited that, should the notion of using Access be successful, it would be a simple solution that could be shared with other libraries and archives, a primary goal stated in the grant application.

Further research was conducted to see if such a database already existed, to avoid unnecessarily creating one from scratch. The National Library of Medicine (NLM) offers a free Access database to members of the History of Medicine Finding Aids consortium. Although Bok Tower Gardens is not a member, John Rees, archivist and digital resource manager at the National Library of Medicine, was kind enough to share the database. He explained that although they have switched to a web application and no longer use the Access database, it should still be an effective tool for creating EAD finding aids.

NLM's Access database proved to be a valuable tool, allowing the user to input information into unintimidating forms and automatically creating EAD finding aids that could be exported in XML. Because the database was created for use in medical libraries, it needed modifications to suit the needs of the ABCL's collections.

Some changes needed were merely cosmetic, such as removing the NLM's logo and replacing it with BTG's (Figure 2). More fundamental changes tested the limits of the PA's coding skills to ensure that both the visible and behind-the-scenes elements were correct and without error. The PA continued to work with Mr. Nemmers to ensure that the EAD output followed the "Best Practice Guidelines for the Implementation of EAD Version 2002 in Florida Institutions," a standard selected for use in the grant application process.

Fig. 2: Edited NLM Access front page

The PA also added the functionality to export the finding aids as PDFs. With this added feature, an organization or individual can now input their information into the Access database easily using its convenient input forms, and choose XML or PDF as export options with a single click (Figure 3). The resultant export can then be shared or uploaded in other systems.

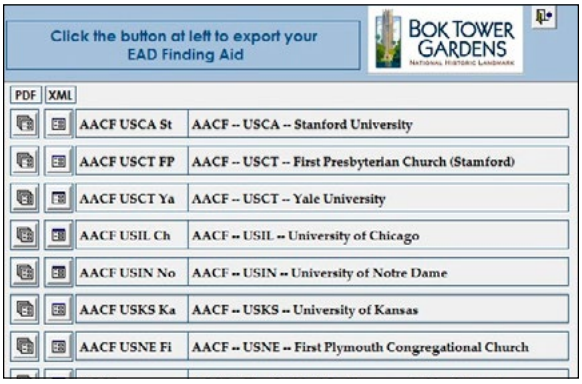


Fig. 3: One-click functionality for PDF or XML export

The NLM's original database instruction manual had not been updated to reflect later changes, creating minor obstacles in learning the database's nuances. These were quickly overcome and a new version of the manual was created.

Once we established the system to publish and how to create finding aids, we turned our attention to the daunting task of physical processing and arrangement. Three interns were hired for the first summer of the grant project to help process the collection. Following the workflow created by the PI and PA, the interns started by moving the entire GCNA collection from its two locations to the temporary workspace. Items that would clearly be shifted into the ABCL files were set aside, while the sorting of the remaining GCNA collection commenced.

Immediately, the PA needed to devise a system to allow multiple interns working on different days to easily sort the materials while always knowing where the last person left off and where work should begin. The workspace was a large table where the team sat, each with a stack of documents to be sorted into the predetermined series. The table was sectioned off using masking tape and marked to keep the delineation of the different series (Figure 4).

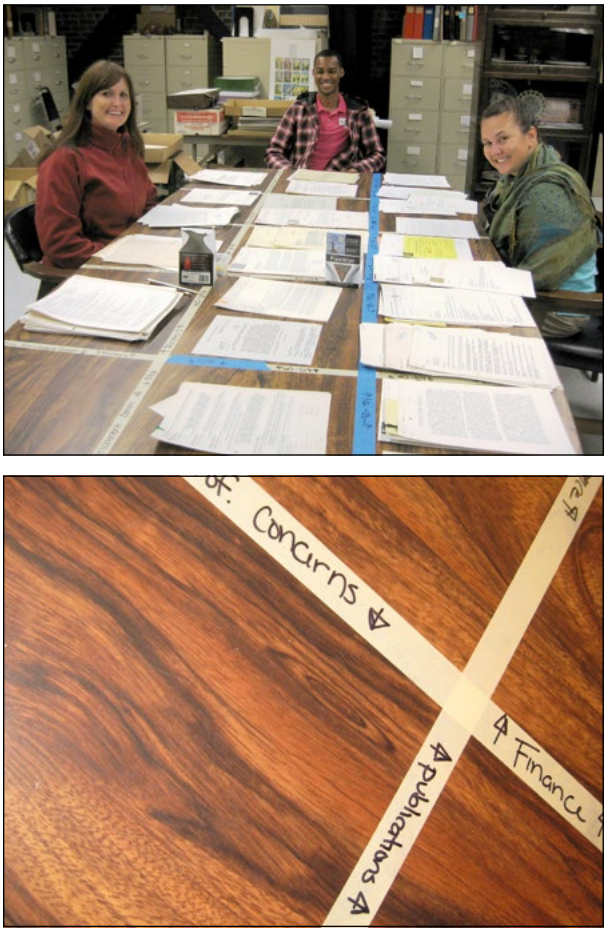


Fig. 4: Interns at work around our divided table

During the first few days of sorting the GCNA collection, it was discovered that this hidden collection was not as simple as it first appeared. The Guild activities and collected materials were complex, and it became clear that the hierarchy initially created was not the best solution for arrangement. The end user's potential use of the collection was constantly kept in mind. Exercising flexibility and remembering the old adage that getting it right is more important than getting it done quickly, the hierarchy was readjusted and sorting corrected.

After the initial broad sort by series, each was then sorted by subseries. This stage of the process raised many questions and challenges to consider moving forward:

- Documenting copyright permission to the GCNA vs. from the GCNA;
- Retaining the personal experiences of individuals attending the annual congresses while limiting space-consuming and unnecessary duplicates;
- Storing music manuscripts that do not fit in the vertical cabinets;
- Accounting for gaps in the collection, especially in the later years as more business was conducted electronically; and
- Capturing the grey areas of the organization, which often blur not only the responsibilities of different individuals and committees but also their professional and personal lives.

Fig. 5: Several examples of bell-related findings

Anyone who has processed an archival collection knows that many interesting items are uncovered when sorting through archival materials, and this collection was no exception. The team learned more than they ever expected to about the GCNA, the intriguing personal lives of the members of the Guild, and, of course, bells, bells, bells, and more bells. We discovered items unrelated to the scope of the collection, which had been retained simply because they contained the image of a bell. In fact, a new processing section was created for just such items: "Oh look, there's a bell!" (Figure 5).

The first summer of the project, with the assistance of the three interns, approximately 60 linear feet of materials were processed, including the most at-risk materials, which were housed in deteriorating boxes. The physical processing included rehousing items into acid-free folders as needed. Doing this during the sorting process allowed for appropriate grouping and descriptive labeling.

After completing the sort of the GCNA collection, we moved the materials back into their vertical filing cabinets, with some room left for growth of this active collection. Then the sorting of the second portion, the ABCL vertical files, began. Processing of these materials is ongoing as we create were finding aids for each of the carillons in North America, as well as many foreign carillons, notable carillonneurs and members of the GCNA about whom the ABCL has information, each bell foundry that produces carillon bells, and a smattering of other related topics.

Creating Finding Aids

As the physical processing of the collections advances, the creation of the finding aids will begin. Some finding aids for individual carillons were created during the employment of the summer interns. This allowed them to gain experience with the process, adding to the skill set they could list for future employment. The interns each provided valuable testing services, too, because they used the instructional manual to create the finding aids in the Access database. This process brought several issues to light that could then be fixed and retested prior to intensive finding aid creation. The creation of finding aids will include development of controlled subject headings; additional research to properly represent the instruments, individuals, and foundries in the Biographical and Historical Note section; and the inputting of the information into the Access database.

Though CONTENTdm was selected, initial testing proved that its functionality for finding aids is rather limited at this point. Conversations with CONTENTdm representatives confirmed that an XML finding aid, once uploaded using the finding aid wizard, cannot be updated or altered later. The same is true of a PDF that has been uploaded as a "compound object." This poses a significant problem because the ABCL collections are active

and will require regular updates. After additional experimentation and further conversations, it was determined that a PDF uploaded as a "single object" can be replaced. It was this discovery that prompted the development in the Access database for PDF export functionality. Uploading PDFs as a single object in CONTENTdm will allow regular updates while retaining the metadata and static URL for the item.

As processing continues, all finding aids will be contained in a single CONTENTdm collection, allowing them to be fully searchable. Once the finding aid collection is established, they will then be shared via OCLC's Digital Collection Gateway onto WorldCat. This method should allow for future linking between finding aids and any digital items created that are referenced therein.

Looking Forward

With an eye to the future, the PI and PA have discussed other avenues and outlets for the information contained within the newly created finding aids to become available not only to the carillonneur and academic communities but also to the general public. Some options being considered include seeking out a partner, such as the Florida Virtual School (home of the Florida Opening Archives program) or the University of Florida, to serve as an additional host for the finding aids, and the addition or creation of another type of web venue such as a wiki or blog. Attendance at regional meetings has revealed possible outlets including adding documented information to Wikipedia though the University of Miami's Remixing Archival Metadata Project.

The mission of this project is to bring a hidden collection to light, thereby increasing the global knowledge about carillons and drawing users into the other collections held at BTG. Bok Tower Gardens is committed to being a leader in carillon scholarship, and this is one giant step toward that goal. As more items are processed and brought to light, new research avenues can be discovered. Links between previously unknown facts and people can be made, and a new era of discovery can begin. The project will also help link carillon research to other subject interests such as engineering, general music scholarship, and art. Through flexibility in processing methodology and always keeping our end users in mind, planning and implementation of this project will assure that these materials can ring out to users for many generations to come.

References

The History of Medicine Finding Aids Consortium: http://www.nlm.nih.gov/hmd/consortium/about.html

Nemmers, John R. 2006. Best Practice Guidelines for the Implementation of EAD Version 2002 in Florida Institutions. Available at http://fclaweb.fcla.edu/uploads/Priscilla%20Caplan/FloridaEADguidelines.pdf.

Discovering the Future: The New York World's Fair Collections of 1939 and 1964 at the Museum of the City of New York and Queens Museum

Annie Tummino, Project Archivist, Museum of the City of New York and Queens Museum (past); Project Archivist, Columbia University Rare Book and Manuscript Library (current)

Abstract

This paper discusses the major components of the 18-month CLIR Hidden Collections project to process the World's Fair collections at the Museum of the City of New York and Queens Museum. Challenges and strategies for collaboration, staffing, processing, cataloging, metadata sharing, and publicity are shared.

In October 2014, the Museum of the City of New York (MCNY) and the Queens Museum (QM) completed an 18-month project funded by a Hidden Collections grant from the Council on Library and Information Resources (CLIR) to process the World's Fair Collections held at both institutions. With the 50th anniversary of the 1964 World's Fair and the 75th anniversary of the 1939 World's Fair in 2014, interest in the topic is at an all-time high.

Together, the museums hold about 12,000 items (approximately 360 boxes and 450 oversized/irregular objects) documenting the New York World's Fairs of 1939–1940 and 1964–1965. Items include books, pamphlets, printed ephemera, rare architectural blueprints, original artworks, film and audio recordings, photographic prints and negatives, architectural models, textiles, and realia.

Primary goals of the project were to create finding aids uniting both museums' collections, cataloging 1,650 highlights from both collections at the object level, and promoting the collection to the public and within the archival/museum community. I was hired as project archivist along with two part-time employees to implement these objectives under the leadership of the MCNY archivist and QM registrar.

MCNY connects the past, present, and future of New York City. It celebrates and interprets the city, educating the public about its distinctive character, especially its heritage of diversity, opportunity, and perpetual transformation. The QM, founded in 1972, has an intimate connection to its community and the history of its site. The QM building was originally designed to house the New York City Pavilion at the 1939 World's Fair, and its Panorama of the City of New York, of worldwide fame, was commissioned by Robert Moses for the 1964 Fair.

Benefits of Collaboration

The goal of creating joint finding aids for the World's Fair materials at MCNY and QM was to facilitate more efficient researcher access to these two related collections. The CLIR grant allowed these previously hidden collections to emerge into the light of day together, providing a single point of entry for researchers looking to learn more about the World's Fairs. The joint finding aids (one for the 1939 Fair and one for the 1964 Fair) elucidate where the holdings are similar and where they diverge. Additionally, local

versions of the finding aids were created to aid the reference process at each institution. Links to the joint and local versions of the finding aids can be found in Appendix A: Project Links.

The joint nature of the project allowed the strengths of each institution to benefit the other. MCNY's collection is considerably smaller, but it is embedded within a more formal manuscripts and archives program. Lindsay Turley, MCNY's assistant director of collections, provided input on the processing plan and training in MCNY's cataloging, processing, and preservation standards; this helped me develop standards used at QM. Dealing with the smaller MCNY collection first also allowed us to refine our processing plan for the immense QM collection. Louise Weinberg, the QM registrar, provided valuable context regarding the provenance and history of the collections and the World's Fair site.

When the grant proposal was written, neither institution had an online platform for sharing archival finding aids.[1] MCNY has an online collections portal (http://collections.mcny.org); however, the portal is not equipped to share metadata for objects without an accompanying digital image, and the grant did not cover digitization. The QM's collections management system is only available internally. Thus, the Queens Library was brought into the project as a consulting partner to share the finding aids and item-level catalog records through their online database. As Queens is the borough that hosted both fairs, its library is uniquely positioned to serve the needs of the local and scholarly community.

1 Since the start of this project, MCNY launched a "catablog" to share finding aids. The local versions of the World's Fair finding aids are shared on the MCNY catablog as PDFs, while the joint finding aids are marked up in EAD (Encoded Archival Description) and shared on the Queens Library site.

Diversity Fellowship Model

The CLIR World's Fair project built upon the previously established Diversity Fellowship program at QM, which aims to increase diversity in the archives and museum professions. The fellowships are intended for candidates from minority backgrounds or of recent immigrant origin. The fellowship pays a stipend of $10,000 for a year of part-time work.

Hoang Tran, a recent graduate of Drexel University's Master's in Library and Information Science program, and Richard Lee, a student in Pratt's School of Information and Library Science, were selected for the fellowship from a pool of highly qualified candidates. Wendy Jimenez, a student in City College's Art History and Museum Studies Master's program, replaced Hoang Tran approximately two-thirds of the way through the project. Lee is pictured in Figure 1.

The fellows carried out a wide variety of tasks, including processing, rehousing, drafting portions of the finding aids, contributing to the project's Tumblr feed, researching provenance, and item-level cataloging.

Fig. 1: Fellow Richard Lee at the Queens Museum

Processing Decisions

The original project plan called for intellectually uniting all World's Fair material from both fairs under the rubric of one finding aid, but it quickly became clear that this would have been confusing and unwieldy. Bringing together such diverse object types from across two institutions already provided more than enough complexity. Thus, the decision to create two finding aids, one for the 1939–1940 Fair and one for the 1964–1965 Fair, was a key early decision.

A second challenge of the project was to create a processing plan that would satisfy the needs of both institutions. As outlined in the grant proposal, I spent the first 4 months of the project at MCNY, followed by the next 14 months at the QM. To ensure a smooth transition from one institution to the other, the project staff made early visits to the QM and surveyed its collections. We were pleasantly surprised by the similarities of the materials, which allowed for more seamless intellectual integration. However, we discovered that some object types, such as architectural elements and audiovisual materials, existed only at the QM. Additionally, the QM held large quantities of unique materials on particular exhibitions, such as the Panorama of the City of New York of the 1964 Fair, which later became a permanent feature of the museum.

It was also a challenge to scour both museums' collections for all World's Fair materials. We searched both museums' collections management systems, physically surveyed their spaces, consulted legacy documentation, and harvested the institutional knowledge of the staff. At MCNY, World's Fair materials located in other artificial manuscript collections (e.g., the pamphlet collection) were removed and physically integrated into the World's Fair boxes. Items found in the decorative arts, costumes and textiles, and theater collections were intellectually described in the finding aids, though they continued to "live" in their respective departments. At QM we had to inventory items stored offsite and in the Visible Storage Gallery. It seemed that every couple of months, new boxes and materials would appear, unearthed from various nooks and crannies of the Queens Museum. Although some of the items we located at each museum had already been cataloged (or at MCNY, even digitized), they had never before been presented to the public as part of a unified World's Fair collection.

Much of the World's Fair collections at both museums had been loosely grouped according to subject and material type, with multiple gifts intermingled. The project team weighed the pros and cons of maintaining or restoring provenance and decided to create a subject-based arrangement for most of the materials. This arrangement reflected prior institutional practice and the nature of the collections, which consist largely of mass-produced ephemera. Maintaining provenance in an artificial collection comprising so many different gifts would create major headaches for researchers. Who would want to look in dozens of different places to see all the pamphlets on the General Motors exhibit?

To remain true to the character of the fairs, we turned to the official guidebooks to organize materials as the fairs themselves had been physically and intellectually organized. The 1939–1940 World's Fair was divided into seven "zones"; the 1964–1965 World's Fair into five "areas." As we used the guidebooks to familiarize ourselves with the organization of the fairs, we were occasionally surprised by what we learned. For example, the infant incubator at the 1939 fair, which housed premature infants viewable through glass windows, was located not in the Medical and

Public Health Building as we had guessed, but rather in the amusement area. Such distinctions demonstrate how public ideas of "amusement" have changed over time, illustrating the essential nature of primary sources when describing and investigating the past.

Not all of the items in the collections were intermingled. Provenance was retained for most of the photographic collections, and several gifts of unique materials with clear relationships to each other—such as Series I:B Donald Oenslager papers—were retained as their own series or subseries. Overall, the finding aids for the collections reflect an eclectic mix of subject, object-type, and provenance-based arrangements woven together into a cohesive narrative, with the overarching goal of facilitating ease of access and providing proper context. Administrative, biographical, and historical notes were incorporated throughout where pertinent, and complete donor appendixes were created for each institution.

Finally, rehousing the materials in these collections was no simple task, considering the diversity of object-types represented. For example, three-dimensional objects were stored in boxes with partition trays lined with volara polyethylene foam (Figure 2), and textiles were stored in flat boxes lined with abaca fiber tissue.

Fig. 2: Rehoused objects at Queens Museum

Cataloging and Sharing Metadata

As specified in the grant proposal, 150 collection highlights were cataloged at the item level at MCNY, and 1,500 were cataloged at the QM, reflecting the relative size of the collections.

At MCNY, archivist Lindsay Turley trained project staff in the use of Argus, the museum's collections management system, and on their cataloging standards, which are based on the Cataloging Cultural Objects (CCO) data content standard. We created a spreadsheet to identify highlights while we were processing and carried out the cataloging toward the end of our four months at MCNY. The provenance of each item was investigated and numbered based on whether it could be connected to a deed of gift.

The Queens Museum uses the PastPerfect collection management system, but did not have cataloging standards in place. To ensure consistency, I created a QM cataloging guide, making decisions about which PastPerfect fields to use and how to populate them (see Appendix A for a link to the guide). I consulted CCO for data content instructions, selected authority lists to use for particular elements, and developed a local vocabulary for Fair-specific terms. Because the QM staff is squeezed for time, I made the guidelines as detailed as possible for the benefit of future interns and volunteers. At QM we cataloged about 500 commemorative objects, 900 photographs, and 100 printed items, giving priority to the most unique materials. As at MCNY, we had to investigate provenance and number items before cataloging them.

As previously noted, the finding aids and item-level catalog records were shared with the Queens Library to make them publicly accessible. Sharing the finding aids was relatively simple, as the Queens Library staff agreed to mark them up

in EAD for posting. However, sharing the item-level catalog records was a bit more complicated.

Queens Library uses VRA Core 4.0 XML to catalog their photographs and archival ephemera. Although PastPerfect and Argus can both export records into basic XML, creating crosswalks and further transforming them into VRA Core records would be a fairly complex task. Thankfully, we were able to settle on an alternative option that worked for all parties: sending the records to Queens Library in Dublin Core XML. The simple 15-element Dublin Core standard may not offer the most complexity, but it is ideal for interoperability because it allows for a straightforward 1:1 relationship. Moreover, the Queens Library could readily ingest such records into its database, and PastPerfect's export utility includes a Dublin Core XML output. See Appendix B for the final version of the crosswalk deployed during the project.

Upon testing, I found that the Dublin Core metadata crosswalk in the PastPerfect software differed significantly from the crosswalk I had created in consultation with the Queens Library staff. However, the PastPerfect technical support team was willing to modify the export behind the scenes to match the project specifications. Since we used all four of PastPerfect's "catalogs" (objects, library, archives, and photographs), we essentially had to create four versions of the crosswalk.

The 150 highlights cataloged at MCNY became the test batch for the project. Since Argus (the software used at MCNY) did not have the capability to export records as Dublin Core XML, I exported them as CSV files, imported them into PastPerfect, and finally exported them as Dublin Core XML. Upon review, the Queens Library metadata team identified some minor problems that we were able to fix by modifying our standards. Once all issues were resolved, we set up a workflow wherein we sent 300 QM records to the Queens Library on the first Wednesday of every month over the course of five months.

Publicity

The project team used a variety of methods to publicize the collections. Several World's Fair-related posts were published on MCNY's blog over the course of the project (Figure 3), resulting in cross-posting and publicity on other blogs and news outlets. We also repurposed a Tumblr page that had been previously established by MCNY to advertise a World's Fair exhibition for use during the CLIR project (Figure 4). We posted two to four times a week, sharing gems we discovered while processing and celebrating the fiftieth and seventy-fifth anniversaries of the fairs. The Tumblr's audience grew from about 70 followers to more than 30,000 by the end of the grant. Tumblr posts were also shared through MCNY and QM's Facebook and Twitter accounts.

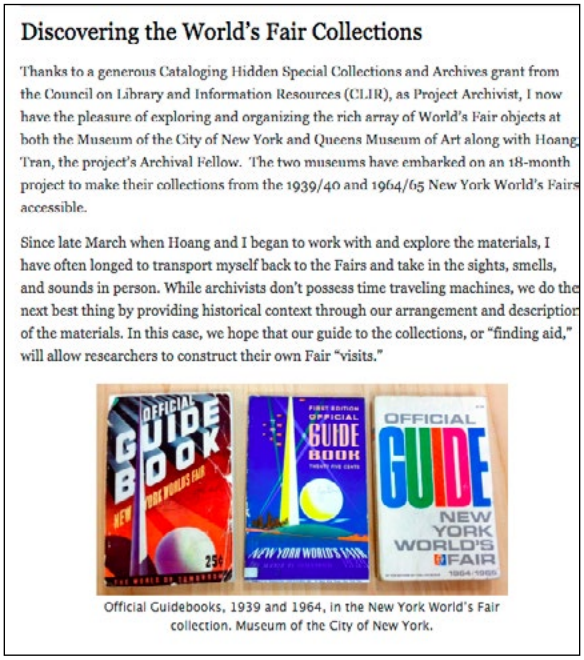


Fig. 3: Screen shot of MCNY blog

To publicize the project to the museum and archives community, we announced the opening of the collections on a number of professional listservs. I also contributed an article to the "News" section of the website of the Metropolitan New York Library Council (METRO), the largest reference and research resources (3Rs) library council in New York State (see Appendix A for link to the article). Finally, lessons from the project about provenance issues were shared at CLIR's Hidden Collections symposium and on a panel at the MARAC/NEA (Mid-Atlantic Regional Archives and New England Archivists) conference in March 2015.

Fig. 4: Screen shot of project Tumblr

Appendix A: Project Links

Finding Aids

1939–1940 New York World's Fair (MCNY only version): http://mcnycatablog.org/2013/10/01/collection-on-the-1939-1940-new-york-worlds-fair-1934-1993/.

1939-1940 New York World's Fair (QM only version): https://anniearchivist.files.wordpress.com/2015/06/1939wf-finding-aid-qm-10-10-2014.pdf.

1939–1940 New York World's Fair (joint version unifying MCNY and QM): http://digitalarchives.queenslibrary.org/vital/access/manager/Repository/aql:5891/EAD.

1964–1965 New York World's Fair (MCNY only version): http://mcnycatablog.org/2013/10/01/collection-on-the-1964-1965-new-york-worlds-fair-1959-1967/.

1964-1965 New York World's Fair (QM only version): https://anniearchivist.files.wordpress.com/2015/06/1964wf-finding-aid-qm-10-10-14.pdf.

1964–1965 New York World's Fair (joint version unifying MCNY and QM): http://digitalarchives.queenslibrary.org/vital/access/manager/Repository/aql:5893/EAD.

Catalog Records

http://digitalarchives.queenslibrary.org/vital/access/manager/Collection/aql:5839.

Standards

Queens Museum Guide for Cataloging Collections in PastPerfect: http://www.clir.org/hiddencollections/resources/queens-museum-guide-for-cataloging-collections-in-past-perfect.

Social Media

Collections Tumblr: http://nyworldsfaircollections.tumblr.com/

Posts on MCNY blog: http://mcnyblog.org/tag/worlds-fair/

Post on METRO.org blog: http://metro.org/articles/a-glimpse-of-new-yorks-past/.

Appendix B: Metadata Crosswalk

PastPerfect Fields to Dublin Core

Dublin Core Elements	PastPerfect Object Catalog Fields	PastPerfect Photo Catalog Fields	PastPerfect Library Catalog Fields	PastPerfect Archives Catalog Fields
Type	Object name	Object name	Object name	Object name
Format	Dimensions, material, medium	Print size, dimension details, medium, processing method	Physical description	Extent and medium of the unit of description
Title	Title	Title	Title	Title
Description	Description		Summary	Scope and content/abstract
Subject	Subjects, search terms	Subjects, search terms	Subjects, search terms	Subjects, search terms
Creator	Artist	Photographer	Author	Creator
Contributor	Artist 2, Artist 3	n/a	Author added entry	Other creators
Publisher	n/a	n/a	Publisher	Publisher
Date	Date	Date	Publish date	Date
Identifier (ID)	Object ID	Object ID	Object ID	Object ID
Source	Your institution	Your institution	Your institution	Your institution
Relation	Collection, related objects	Collection, related objects	Collection, related objects	Collection, related objects
Language	n/a	n/a	Language	Language
Coverage	Place of origin	n/a	Publication place	Publication place
Rights	Legal status	Legal status	Legal status	Legal status

Preserving a Montana Senator's Image: The Lee Metcalf Photograph and Film Collections Project

Matthew M. Peek, MA, CA, Montana Historical Society Research Center Photograph Archives

Abstract

One of the greatest challenges for archivists processing the papers of post-World War II U.S. congressmen is organizing, describing, and promoting their photographs. With the explosion in then-new public media forums (i.e., TV and increased newspaper publication) and cheaper, repeatable photographic processes, U.S. congressmen began taking and collecting more and more photographs for use in campaigns and public relations. The 1950s, especially, ushered in a new era of awareness by members of Congress of the power of images; the importance of documenting their work, campaigns, and constituency; and the need for still images for films, commercials, and congressional hearings.

After World War II, congressmen began to select their own images and to determine the means through which the public and the press digested the images of public officials at work. Congressmen began collecting negatives and having the same image cropped in several ways to reflect different visions or uses. This led to multiple photographic prints showing different actions, all born from a single image. This new use pattern has been little documented in the archival world, and some institutions struggle with how to identify and organize their congressional photograph collections.

Image collections from this time period present two challenges. First, little was written about the subject's life because of the newness of the congressman's service (and depending on the congressman, nothing may have been written since or the records may not be available for such publications). Republican and Democratic congressmen had different approaches to the use of their images, but all selected similar mechanisms by which to record and promote events. Second, many of the people documented in the photographs are still alive and have different views about the work archivists are performing on the collections. Some are afraid of how images will affect their public profile; congressman's families attempt to guard the legacy of a public career through careful selection of the images they make public. These mechanisms are illustrated by the Senator Lee Metcalf Photograph and Film Collections Project.

Project Description

In 2012 the Montana Historical Society (MHS) Research Center received a Council on Library and Information Resources (CLIR) Cataloging Hidden Special Collections and Archives grant for the Lee Metcalf Photograph and Film Collections Project. Senator Metcalf was the longest serving acting president pro tempore of the U.S. Senate, serving from June 1963 to January 1978. He was a member of the U.S. House of Representatives from 1953 to 1960, elected as a U.S. senator in 1961, and served until his death in January 1978. He was one of the most important Democratic senators of the mid-twentieth century. Yet little is known about him because he was a very private

person and few authors published anything about his work. At the time of Metcalf's death, MHS already had collected 302.6 linear feet of his office papers, which were donated by the senator in the 1970s. However, Metcalf left no personal papers and has no living children to share their memories. To understand his work, life, and papers more fully, his visual record needed to be properly and completely identified.

Early Project Activities

The first year of the Metcalf Project included researching, sorting, identifying, organizing, describing, and preserving the original 4,454 (3,900 after processing) photographic prints, negatives, and slides documenting Metcalf's life and work. The second year involved cleaning, preserving, researching, identifying, and rehousing the 300-plus film reels. In processing the collections, I also aimed to discover new visual materials for celebrating the fiftieth anniversary of the Civil Rights Act, the Wilderness Act, the 1964 Montana Flood, and other events in which the senator had been involved. Another aim of the project was to intellectually tie this project's materials to the Lee Metcalf Papers MC 172 and other Metcalf materials held by MHS.

I spent the first six months researching, preserving, and organizing the photographic prints into groups by subject, date, and purpose. This included surveying the minimally processed MC 172 Lee Metcalf Papers, reading more than 6,000 newspaper articles on his life and work, visiting and researching online multiple archives and historical societies, and researching his use of photos and films during his congressional career. I also interviewed in person, by telephone, and by e-mail more than 55 people who were former friends, staff members, conservation activists, and state politicians with connections to Metcalf.

Results

After processing, the Lee Metcalf Photograph Collection (Lot 31) was housed in the MHS Photograph Archives. The collection is nine linear feet and comprises 3,900 images. When I began the project, I found that 25 percent to 30 percent of the images had already been identified in some fashion. However, not all the information was accurate, and there was no notation about who had identified many of the images. By the end of the project, I had identified at least two pieces of information for the photographs—subject matter, date, people, or location—for 85 percent to 90 percent of the images. The Lee Metcalf Photograph Collection currently constitutes one of the largest image collections of a twentieth-century Montana congressman.

Process

The existing arrangement of the photograph collection was based largely on the Lee Metcalf Papers' folder subjects and categories, which greatly expanded the utility of both collections in relation to one another. For example, Metcalf's papers contained two folders labeled Department of Agriculture: Forest Service, Bolle Reports (1970–1971). There were original prints of images used in the 1970 Bolle Report; therefore, I created a corresponding folder titled U.S. Senate—Congressional Projects: Department of Agriculture (Bolle Report) (1970). Such arrangement allows better coordination for researchers between the two collections. This type of organization of images and documents was the original intention of the Metcalf staff's filing system before the photographs and papers became separated over time.

I took time to research the senator's life because of his relative obscurity in the modern historical narrative. Only one book was written about him,

in 1965, and one book chapter, in 2003. Little biographical or historical material and no personal records exist. Little was known of his family history, family life, childhood, or college years. This was a significant challenge, as a number of photographs in the collection are from Metcalf's early life.

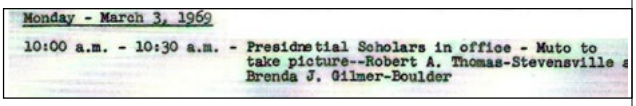
After combing through his papers, I compiled a list of people still alive who knew Metcalf. I scheduled times to talk with them about how he used photographs in his career, his personal life, and his political history. These talks and research helped narrow the research focus to themes and subject areas commonly identified in his photographs. With these subject areas identified, I discovered more information faster about the images in the collection. I found online resources that helped to identify elements of his history in Congress. I conducted informal interviews with former staff and friends, who allowed me to take notes for the project. These steps created a network of individuals who were introduced to the project, the importance of Metcalf's images, and the work of archivists to preserve history.

These interviews proved especially informative once I started working more closely with the photograph collection. For instance, among Metcalf's photographs I found a series of ten contact print sheets of various scenes from around 1972. They were not accompanied by matching photographic prints or negatives, and contained no indication of the images' purpose. I interviewed Karl Englund, son of Metcalf's long-time administrative assistant, Merrill "Brit" Englund. The son said that in 1970–1971, Metcalf and his staff started gearing up for his 1972 Senate reelection campaign. Metcalf did not like having new pictures taken of himself and in the 1960s often used photographs of himself from the 1940s and 1950s. Brit Englund, according to his son, forced the senator to go to new places, attend events, and meet with individuals in order to have new photographs to use in the 1972 campaign. These contact sheets are the images made during those photo shoots, and they show Metcalf's markings, indicating which images he liked the most for use in the campaign. I later found those images published in newspaper campaign advertisements.

During my research I learned much about how Metcalf used his photographs for political messages, personal relations, public relations, and influencing legislation. Metcalf's office managed publication of the Montana congressional delegation's newsletter, *A Montanan's Washington Notebook*, which began as a Democrat-focused publication, started in 1956. The publication did not use large numbers of photographs until Metcalf became senator and took over the publication in 1961. Photographs were published in the newsletter from 1960 to 1977. Metcalf retained photographs he had taken on Capitol Hill in a photo file in his office for use in this newsletter. Many of the photos I worked on in the Metcalf collection were used in the newsletter.

Senator Metcalf shot a series of television films called "Report from Washington" (1963–1965)—later called "Washington Report" (1965–1967)—in which he used photographs for visual evidence related to legislation, public programs, and federally funded Montana projects on which he worked. These films help date the photographic images, and the photographs explain the subject content of Metcalf's films (which were mostly unidentified). Metcalf used those same photographs in his television campaign commercials and in published handouts and newspapers during his Senate campaigns of 1960, 1966, and 1972. Information on the senator's use of images in publications, campaigns, and television all came from records located in the Lee Metcalf Papers at MHS.

Metcalf's papers also provided information that helped confirm the descriptions and dates of his photographs. For instance, to confirm the date of a Senate Democratic Photograph Studio negative of presidential scholars Robert Thomas and Brenda Gilmer visiting Lee Metcalf, I checked Senator Metcalf's daily schedules for 1963 to 1974 (see below for more information on the studio). The original Senate negative sleeve in which the negative was housed was dated March 3, 1969. I did not know if this was the date the Senate photographer's staff filed the negative, or the date the image was taken.


Monday - March 3, 1969

10:00 a.m. - 10:30 a.m. - Presidnetial Scholars in office - Muto to take picture--Robert A. Thomas-Stevensville
Brenda J. Gilmer-Boulder

Fig. 1: Schedules, Lee Metcalf Papers, MC 172, Box 426, Folders 2-3

I found the above notation (Figure 1) in Metcalf's schedule, and the description and dates matched the negative's sleeve for the date of the students' visit, March 3, 1969.

Metcalf's office regularly published photographs of the senator visiting constituents, talking with other politicians, testifying in committees, hitting the campaign trail, and attending Montana events and dedications. The *Billings Gazette, Kalispell Daily Inter Lake, Hungry Horse News, Helena Independent Record,* and the *Butte Montana Standard* were the newspapers that most commonly published Metcalf's photographs. The photos, intended as visual press releases, contained captions created by Metcalf's staff. Staff typed descriptions on carbon paper that were taped to the back of photos sent to the newspapers. Most of the time, newspapers published these captions and image descriptions word-for-word.

Online Tools and Sources

As part of my strategy for identifying Senator Metcalf's photographs, I used various online tools, publications, and websites that provided information for matching or explaining the subject content. One such website is the American Presidency Project, created and maintained by the University of California at Santa Barbara. It provides online, searchable text of presidential speeches, events attended, presidential dedications, and political party platforms. This was extremely useful when looking for subject content, the purpose of presidential visits on specific days, or special events attended by the senator. HathiTrust was very useful for locating online committee hearing transcripts, which provide lists of congressmen present, dates, staff members, and names of those testifying. I used this information to confirm the identity of photographs with notations about committee hearings, or about the people present at those hearings.

I used Wikipedia for comparative photographs of national politicians. Wikipedia provides lists by session for all U.S. congressmen. As an example, the Wikipedia page for the 89th Congress (January 3, 1965–January 3, 1967) gives names by state for all senators and representatives, provides sample photographs of the leadership, and notes mid-session replacements. Such information for congressmen is difficult to find in published histories on federal politics after the 1940s, and proved extremely useful (despite concerns about Wikipedia's reliability) in narrowing search parameters to identify specific congressmen in photographs with Metcalf.

Other archives and presidential library digital collections are also very helpful in finding images of congressmen from different time periods, at various events, and in different styles of dress. Over the past five years, presidential

libraries have increased the number of official White House and congressional photographs of presidents hosted on their websites. There are few places—other than presidential photograph collections—where the descriptions and dates for photographs are so detailed and well documented. For congressmen who are a little more obscure than, say, Hubert H. Humphrey, presidential library photograph collections are often the only online source for finding representative images of those congressmen.

At present, large numbers of online photographs are available for presidents Harry S. Truman through Gerald R. Ford. The John F. Kennedy Presidential Library and Museum currently has one of the best online collections of presidential photographs. Although not having individual images identified, the Gerald R. Ford Presidential Library and Museum's Digital Library has scanned and made available online descriptions for photographs by sets of images in contact sheets and rolls of film taken at the same time, through a monthly calendar browsing option. If you know a date for an image of a president but do not know the event or names of those in the photograph, presidential library image collections are the best bet for finding image descriptions.

With the greater availability of digitized twentieth-century newspapers via subscription databases, it is easier to find events and photographs, and to confirm information about individuals for congressional photograph collections. My greatest asset on this project was NewspaperArchive, which contains full-text, searchable, OCR (optical character recognition) newspapers dating from the 1860s to the 1990s, with summaries of search results. NewspaperArchive happened to have all but two of Montana's major newspapers.

NewspaperArchive provides chronological searching options to narrow search dates. For example, I needed to date a Metcalf photograph taken sometime in 1963. In NewspaperArchive, I limited the search dates to January 2, 1963–December 31, 1963, and used the search results to indicate whether a photograph was present (for this example, it helped me identify the image Lot 31 B16/7.01). An archivist can find photographs in the digital newspaper articles by finding in the search summaries such statements as "Metcalf visited with . . . last Wednesday" or "Metcalf shown discussing." These terms indicate a visual element in the article, and will likely have a photograph of the topic mentioned in the search summaries. One has to use multiple subject term combinations to locate photographs or information in the database, but the results are stunning.

I also used NewspaperArchive to find period photographs of people thought to be in photographs to confirm the identity and time period. For example, in a photograph taken of the Montana delegates to the 1958 National 4-H Conference in Washington, D.C., I could not identify the 4-H students individually, but I knew their names collectively from typed copy attached to the photograph. Through newspaper searching, I stumbled upon a March 7, 1958 *Billings Gazette* article containing photographs of all four 4-H delegates. I used the newspaper photographs to match the students' faces with the students in the Metcalf image.

In 2008, the U.S. Senate Historical Office sent 962 negatives to the MHS Photograph Archives of Senator Lee Metcalf. The office had found in a U.S. Capitol building basement the negatives for dozens of senators, identified only by senator last name written on the sleeves. The office had no idea who took these negatives, and there were no descriptions. In 2008, the Senate Historical Office began repatriating the negatives to cultural heritage institutions that possessed a

senator's original papers or photograph collections. The sleeves of the negatives sent to us had notations about print sizes and dates, though it was not known if these were the dates the photograph was taken, the negative was printed, or the negative was filed. The only recorded Senate photograph studio was the Republican Party's studio operated by Arthur E. Scott. My question was simple: If the Republican Party had its own photographer(s), did the Democrats, too?

Using Senator Metcalf's papers, interviews with former staff members, and notations on and letters attached to photographs, I identified Al Muto as the photographer for some of the Senate negatives. The 1964 Democratic National Convention published record reported that beginning in the early 1960s, the Democratic Senatorial Campaign Committee hired Muto to take photographs of Democratic senators. President Lyndon B. Johnson had used Muto as a personal photographer since the early 1950s and continued to do so into his presidency. An internally created House lobbying bloc for liberal legislation and public policy, the Democratic Study Group (DSG) in Congress began hiring Muto after he was laid off from the Associated Press in the late 1950s and the DSG was formally incorporated in 1959. Then-U.S. Representative Metcalf, a cofounder of the DSG, began using Muto regularly in early 1961.

Muto's title in the early 1960s was "Photographer —Democratic Senatorial Campaign Committee," according to notes found in Senator Metcalf's papers. There is no record of what Muto's title was through the remainder of his time as a Senate photographer. From 1962 to 1972, the Senate Democratic Photograph Studio operated in the U.S. Capitol Building. (I created this studio name to parallel the similar name later given to the Republican Senate studio.) Eventually, the studio was operated by public funds and was no longer supported by the Democratic Campaign Committee. In 1972 the Democratic and Republican studios were combined, but the studio continued employing separate party photographers. Brothers Al and Frank Muto, long-time accomplished press photographers, were the two primary photographers in the Democratic studio until 1975.

When Metcalf's staff needed copy prints from negatives held by the Democratic studio or by Metcalf's office, they sent a letter to the Mutos or Dev O'Neill (the House Democratic photographer) requesting reprinting or resizing services, as seen in the following note from Senator Metcalf to Al Muto (Metcalf 1967): "The Democratic Study Group, of which I [Lee Metcalf] am an alumnus, would like two 8 x 10 glossies of each of the enclosed negatives. I'll appreciate a separate bill, because DSG is going to pay it." The Democratic studio kept most of the original negatives for the images they took, and sent prints to each senator per the number requested (Figures 2 and 3). By using visual markers in the photographs and the corresponding dates printed on the back of prints (dates usually marked in blue ink) created by the Senate studio, I identified and paired 368 prints with their corresponding Democratic studio negatives (Series 24 of Lot 31 Lee Metcalf Photograph Collection).

After talking with former Metcalf staff and interns, I learned the senator had specific places in his office where he liked photographs to be taken for certain occasions, or with certain people. For instance, he had a framed Native American headdress on the wall just inside the door to his personal office—a gift from a Montana tribe. He liked to have pictures taken in front of it with staff members and summer interns. If I had an unidentified photograph of a young person with

Metcalf in his office and he or she was standing in front of this framed headdress, I could reasonably assume the person was a staff member or intern. This information helped limit my search for the name of the staff person in newspaper articles and Metcalf's papers.

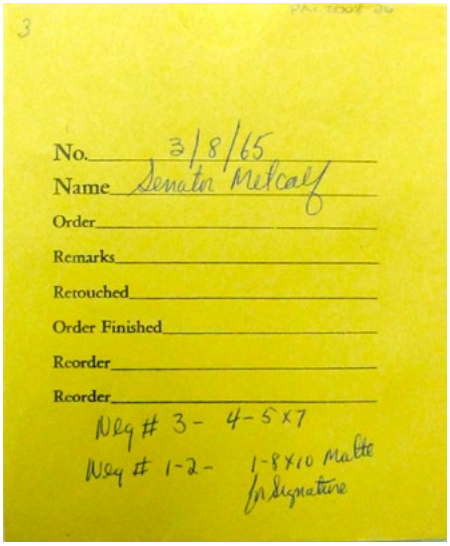


Fig. 2: Senate negative sleeve for images Lot 31 B16/15.01-.03. Note markings for print sizes based on the negatives

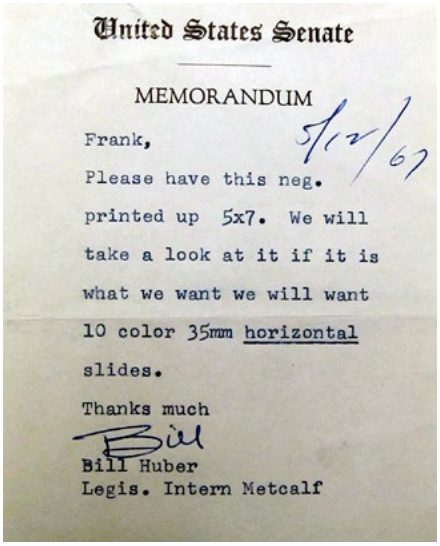


Fig. 3: Note attached to Senate negative sleeve, giving Senate Democratic Photo Studio instructions about image reproductions, May 12, 1967—here addressed to Frank Muto.

Metcalf also liked to be photographed with visiting constituents while sitting on the leather couch or armchair in his office. Again, for a photograph matching this type of scene, I narrowed my search to visiting constituents whose names appeared on Metcalf's office schedules, newspaper articles, and thank-you letters. The Senate Democratic Photograph Studio photographers also had a specific look and style of photograph for different subject matter that they preferred to use for different senators. Studying Metcalf's senate photographs taught me ways in which the Muto brothers liked to portray these public representatives. Metcalf's former staff members relayed to me that all of the Democratic senators in the 1960s they knew also had specific places in their offices for taking photographs.

A final note is necessary about the Senate Democratic Photograph Studio. Because archivists and historians have not been familiar with its operation or existence of the Democratic photographs, many archives that received negatives from the U.S. Senate Historical Office have not known who took the images. The images, therefore, remain limited for use by researchers and historians, because of questions about copyright and creator. Because the Senate Democratic Photograph Studio was paid for from public funds, all of the images taken by the photographers are in the public domain. For MHS, this means having 962 images from Metcalf's Senate negatives that can instantly be made available for publication. In other parts of Metcalf's collection, the photographs are unknown or were taken by a newspaper photographer, and are still copyrighted.

In closing, identifying and describing twentieth-century congressional image collections is becoming easier and faster, as digital content has been added online nationwide. Online resources,

such as EAD-based finding aids for other congressional photograph collections and digitized newspapers, are helping archivists better coordinate descriptive information for a congressman's photographs. A 1989 archival leaflet on processing congressional papers recommended the following for photographs: "Member's audiovisual material will probably contain duplicates which are unidentified, negatives with no matching positives, and positives with no matching negatives... Often, the identification of photographs is left to researchers due to limitations on processing resources" (Boccaccio and Carmicheal 1989).

Once considered too time-consuming to process, congressional photograph collections can now be processed using more advanced research techniques. These include capturing the memories of a congressman's office staff, using digital databases, and using visual clues in images without a great amount of labor. In our visually dominated age, it is vital that archivists identify and describe a congressman's photographs in as much detail as possible, so an institution and its users can benefit from information and content that corresponds with a congressman's use and creation of his documents.

The descriptions of a congressman's photographs are often the only record of his involvement with particular people, projects, and committees. Using approaches similar to those employed in the Lee Metcalf Project, archivists who process congressional photograph collections no longer need face drudgery. Nor do the photographs have to be set aside for lack of information about their content. The Lee Metcalf Project serves as an example of the cultural and research value that a congressman's images can bring to the historical record, once strategies are developed with modern tools to accelerate the processing and description of congressional photographs.

References

Boccaccio, Mary and David W. Carmicheal. 1989. "Processing Congressional Collections," Technical Leaflet Series Number 4, Mid-Atlantic Regional Archives Conference, 7-8. Available at http://www.marac.info/assets/documents/marac_technical_leaflet_4.pdf.

Metcalf, Lee. 1967. Letter to Al Muto, August 4, Lee Metcalf Papers, MC 172, Montana Historical Society Research Center, Archives, Helena, Montana.

Websites Noted

American Presidency Project: http://www.presidency.ucsb.edu

HathiTrust: https://www.hathitrust.org

Audiovisual Collections

Accelerating Exposure of Audiovisual Collections: What's Next?

Karen Cariani, Director WGBH Media Library and Archives, WGBH Educational Foundation

Sadie Roosa, Production Assistant, WGBH Educational Foundation

Jack Brighton, Director of New Media and Innovation, Illinois Public Media, University of Illinois

Brian Graney, Archivist and Head of Public and Technology Services, Black Film Center/Archive, Indiana University

Abstract

The paper presents issues of exposing large audiovisual collections online with examples from Hidden Collections initiatives. User interest in the exposure of digitized objects must be constantly weighed against the cost of preparation, management, and storage.

The Boston Local TV News Digital Library project is a core example. As a collaborative project, it has been a signal success, exposing about 50,000 individual film and video news stories created between 1960–2000. The four institutional partners addressed issues faced by libraries and other institutions with large moving image collections: how to effectively and economically expose large collections for better access, solve rights issues that may affect online access and use, and build awareness of and use for educational needs. The project showed that exposing minimal item-level descriptive data increases awareness and discoverability of collections.

The paper is enhanced by examples from other moving image Hidden Collections projects, with particular attention to overcoming technology barriers, funding structures, and staffing limitations. It will focus on the importance of sharing metadata standards for audiovisual works, in particular the use of PBCore, and will report from the PBCore user community. Lastly, presenters address the effectiveness of metadata workshops in a period of transition.

▶ Exposing Large Collections of Media

Karen Cariani, Director WGBH Media Library and Archives

The challenge for analog audiovisual collections in the digital networked age is their relative invisibility compared with numerous online resources that can be easily searched and viewed through public search engines. If an archive record does not appear in a search, it may as well not exist. Any plan for a sustainable online archive must take into consideration this extremely competitive environment for the public's attention. Because of the volume of individual surviving TV news collections, in particular, and each collection containing millions of feet of film and tens of thousands of stories, it is not economically feasible to catalog and digitize the full extent of each collection to expose each archive's content. The solution is not to select narrow topics and abandon the volume of the holdings to dark archives, thereby restricting future use; instead, the challenge is to illuminate the breadth of the collections by determining the information necessary for research within entire collections. The Boston Local TV News Digital Library project demonstrated and tested how to organize widely disparate news stories from separate collections, expose

the full range of content from 1960-2000, and contextualize the stories for public access, educational, and scholarly use.

Despite rapid advances in the world of digital libraries, moving images are usually the last to benefit from effective standards-based work—and television news is further deferred, offered in small topical selections, because of the size and technological challenges of large existing collections. Large media collections are hidden because they are on inaccessible formats, such as film and videotape formats of 2", 1", and ¾", which are all now considered obsolete. The equipment to watch the content stored on analog formats is disappearing. It is no longer being manufactured, and existing equipment is breaking down and being discarded. In addition, because of the deteriorating magnetic signal on the tapes and disintegration of the film, playing a tape to determine its content may be the last time the tape can ever be seen, or the last time signal on the tape can be captured to make a copy.

So how can we discover what is on these tapes and film and in these media collections? How can we let scholars and researchers know about this content? The dilemma for media material is that content on obsolete formats needs to be digitized to be accessible and then it needs to be described and cataloged to be identified and discovered. Videos are opaque, so the content is unknown without playing them. Titles alone do not necessarily tell you the content. You cannot scan tape or film for searchable data. Most research and searching on the web occurs through text search engines. With no robust descriptive metadata for the media, having many hours of digital content available on the web does not make it discoverable. To improve findability and increase use and access for scholars and researchers, the content needs better descriptive information that can be indexed by search engines.

There are a number of attributes and data that could be gathered without having to fully watch or listen to a program and that would greatly enhance accessibility. Collecting, posting, and exposing a minimal amount of data can be helpful. Minimal information that can be gathered from the physical item—on tape labels and film cans—is better than nothing. Verifying or documenting the correct title(s) and title types from the title card and dates associated with the content, identifying primary creators and contributors by just watching or listening to the credits, or only documenting the statement in the copyright notice would help with rights assessment. Additionally, noting the type or genre of program that might not be obvious from the title—such as dramatic performance, news, instructional, public affairs, documentary—would be extremely helpful and would aid researchers in their search for material. Having this data as part of the archive will greatly enhance discoverability of the collection for scholarly and humanities work.

To increase access for scholars at WGBH, we have been testing the value of publishing minimal data for our media archive. We launched our archive "catalog" online. It has minimal data, often just a title, in the hope that it is at least now discoverable. When possible, we have published transcripts of interviews, which are exposed to search engines, further enhancing discoverability. We have worked with our users to add information to our minimal records, enhancing the catalog even more. But most of these records are titles of full programs that may have been referenced somewhere at some time. Our news collection posed a more complex challenge. The first was its volume, and second was the variety of topics, which covered more than 10 years of daily news in Boston. Daily news does not have a

succinct title other than on a collection level that represents 10 years and thousands of items.

Television news provides a daily chronicle of our communities since the late 1940s, especially through local TV news. Across the United States, collections of local television news on film and defunct analog video occupy shelves at TV and radio stations, colleges and universities, public and specialty libraries, local and state historical societies, and archives and museums. These records of local TV news represent an as-yet largely untapped resource for study and use. Despite their potential value to the public and school audiences, few other records of contemporary culture have been as unavailable as twentieth-century television news. In its *Television and Video Preservation 1997* report, the Library of Congress stated, "The most devastating losses have already occurred among news film and videotape files of local television stations across the United States." Compounding the access problem is the volume of those individual surviving collections and the overwhelming cost of making them accessible and useful. To preserve and enhance the value of audiovisual news collections, they must be organized, described, put into context, and made known to users.

Daily local TV news has been created and produced for more than 50 years—30-60 minutes, every day. It is a lot of content, a lot of tape, a lot of hours to view to catalog. For time-based media, the content needs to be seen or heard in real time for a human to be able to describe it and tag it. As an example, the 40,000 hours we have digitized for our American Archive of Public Broadcasting project would take one person working 8 hours, 7 days a week, 30 years to catalog. The resources necessary to fully describe and catalog hours of programming are overwhelming and unattainable.

The few local TV collections offered in digital libraries are relatively small-scale and topically focused, chiefly on the civil rights era. These and other projects offer important documentation, yet they do not expose entire collections. The Boston TV News Digital Library: 1960-2000 sought to expose entire news collections. Developed by the WGBH Media Library and Archives, with Northeast Historic Film, Cambridge Community Television, and the Boston Public Library, the project yielded the first online resource offering a city's commercial, noncommercial, and community cable TV news heritage to educators and the public. The purpose of the collaboration was to use, test, and demonstrate open source tools to assist custodians of similar resources, while creating an online library offering 40 years of urban moving image materials, resulting in approximately 70,000 news records. Each of the four partners brought a television collection to the digital library ranging from hundreds of analog videotapes to four million feet of 16mm film.

Online access to TV news collections poses complicated rights issues. Where users once expected to make research visits to collections, global exposure via the web changes the range of potential legal risks. Consequently, an examination of rights issues for online access to local TV news in public collections was also greatly needed. The ephemeral nature of TV news—it airs and never will be seen again—made releases and rights issues less risky for broadcasters, so releases and permissions were generally not secured. As part of this project, we worked with our legal team and the Berkman Center at Harvard University to answer some questions about allowing online access to the news collection. We wanted to develop a guide for archivists that would help them identify and think about the risks, to inform decisions about what they can do with their own collections.

As a collaborative project, it has been a signal success, exposing around 50,000 individual film and video news stories from between 1960 and 2000. The four institutional partners addressed issues faced by libraries and other institutions with large moving image collections: how to effectively and economically expose large collections for better access, solve rights issues that may effect online access and use, and build awareness and use for educational needs. The project showed that exposing minimal item-level descriptive data increases awareness and discoverability of collections.

An Example of Exposing Large Media Collections: The Boston Local TV News Project

Sadie Roosa, Project Assistant, WGBH Media Library and Archives

Using the Boston TV News Digital Library as an example of a large audiovisual collection, we can explore some of the obstacles and challenges one might face while processing, cataloging, and making accessible any similar collection. We can also use the project's successes to guide the work needing to be done on other large audiovisual collections.

At the beginning of the project, what we had was essentially four separate audiovisual collections. Each collection was too large for any of the individual holding institutions to handle as part of their everyday workflow. The Boston Public Library (BPL) holds the WHDH Channel Five collection containing more than 23,000 assets, and Northeast Historic Film (NHF) holds the WCVB Film Collection containing more than 15,000 assets. Both of these collections were completely hidden from users, with no electronic inventories or catalogs. Additionally, WGBH holds the Ten O'Clock News collection with nearly 8,000 assets, and Cambridge Community Television (CCTV) holds its own collection of nearly 5,000 assets. Both of these collections were partially cataloged in Filemaker databases but were only available onsite at the holding institution. The collections were not well advertised as a resource to potential researchers on the collection level, let alone the item level. We decided that the best approach to handling such large collections would be to bring them together, sharing our technical and personnel resources.

With 51,000 assets, 38,000 of which did not have any form of electronic record, we needed to figure out how to make these collections useful. The initial focus of our energy was on the two completely hidden collections. We designed PBCore compliant Filemaker databases to capture and manage the new electronic records. We also took the existing Ten O'Clock News and CCTV Filemaker databases and mapped the existing fields to PBCore. While each collection had particularities in what data were available and relevant, by mapping it all to the same standard, we ensured that these disparate collections could exist as one cohesive collection that could be searched as a whole. Bringing the collections together provided many benefits, including providing more context for each individual collection and allowing for easier comparison between different styles and eras of news reporting.

Rather than beginning by fully cataloging each item in the two hidden collections, which would take considerable time and require the use of playback devices not easily available, we created item-level records based on what was available to us without any playback. Project assistants and library science interns transcribed a set of index cards that had served as a basic catalog for the WHDH collection. NHF's archivist created item-level electronic records for each film reel, recording the can labels and the slugs written on

the leaders of the films. By the end of the transcription process, many of the electronic records had dates, and some had significant keywords and names of places and people, but they were remarkably sparse overall.

While we knew the data were far from perfect, and we expected many hours of grooming and normalization ahead, we decided not to wait to publish the records online. There is no clearly defined standard for what data are "good enough" for users to be able to search. We wanted to push the limits and see if the barest item-level records, essentially just a date and a keyword or two, a full sentence if we were lucky, would be enough to catch anybody's interest. We discovered that for some users this really was enough. If they knew exactly what they were looking for, they could narrow down our collection based on date and then scan the sparse descriptions for the story they needed.

As not all researchers could conduct such narrow searches, we continued to improve the records by adding subject, name, and location headings that we authorized in a controlled vocabulary consisting of many Library of Congress subject headings and name authority files, as well as many locally authorized headings. At this point we still had not played back any of the items, so all of this cataloging was being done based on the sparse descriptions available by just looking at the index cards, film cans, tape cases, and other labels. Because we were not viewing or listening to any of the items yet, we were able to move through this minimal cataloging at a rapid pace. Using this method, we were able to make all of the records slightly more searchable much more quickly than if we had fully cataloged even one of the individual collections, let alone the combined collection of all four. As we improved the records, we refreshed the data on the website, making it more searchable and user-friendly, not only because of the normalized names and topics, but also because the website links the headings, which allows users to go from a record with a specific heading to a list of all the records with that heading.

As we worked on this minimal cataloging, we also had the project staff and interns pick out especially interesting-sounding news stories that they came across. Based on the date and sparse description, they would research newspaper databases and other resources to determine what the news story was likely about. They would then write up a post for the project blog. By featuring these more in-depth looks at our content on the website, we provided richer descriptions for at least a few items, and made our collection more visible and attractive to potential users. Users found our blog posts in search engine results, read the posts containing richer descriptions, and then searched all of the sparser records for additional content of interest.

We decided that because of the fragile state of many of the items, especially the ¾-inch videotapes, and the scarcity of easily accessible playback equipment, especially for the BPL's film, we would wait to view and more fully catalog any of the content until it had been digitized. The risk we took in choosing to digitize before viewing was that we might end up digitizing content that had been mislabeled or was actually less significant than the labels indicated. However, we decided that issues with playing the analog formats and the amount of time it would take to view items before digitizing were enough of a deterrent. We also were not comfortable making the blanket judgment that a user would find certain content more valuable than other content.

It was clear from the beginning of the project that we would never secure the funding and other resources to digitize the entirety of any of these collections. Rather than deciding internally which items would receive priority for digitization, we conducted an experiment to see if we could get enough public attention to have our users decide the priority through a process of user voting at the item level. With only the sparse descriptions and the blog posts highlighting a few of the items, we had a total of 1,184 votes from the public, spread over 825 items. Based on these votes, and some input from our advisors, we began digitizing the prioritized items.

Once the items were digitized, they were much easier to view for cataloging. Project staff and interns viewed them, wrote short descriptions, and added subject, name, and location authority headings. Because we viewed the electronic copies of the items, we also manipulated the content to make cataloging easier. For example, we played content at two or four times the regular speed, so that speech was still understandable, but the viewing process took half or a quarter of the time. We refreshed the website with these richer descriptions and additional headings, and added the digitized media to the records so users could stream the videos in the site.

When the project ended, we had 1,593 items digitized and streaming on the site, with enhanced descriptions and headings. The other 49,000 records still only have the sparse descriptions and minimal headings, but they are available to search along with the records for the digitized content. We implemented technology that allows users viewing the records to request digitization on an item-by-item basis. Without additional grant funding to support more digitization, we require users to cover the costs. Overall, we are pleased that all of the records, not just the ones we were able to view and fully catalog, are searchable by users and available to view in some manner.

Accelerating Exposure of Audio Visual Collections, What's Next? *Confessions of an Accidental Archivist*

Jack Brighton, Director of New Media & Innovation, Illinois Public Media | College of Media

"My main theme is the extension of the nervous system in the electric age…"

Marshall McLuhan - *Letters of Marshall McLuhan* (1987), 300

The ability to record and play back moving images and sound has been part of the human technology repertoire for barely more than 100 years. Electricity was not commonplace anywhere until the turn of the twentieth century. The combination of electricity and media technology brought an acceleration of change in our capacity to see and hear, paradoxically both extending and annihilating distance and time. It also allowed us to create new forms of time-based narrative arts, and increasingly intervenes in every economic, political, and social interaction.

And most of this happened before the Internet.

If "all media are extensions of some human faculty," as McLuhan has it, then archival media are an extension of our memory and imagination. For this to actually mean anything, preservation of moving images and sound is necessary but not sufficient: its value in human culture depends on ready access.

I began my professional life in public broadcasting as a journalist and creator of electronic media. It is toward the realization of "persistent access" to the vast store of our multimedia memory that my career has increasingly become focused.

Fortunately We Have No Resources

Exposure to archival practice came from my struggles as a media producer in the emerging digital age. I began designing websites with streaming and downloadable multimedia in 1997, and quickly realized that without an archival plan the situation was becoming hopeless. I saw how quickly technology was changing, and suspected that the media we published on the web at that time would be unplayable within a few years. I placed bets all over the board by publishing RealMedia, QuickTime, Windows Media, and Flash but I assumed all these formats would sooner or later be obsolete. I guessed wrong how quickly that would happen. But I got one or two things right: I saved the original physical formats and high-resolution digital files derived from the masters. And I began to maintain a database of everything I saved.

Fortunately nobody saw me doing this. As a lowly producer, I had no budget. Therefore, no one could cut my budget or tell me to stop. I was the guy running the station website in his "spare time," and no one complained so long as I did the magic. But I soon noticed something truly disturbing: the problem of persistent media was growing more difficult, more complex, and just plain larger every day. And the rate of change was accelerating.

And that is not counting our shelves and desk drawers everywhere filled with analog media in fragile formats with barely a peeling label to tell us what they are.

Toward Content Management and a Life of Data

You build a website with content and metadata. The presentation is just the access layer. A web page is accessed by humans by way of a browser that parses the HTML and presents a view we can understand and use. The same web page is accessed by machines (such as web crawlers and other applications) by virtue of well-structured data they can parse and reuse. As I began to understand these things, I learned from my colleagues at other public media organizations that many of us were becoming deeply invested in technical solutions that could manage this data, and present it to humans and machines, at the cost of a very large contract and ongoing vendor dependency.

The solution I pursued seemed half-baked by comparison with these expensive, proprietary systems: a simple MySQL database tied to a file system. With this as a data core, we can build applications that speak the language of the Internet. We can present web pages to humans; RSS/podcast feeds to iTunes; Dublin Core and PBCore records to library and archival services. The most difficult challenge is getting good data to begin with. So that is what I focused on, and sought external funding for with some success.

Today we openly share well-structured data with a growing number of public institutions and archival services, including the American Archive, the Pop Up Archive, Collective Access, and the Internet Archive. We provide open access to high-resolution digital media files and all the metadata we can capture and share. In some cases our partners add to and enhance the existing metadata. For example, the Pop Up Archive provides a speech-to-text transcription service that feeds back into our core data. Most of this happens automatically using PBCore as an exchange format between systems.

The way the automation works is as follows:

- Content producers create media files, such as radio news stories, video productions, interview segments, etc. Most of these assets are finished products for broadcast or online distribution. They also might be products of oral history interviews, youth media projects, and special collections based on election cycles and other topic-based productions.
- Producers export media files to a watch folder, from which scripts route the files to software that encodes them to web-friendly formats. Source media files are then moved to network storage where they become subject to our preservation protocols.
- Producers upload the web-friendly media files to our website content management system, which stores the files in an NFS mount addressable by http. While uploading these files, producers add a variety of (mandatory) descriptive metadata. The location of the files and all other metadata then resides in our MySQL database.
- Our website content management system provides an HTML front end for public access to all uploaded media files and associated metadata. The assets are available for browsing and search on the Illinois Public Media website, and exposed to social media services via Open Graph and other metadata formats.
- The website CMS also provides full access to all content and metadata in the form of PBCore records using the pbcoreCollection element for exchange of multiple records as a "feed."
- Services like the American Archive and the Pop Up Archive consume Illinois Public Media's PBCore feed for automated ingestion and processing of the media essence files and associated metadata.
- Where external services add to our existing metadata, e.g., text transcripts and SRT files from the Pop Up Archive, the enhanced metadata is added to the MySQL store of data at Illinois Public Media where it is used for search optimization, video captioning, and progressive enhancement of our website user experience.

I would emphasize one key feature of this system: it requires no additional effort on the part of producers beyond publishing media content on the Illinois Public Media website. The system itself could be described as a system of shareable metadata and media files. It just happens to also produce a website.

At Illinois Public Media we do not aspire to fully solve the archival problem. We simply try to be an awesome version of the best data source ever. I think it is important to keep role and scope in mind as more public media and cultural heritage institutions take on the challenge of preserving and creating access to their media collections. Illinois Public Media will serve as the source and authoritative voice for our collections, and if we cannot stand up our own trusted repository it is within our mission to share it with other public institutions that can.

If You Do Not Know What You Have, You Do Not Really Have It

The complexities of persistent access in the age of rapid change in media technology can seem overwhelming for public media and other cultural heritage institutions. The situation is more dire with the deterioration of our legacy analog and physical media collections. We are racing against both time and scope, and for some significant portion of our audiovisual heritage, we will lose this race.

If we have any remaining arguments about which system or tool is best for managing media content and metadata, we can enjoy that conversation over drinks. The most important thing is to get the data about what we have, and put it in structured form so it can be accessed and shared. Media census projects, like those recently completed at Indiana University and the University of Illinois at Urbana-Champaign, are needed to begin answering the question of what there is to preserve. We can then marshal that data to prioritize the work of preservation and "persistent access."

Persistence of Vision

> "The written symbol extends infinitely, as regards time and space, the range within which one mind can communicate with another."
>
> Samuel Butler, *Life and Habit*, London: Trubner & Co, 1978

We live at a time when all previous forms of media are potentially hyperlinked and accessible. We added moving images and sound to the written symbol as means to extend perception, communication, knowledge, and imagination. It may seem lofty to claim that our media archives are to culture what memory is to the human brain. But to the extent that we have not yet embraced caring for these as means of extending our senses, our vision and reach are impaired.

The work still to be done is far from trivial, but my experience has been that it begins by focusing on the data. The technical means of handling data will change, but we can make the data accessible as systems and standards evolve. Small institutions like Illinois Public Media cannot do everything required for preservation, but we can serve as a node in a larger preservation ecosystem. In terms of preserving an aggregate of potentially all institutional media collections, resources can be allocated at different levels depending on roles and scope. The thing that has brought such rapid change, the Internet, is also the means of connecting the levels.

We may have limited resources, but fortunately we have an architecture of collaboration we're just now beginning to understand.

▶ Hidden Collections at Northeast Historic Film

Brian Graney, Archivist, Black Film Center/Archive, Indiana University

Northeast Historic Film (NHF), a regional moving image archive based in Bucksport, Maine, was among WGBH's partners on the Boston Local TV News project, and was at the same time exploring related questions about the exposure of large audiovisual collections over the course of two sequential projects funded in 2009 and 2010 through CLIR's Cataloging Hidden Special Collections and Archives program: Intellectual Access to Images of Work Life, 1916-1960; and Moving Images 1938-1940: Amateur Filmmakers Record the New York World's Fair and Its Period. Both projects were directed by Karan Sheldon, a co-founder of Northeast Historic Film.

The 2009 project Work Life described 50 unique hidden collections of nearly 800 amateur and industrial films documenting work and labor in the first half of the twentieth century "to help advance discovery and inclusion of moving images as primary source materials in the universe of traditional library and archives collections."

This mission had roots at NHF in cataloging efforts extending back to its 1995 Collections Guide, which made use of descriptive standards Archives, Personal Papers, and Manuscripts (APPM) and Archival Moving Image Materials (AMIM). More recently, the 2007 project, Finding and Using Moving Images Online, funded as a

National Endowment for the Humanities Digital Humanities Start-up Grant, demonstrated the value of treating films with integrity as primary source materials for the humanities by selecting and presenting archival moving images explicitly as unedited texts within a hierarchical structure rooted in provenance.

With the Work Life project, NHF continued this approach by formally adopting an efficient metadata model integrating provenance-based archival description of moving image collections with descriptive and technical metadata at the item level, based on the widely supported standards *Describing Archives: A Content Standard* (DACS) and Encoded Archival Description (EAD), along with the then-emerging audiovisual metadata standard, PBCore 1.3. NHF also worked with database developers to implement a custom profile of the open source CollectiveAccess platform, designed for web-enabled production and presentation of finding aids with embedded images and digital video; search and navigation across collections and items; and export of XML-based EAD and PBCore records. Elements of the open source profile developed for this project have since served in CollectiveAccess implementations at large and small institutions, including the Academy Film Archive and the Chicago Film Archives.

NHF's 2010 CLIR project set out to describe hidden collections of amateur film across three institutions, including NHF, the George Eastman House International Museum of Photography and Film, and the Queens Museum of Art. The emphasis on amateur film shot at the 1939-1940 New York World's Fair aimed to establish the subject within a broader research context of amateur filmmaking of the period. There was at the same time a strategic emphasis on promoting wider implementation of the descriptive practices and tools developed through the Work Life project and dissemination of related resources, tools, and strategies, primarily through a series of cataloging workshops held over the project year in Rochester, New York, with the L. Jeffrey Selznick School of Film Preservation; in Boston, Massachusetts, with the Northeast Document Conservation Center and the Simmons College School of Library and Information Science; and in Austin, Texas, with the Association of Moving Image Archivists.

There was throughout both projects an underlying effort to interrogate the practices and priorities at the archive in order to establish efficient workflows in areas of support for project cataloging—such as inspection, conservation, and reformatting—and in cataloging itself. In evaluating this, we returned regularly to discussions of what might constitute efficient processing for moving images, particularly for home movies and amateur films.

In the case of the World's Fair project, we resolved this question situationally within the thematic framework of the project, which privileged above all films of the 1939-1940 New York World's Fair, examined them alongside other amateur films of the 1938-1940 period, and established all of the films as individual items within the context of their individual collections, which, on the whole, were broader in their subjects and dates of creation.

As in the earlier Work Life and Moving Images Online projects, DACS-based collection level records were created across the board, meeting the standard's single-level minimum element requirements and enhancing records with some additional elements recommended for optimal description, including administrative and biographical histories of the creators, and geographic, subject, and name-based access points. Each

finding aid also included a full collection inventory, corresponding to the individual PBCore records, based on minimal Intellectual Content and Instantiation element sets adopted for the project.

Within the PBCore Intellectual Content class, we determined the level of description to employ by considering the relevance of each item to the project themes. New York World's Fair films received highly detailed shot-level descriptions; 1938-1940 films received rich item-level descriptions; and other reels within the project collections that fell beyond the scope of its principal themes received minimal item-level descriptions, often based solely on notes drawn from the film cans, leaders, or inventories in the collection files. Among this last group, we made exceptions to the minimal approach for films that were discovered to be of particular significance and potentially high user interest. Two examples of these are the mountaineering films of Charles and Oscar Houston, including film of the landmark first ascent of Alaska's Mount Foraker in 1934; and 1927 amateur footage shot by Simmons Brown of Charles Lindbergh and "The Spirit of St. Louis" at Orchard Beach, Maine.

PBCore Instantiation class elements, associated with each Intellectual Content record, were populated for all original film items and for video transfers derived from them. In the case of the thematically prioritized films, the archive's technical staff provided highly detailed inspection reports, including information about the stock types, date codes, editing structure, and distinctive aperture marks, which can help determine the camera used to shoot the film. While far from minimal in practice, this strategy was directly responsive to stated user interest in aspects of the material film artifacts not otherwise captured during the creation of digital surrogates or content description.

We configured the CollectiveAccess database modules in our profile to correspond to our metadata schema, and provided mapping schemas to the developer for generating valid EAD and PBCore XML documents for harvesting from the CollectiveAccess records.

Finally, as an important supplement to the metadata strategy, we embedded visual and audiovisual digital media in the records at different levels, again according to the thematic priorities established above. Collection-level displays featured single frame enlargements and brief streaming video excerpts considered to be representative of the collection; 1938-1940 item record displays were each illustrated with a frame enlargement, which appeared as a thumbnail in item-level search results; and the item record displays for New York World's Fair films featured multiple frame enlargements and full reel transfers in streaming video. While this fell outside the parameters of the cataloging project, our rationale here was that, while researchers could expect a reasonable level of recall on their searches based solely on textual descriptions and access points at the collection and item levels, their subsequent determinations of relevance—particularly for researchers working with home movies and amateur films—were often most easily narrowed by primarily visual means.

References

American Archive: http://www.cpb.org/features/americanarchive/

Boston Local TV News Digital Library project: http://bostonlocaltv.org/

Collective Access: http://www.collectiveaccess.org/

Internet Archive: https://archive.org

Pop Up Archive: https://www.popuparchive.com/

Putting Archival Audiovisual Media into Context: An Archival Approach to Processing Mixed-Media Manuscript Collections

Megan McShea, Archives of American Art, Smithsonian Institution

Abstract

The Archives of American Art has completed a three-year project, funded by the Council on Library and Information Resources (CLIR). The project focuses on collections documenting postmodern art movements that rely heavily on audiovisual documentation. The media-rich nature of these collections has allowed project staff to investigate ways to keep audiovisual media from being hidden within larger mixed-media collections in manuscript repositories. Through this project, the Archives has developed tools for processing mixed-media collections that integrate traditional archival processing procedures with best practices for archival audiovisual media. Our approach includes developing media-specific tools for surveying collections, defining levels of processing, collecting metrics to help in planning, and creating guidelines for arrangement strategies and for describing audiovisual media in EAD. This paper outlines some of the details of these tools and guidelines and how processing archivists can use them to ensure better physical and intellectual access to audiovisual materials.

Introduction

The Archives of American Art is a research unit of the Smithsonian Institution that has been collecting archival materials documenting the history of art in America since 1954. Like many organizations in the manuscript repository category, the Archives does not primarily collect audiovisual material, but does collect it in substantial quantity and variety, nonetheless. At latest count, the Archives has more than 15,000 analog and digital audiovisual objects in more than 800 of about 5,000 total collections. Our recordings exist in about 40 different audio, video, and motion picture film formats. About 90 percent of these materials are in analog audiovisual formats.

When we started to investigate the audiovisual holdings, the state of intellectual and physical access to them varied. In general, these holdings were poorly understood and had been inconsistently and often inaccurately documented and described over the years through several generations of staff. This was as true for collections considered processed[1] as for those that were unprocessed. As such, they constituted a sort of insidious, hidden backlog of inaccessible material within collections that were no longer considered part of the Archives' processing backlog.

To address this issue, a general one across the profession, the Archives undertook a three-year project to investigate methodologies for processing mixed-media collections. These are collections in which archival audiovisual documents and paper documents exist together.[2]

1 The Society of American Archivists' *Glossary of Archival and Records Terminology* defines archival processing as "the arrangement, description, and housing of archival materials for storage and use by patrons" (Pearce-Moses 2005).

2 More information on the project, Uncovering Hidden Audiovisual Media Documenting Postmodern Art, is available on the Archives of American Art's website at http://www.aaa.si.edu/collections/projects/clir. Benchmarks, ratings, rehousing guidelines, and description guidelines described here are all available in the project's technical documentation at http://www.aaa.si.edu/collections/documentation/av.

The primary goal of our Hidden Collections project was to create tools for processing archivists that would help integrate best practices for audiovisual material with best practices for traditional archival processing. The aim is to provide guidance that will enable arranging and describing mixed-media collections so that audiovisual material is just as accessible, both intellectually and physically, as other kinds of records in the collection, and to accomplish this via traditional processing workflows.

Project Goals: Tools and Guidelines

We created tools for planning processing, and guidelines for arrangement and description, that could be adopted by any archivist undertaking traditional processing of collections with an audiovisual component. The three planning tools include: (1) benchmarks defining what is expected of the processing archivist at different levels of processing; (2) metrics on processing rates and extent changes to understand how long media-specific tasks could be expected to take, and how much the size of collections changed when processed; and (3) a ratings system to help archivists assess the pre-processed state of media in collections and its needs.

In addition, guidelines developed during the project address arrangement and description of audiovisual material, including instructions for when and how to replace media housing and how to seat media in collection containers, detailed guidelines for description of audiovisual media in Encoded Archival Description (EAD), and some broad guidelines for intellectual arrangement of media in collections.

Planning Tools: Benchmarks

The benchmarks developed for this project were designed to make explicit what is expected of processing archivists at each level of processing, specifically for audiovisual media. At the Archives, levels of processing for all collections are defined as preliminary (tasks completed upon accession), minimal (the standard now for all but specially-funded processing projects), intermediate, and full.

For audiovisual media, preliminary processing involves a brief, minimal identification of media in the accession record for a collection, and a more detailed survey. When a new collection is accessioned that contains audiovisual media, the collecting archivist includes the extent, location, and general content of the media in the accession record. The audiovisual archivist then conducts a more detailed survey, documenting its condition, the condition of its housing, format characteristics such as recording speed and size, date, content description, and an assessment of the uniqueness of the media and its likely rights status. With this information, we document and can track the media and its current state, regardless of when the collection gets slated for further processing.

More detailed processing has three defined levels: minimal, intermediate, and full. Although it is always difficult to create rules that apply to all collections because each collection presents a unique case, a few rules were established to guide archivists in their decision making about audiovisual media in the collection. For instance, for minimal processing, rehousing of media in damaged or unsupportive housing is not required, nor is playback of poorly labeled media. Unlike intermediate or full processing, archivists processing at a minimal level can use the term "unidentified" to describe unlabeled media. They are encouraged to describe labeled media in the aggregate, and they do not have to include every detail available in their finding aids. Instead, they

can list a date range and higher-level description of a group as a whole. With these few established benchmarks, archivists working on minimal-level processing projects, which are typically large-scale and fast-paced, can note media accurately, associating it with related records, without slowing their pace. The benchmarks are meant to be flexible, however, so archivists can use their judgment to decide whether the media merits the extra effort required to provide more granular intellectual access or greater physical protection to collection material found in poor condition.

For intermediate and full processing, archivists are explicitly required to play poorly labeled media, if it is stable and if the playback equipment is available, in order to provide an adequate description. For film, they must inspect the leader and head of the film to discover titles and dates. They are required to rehouse media that is in substandard housing. Description also gets more complex at higher levels, and there are a range of enhancements archivists can consider to provide more granular and nuanced intellectual access.

Planning Tools: Metrics

Metrics were collected throughout the project to investigate how long processing tasks take for audiovisual media, the factors affecting those rates, and changes in overall extent of collections after processing. Data on extent changes showed that an increase in collection size is typical, with the median increase being 24 percent.

Tasks that were timed included surveying material, arranging and rehousing, and writing and entering the description in Archivists' Toolkit, the tool we use to author finding aids. Calculating the data collected, we found that audiovisual portions of collections took an average of 12.7 hours per linear foot to process, and the processing times for mixed-media collections as a whole averaged 15 hours per linear foot. These collections were all considered processed to the "full" level, that is, the highest level of processing.

Although the project collections represent a small sample, it is interesting to compare these numbers to the rates of processing for our collections that do not contain substantial audiovisual material (the latter average 18.1 hours per linear foot for full processing at the Archives). It was significant to discover that large audiovisual components of collections do not necessarily result in longer processing times. In fact, on average they seem to take less time than collections without media. And in the AV-rich collections, the audiovisual portion is taking less time than the non-AV portion. The assumption used to be the opposite. As we gather more metrics, including rates for minimal processing of mixed-media collections, they will surely provide more insights. Already, these preliminary numbers indicate that mixed-media collections need not be excluded from processing projects based on concerns about processing speed.

Planning Tools: Ratings

The metrics also helped identify factors that affected the rate of processing. As with any type of archival material, audiovisual media in unprocessed collections present a wide range of needs. Three factors emerged in the course of the project that seemed to most affect the rate of processing: the extent of rehousing needed, the extent of playback needed to identify content, and the extent of analysis needed to determine relationships among recordings and between recordings and other documents in the collection.

We created two rating scales based on these factors to help archivists assess what level of work will be required to process the audiovisual media in a collection. The housing rating is simply

a range—from a poor rating, indicating that all media need to be rehoused, to an excellent rating, meaning no media need to be rehoused. The audiovisual access rating combines an assessment of how much playback and analysis are needed to properly arrange and describe the media. If everything needs to be played and analyzed to determine the content and the relationships among records, the audiovisual media gets a poor access rating. If nothing needs to be played or analyzed, it gets an excellent rating.

For example, a collection might contain audiovisual media that is completely unlabeled and lacks corresponding paper documentation describing its contents. In this case, extensive playback will be required to arrange and describe it. If everything is well-labeled, seems to have been created in an orderly way, and is accompanied by transcripts or shot lists, it will have a higher access rating. The audiovisual access rating is also related to complexity. One series of interviews in the same format will be a lot easier to arrange than multiple media productions with many versions and production elements in multiple formats, where everything has been boxed together and needs to be sorted out.

Guidelines: Rehousing

Guidelines developed for this project are designed to assist processing archivists with the tasks of rehousing, arranging, and describing audiovisual media. The variety and complexity of archival audiovisual media, and of the repositories that collect it, make it difficult to provide guidelines that can apply to every possible circumstance in every setting. Guidelines written for this project were designed to summarize published standards (Adelstein 2009 and ISO 2000), and to apply those standards to common circumstances specific to mixed-media archival collections.

Rehousing guidelines address both housing and the physical orientation of audiovisual media in storage containers, particularly in cases where audiovisual material is found in containers with other types of records. Archival repositories have a variety of storage resources and therefore will have different approaches to housing and storing special formats. At the Archives of American Art, we keep audiovisual materials in mixed-media collections in their collection of origin, rather than remove them to a special format storage area. Currently, although the general collections storage areas are effectively climate-controlled, there is no storage alternative for materials that would benefit from a cool or cold climate. Without the ability to improve the storage climate for audiovisual media, there is no rationale for separating these materials from their collections of origin.

The rehousing guidelines created for this project are designed to spell out what is expected of processing archivists at the Archives regarding rehousing, and to help them improve the physical stability of audiovisual material for long-term storage as they physically arrange collections. Generally speaking, archivists should apply their knowledge of archival materials and their physical vulnerabilities to the housing of audiovisual media. This means getting rid of deteriorating, dirty, or unsupportive housing, and photocopying or scanning original housing to preserve the information it carries. Specific supplies, seating, and handling instructions vary by media format.

Guidelines: Arrangement and Description

As with any other type of archival record, the arrangement and description of audiovisual media should provide intellectual and physical access to all records, regardless of media, and should preserve and express the relationships between the records within a collection.

The DACS (Describing Archives: A Content Standard) and EAD standards are effective in supporting these goals, but they do not provide consistent or thorough guidance on audiovisual material. And although audiovisual cataloging specialists have made significant progress in standards development for item- and collection-level description of audiovisual media, there is little published guidance for describing archival audiovisual media in finding aids. In fact, the DACS standard even refers to item-level standards for those looking for guidance in describing audiovisual materials in their collections.

Although item-level standards provide a reference point for elucidating and defining certain elements of description in finding aids, a wholesale adoption of item-level standards for describing audiovisual materials can be problematic. Finding aids provide a hierarchical structure meant to express relationships among records in a collection and to make use of efficiencies in multi-level description, where a component can inherit description from higher levels. Item-level descriptive information systems tend to be flat and do a poor job of expressing relationships among records so described. They also tend to repeat common metadata in each record, which is not efficient. Also, as DACS states quite explicitly, the level of description for a particular component of an archival collection is supposed to match the level of processing. The flip side is that more detailed description means more laborious processing. While some recordings might merit such time and effort, many do not, and a high level of detail may mislead researchers as to its significance in relation to other records in the collection that have been more efficiently described. If item-level description is the only tool a repository has for describing its audiovisual material, it may be tied to that process even when the material could just as easily be described in a single aggregated component or a simple list in a finding aid.

Item-level records are very good at capturing the many possible metadata elements of audiovisual media that do not exist for paper records. However, following the "more product, less process" approach now widely accepted as archival best practice, a critical assessment must be made of how much the researcher gains by knowing many of the format details one could include in the description of recordings. At the Archives of American Art, guidelines direct archivists to limit their description to the minimum needed for physical and intellectual access. That is, we include what is necessary for researchers and archives staff to understand what the content of a recording is, how it relates to other documentation in the collection, and how it may be accessed. Many of the details about its recording characteristics and technical specifications can be left out of the finding aid. If researchers are interested in such details, they can find them when they access the material for research.

That said, archivists must ensure that their arrangement and description of audiovisual material are clear and make sense of the material, which can often be complicated by multiple versions, formats, and production elements. If the material does not make sense to the processing archivist, it will not make sense to the researcher. Disambiguating versions and components of archival media is the sense making of processing work. If this work has not been done, the material has not been processed.

Other principles of arrangement and description will be familiar to any processing archivist. First, as with any format, processing should preserve

and express relationships among records. In any mixed-media collection, chances are that at least some of the audiovisual media is related to paper or other types of records in the collection, so media should not automatically be segregated from the paper records. Also, audiovisual media can be effectively described in the aggregate. If a collection contains 30 cassettes of annual meetings with detailed notes on their cases, "annual meetings, 1975-1993, 30 sound cassettes" is an adequate description for a minimally processed collection. Archivists can guide the researcher to key documents in series descriptions that will help unlock the content, and they can go back to heavily used collections to provide more detail in their descriptions if merited.

At the other extreme, overemphasizing format where audiovisual media is concerned can also lead to under-described material. An inventory of media types is not an adequate description of archival recordings, although such description is commonly found in finding aids. The DACS guidance for devising titles applies here, as anywhere: "When devising title information, compose a brief title that uniquely identifies the material, normally consisting of a name segment, a term indicating the nature of the unit being described, and optionally a topical segment ..."(Society of American Archivists 2013, 17). Use names, genres, locations, and subjects to devise a succinct and unique description of the recording content, and express media formats in the physical description area.

A chronic problem in the adoption of EAD has been the tendency to retrofit description to display, leaving the metadata compromised and out of standard. For example, the existence of a copy or a location gets noted in a unit title. The issues this creates become painfully apparent when migrating or sharing metadata between systems. In this era of aggregating and linking descriptive metadata, it is increasingly important to follow standards for tagging metadata so that the code that underlies the archival description is in standard. If a stylesheet is improperly designed and does not display the metadata elements used in audiovisual description (perhaps because these elements are not used for other types of records and have not been accounted for), the stylesheet must be corrected.

The instructions for describing audiovisual material in EAD at the Archives of American Art were designed to be a local document integrating guidelines for audiovisual description with local guidelines for general archival description, with a few notes on arrangement as well. They provide detailed instructions and refer specifically to local practice, and will be updated as tools evolve. As of this writing, the guidelines specify where certain types of metadata are to be entered in an Archivists' Toolkit environment, referring to metadata elements in the EAD 2002 standard. The guidelines were developed through a process that began with a close review of the EAD 2002 tag library and DACS (second edition), a review of the Archives' EAD and DACS implementation, consultation with archivists at other organizations on their use of EAD for describing audiovisual material, and a refinement and interpretation of the standards to establish local rules. Because of the variety of institutional practices and contexts, the Archives' guidelines would likely need review and adaptation to be adopted by other repositories. We also anticipate a significant revision with a future migration to ArchivesSpace and EAD3. Still, the guidelines provide answers to many of the questions that arise in the course of processing audiovisual components of mixed-media collections. As such, they fill a gap in existing standards.

References

Adelstein, Peter Z. 2009. *IPI Media Storage Quick Reference, second edition*. Rochester: Image Permanence Institute. Available at https://www.imagepermanenceinstitute.org/webfm_send/301.

Greene, Mark, and Dennis Meissner. 2005. "More Product, Less Process: Revamping Traditional Archival Processing." *American Archivist* 68:2. Available at http://americanarchivist.org/doi/pdf/10.17723/aarc.68.2.c741823776k65863.

ICA Sub-committee on Descriptive Standards. 2000. *ISAD(G): General International Standard Archival Description.* Ottawa: International Council on Archives. Available at http://www.ica.org/10207/standards/isadg-general-international-standard-archival-description-second-edition.html.

ISO (International Organization for Standardization). 2000. *ISO 18923: Imaging Materials – Polyester-base Magnetic Tape – Storage Practices.* Geneva: International Organization for Standardization.

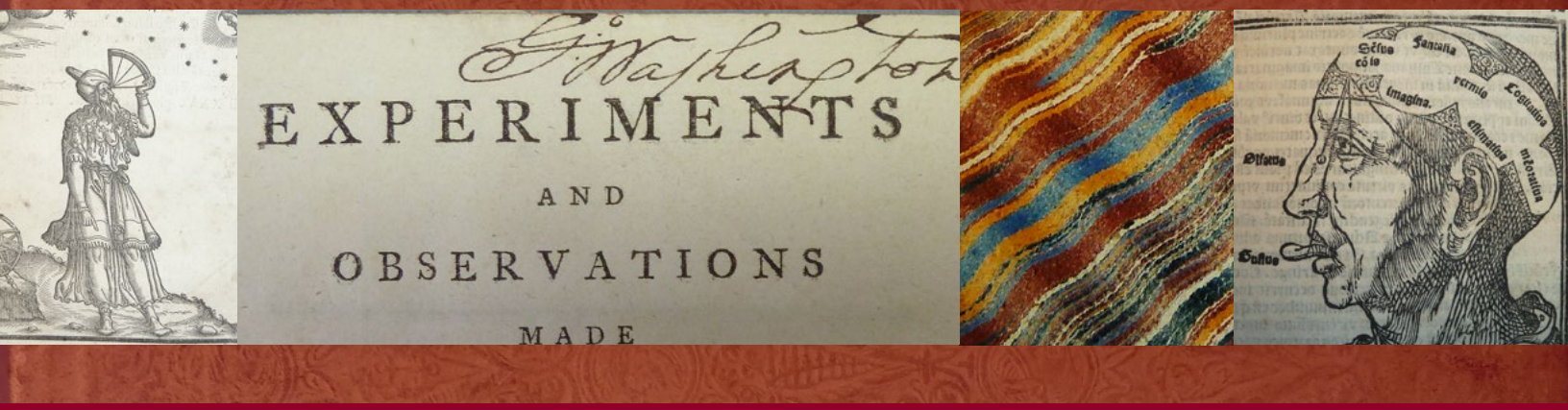
Library of Congress. 2002. *Encoded Archival Description Tag Library, Version 2002.* Available at http://www.loc.gov/ead/tglib/element_index.html.

Pearce-Moses, Richard. 2005. *A Glossary of Archival and Records Terminology.* Chicago: Society of American Archivists. Available at http://www2.archivists.org/glossary/terms/a/archival-processing.

Roe, Kathleen. 2005. *Arranging and Describing Archives and Manuscripts (Archival Fundamentals Series II).* Chicago: Society of American Archivists.

Society of American Archivists. 2013. *Describing Archives: A Content Standard, Second edition.* Chicago: Society of American Archivists.

Science Collections



The Opportunities of Engagement: Working with Scholars to Improve Description and Access at the Center for the History of Medicine

Emily R. Novak Gustainis, Head, Collections Services, Center for the History of Medicine, Francis A. Countway Library of Medicine

Abstract

The Center for the History of Medicine has used its two CLIR-funded grant initiatives to engage researchers at all stages of their professional careers in an effort to understand how archival description can be improved or modified to the benefit of scholars while remaining attentive to workflows that speed processing. This paper considers findings from the Center's "Foundations of Public Health Policy" (2008) and "Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections" (2012) initiatives.

As the Center for the History of Medicine's (the Center) head of collections services, my day-to-day work focuses on the mechanisms of discovery: systems, standards, and practice. While I occasionally staff the reference desk, it requires planning on my part to meaningfully interact with our researchers and find out if what we are doing on the back end is meeting their needs. Researchers are our prime movers, the reason we kick the gears into action and then examine the cogs, so it is no wonder user assessment is instrumental to trying out new ways to do our work. Over the past seven years and through two CLIR grants, the Center has posed a number of questions to its constituents: How do you prefer to learn about collections? What makes a finding aid useful? What is the minimum amount of information you need to determine if a collection is worth your time? And, most recently, how can we help improve access to records containing health information about individuals? These opportunities to engage fuel our data collection, both on our processes and our products.

I will say up front that there are only two ideas to take away from this paper, and I expect they are not revolutionary. One, repositories should increase the depth of description for collections that have limited access; and two, approaches to description should better recognize the needs of different disciplines. In my experience, our inclination as archivists is to provide much more description for open collections because they can be used right now, rather than to expend energy helping our users make tough choices about whether or not to go through the process—when it is even possible—of appealing for access to unprocessed collections. Similarly, our urge to suppress catalog records for unprocessed holdings can drastically reduce opportunities for our researchers to discover collections, resulting in missed opportunities to gauge what is of user interest.

With the Center's 2008 grant, "Foundations of Public Health Policy" and 2012 grant, "Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections," which was proposed under the auspices of the Medical Heritage Library and conducted in partnership with the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions, the Center had the opportunity to explore, respectively, the

descriptive needs of a wide variety of researchers and those of medical and social historians needing access to records containing health information about individuals held by both HIPAA (Health Insurance Portability and Accountability Act) and non-HIPAA covered repositories. The results were polar opposites of one another.

Foundations of Public Health Policy Grant

For "Foundations of Public Health Policy," the Center committed to building a community of interested public health practitioners, scholars, and students to support the acquisition and use of public health collections and to advise us on collections description and outreach. For descriptive assessment, this included recruiting 52 public health and information professionals, students, and historical researchers to evaluate the use of box and folder lists for the collections we were processing, those of public health and public health administration leaders Leona Baumgartner (1902-1991), Allan Macy Butler (1894-1986), Howard Hiatt (1925-), and David Rutstein (1909-1986). The box and folder lists were grouped by series and delivered online as spreadsheets through a take-home exercise and a post-exercise interview. Our objectives were to:

- Assess how useful researchers found box and folder lists containing select metadata independent of any top-level descriptive information associated with a finding aid (such as biographical notes, subject access point, or series descriptions)
- Ascertain whether or not researchers could perform routine information-seeking tasks associated with identifying materials of interest for research use with just a spreadsheet
- Determine whether a spreadsheet provided enough descriptive information to engender "trust" in the resource researchers were using; that is, to find out whether researchers were comfortable with this resource alone as a determining factor to schedule a research visit

Additionally, the scholarly engagement exercise and interviews were helpful to observe the potential challenges MPLP (or at the Center, what we would consider "appropriate level" processing) and other innovations may present to users. Participants, who were divided across the four collections processed for the project, were asked to complete an exercise (Appendix A) consisting of both general and quantitative multiple choice questions and collection-specific, qualitative questions about how they used two versions of an Excel spreadsheet to answer a number of questions. This included an "A" version spreadsheet with nine elements (box number, folder number, series, subseries, sub-subseries, folder title, begin date, end date, and in the case of Hiatt alone, access restrictions) and a "B" version with additional descriptive information provided in a notes field, such as to indicate the presence of photographs. The average completion time of the exercise was one hour, with the post-assignment interview averaging 30 minutes. In retrospect, it was impressive we had 52 respondents, though we did need to eliminate some of the non-U.S. participant data because of differing interpretations of survey terminology.

From our interviews, we gleaned the following:

- By overwhelming consensus, participants agreed that a spreadsheet was a good enough tool to make preliminary selections about what to look at during a research visit, particularly in conjunction with a collection-level record available through HOLLIS, the online catalog of the Harvard University Library.

- Folder lists without contextual groupings (such as series or subseries) would have been less helpful to discovery. Comments from participants confirmed that folder titles can be "deceptive" on their own. Knowing that a folder named "Meetings with Jane, 1990" is part of a series of teaching records or a series of patient files substantially changes a researcher's interpretation of a folder title.
- Inexperienced researchers tended to want more descriptive information and preferred spending less time at a repository. Experienced researchers cared only about access. Moreover, experienced researchers stated that regardless of how an archivist handled a collection, they wanted to conduct their own records review. In other words, they were not going to take anyone's word on what was in a collection or if it was relevant to their work.
- No one was concerned about subject access and only a handful of people mentioned an interest in having more contextual information (beyond groupings) or content-related information.
- Participants cared about record formats. Most liked the "B" version of the spreadsheet with the notes field because the occasional notes provided by processors indicated the presence of photographs or other non-textual items.
- Participants liked the idea of applying a minimal controlled vocabulary to a folder title ("qualifiers") to help disambiguate folder transcriptions, such as "Writings" or "Correspondence." However, multiple qualifiers would have to be employed; if you qualify a folder with "writings," the implication is that the whole thing will be writings, not writings intermixed with correspondence and reference material to support said writings. One individual thought indicating whether or not a preceding or succeeding folder that had a more descriptive title (such as activities) contained the same type of records as a vaguely titled folder.
- Unless archivists are going to provide a substantial number of folder-level scope and content notes, it does not seem worth providing them at all. Most people did not think folders with notes were "more important," and very few people assumed Center archivists were making judgment calls on content.
- While most researchers assumed the spreadsheet could be manipulated, such as to sort by date or add columns, they did not do so. They scrolled and keyword-searched the spreadsheets. For two collections, we provided one massive spreadsheet, and for two collections we provided folder lists tabbed out in Excel by series. The tabbed versions of spreadsheets ended up confusing a few participants, particularly those uncomfortable with Excel. Spending time on novel ways to use the spreadsheet was not of much interest.
- Thoughtful accessioning practices can facilitate access in advance of a full finding aid. Public health processors roughly sorted the records into series and subseries in advance of box and folder listing. The more astute we are about doing this at the point of acquisition or accessioning, the faster we can enable meaningful access in lieu of full finding aids.

As a result of the grant, our then acquisitions archivist—and what is now our acquisitions team—spends much more time grouping and packing records up front and on-site with the donor when feasible. Because all acquisitions are listed before being sent to off-site storage, the result is a higher- quality collection inventory that is easier for public services staff to navigate and a resource that speeds processing planning. We are also far

more comfortable allowing researcher access to unprocessed collections.

Private Practices, Public Health Grant

In 2012, the Center partnered with the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions to process the papers of seven leaders in public health research and advocacy. The grant, "Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections," addressed issues of access in both HIPAA and non-HIPAA covered entities, seeking best practices for archivists facing challenges and confusion over making health-related records in their collections available. Specifically, the project sought to: a) understand the informational needs of scholars seeking to use restricted records and what they consider most valuable to research; b) evaluate how language used by the Center and Hopkins to communicate restrictions in finding aids correlates with the perceived utility of the finding aid and the potential of the collection to satisfy a research need; and c) get feedback on the process for applying for access to collections with restricted records.

To do so, the Medical Heritage Library, represented by the Center's Kathryn Hammond Baker, Scott Podolsky, and Emily Novak Gustainis, and Hopkins's Phoebe Evans Letocha, led discussion sessions, launched an online survey, presented at professional conferences, and ultimately distributed Novak Gustainis and Evans Letocha's jointly authored "Recommended Practices for Enabling Access to Manuscript and Archival Collections Containing Health Information about Individuals" to the research and professional communities for feedback. While this engagement work was under way, Center and Hopkins staff were simultaneously processing collections, including the Center's collections of the papers of Oliver Cope (1902-1994), Stephen W. Lagakos (1946-2009), Erich Lindemann (1900-1974), and Arnold S. Relman (1923-2014), as well as the records of the Harvard School of Public Health's Department of Biostatistics, 1981-2009.

Efforts included:

- A SurveyMonkey survey (Appendix B) on access to health records, "Research Access to Protected Records Containing Health Information about Individuals," was distributed to the Medical Heritage Library governance committee and circulated to professional and discipline-directed listservs. In total, 63 people responded. Data obtained as a result of the survey were analyzed for engendering discussions between archivists and historians and informing the creation of the Recommended Practices document.
- A workshop for Harvard University's History of Medicine Working Group, comprising graduate students and faculty from the History of Science Department, which included an interactive review of finding aids for collections containing restricted records and a discussion of the type of descriptive content researchers might need to evaluate the usefulness of the records for research.
- A lunch workshop at the 2014 annual meeting of the American Association for the History of Medicine (AAHM) to elicit the information historians need to determine whether or not it is worth applying to an Internal Review Board (IRB), what information is most useful to them, and what they think is missing from finding aids. As part of the session, historians Janet Golden, Rutgers University, and Cynthia Connolly, University of Pennsylvania, shared with the audience their research experiences and difficulties using patient records to inform

their research and launched a discussion of how to improve the user experience.

- A presentation at the 2014 Annual Meeting of the Society of American Archivists in Washington, DC. The session, "Partners in Practice: Archivists and Researchers Collaboratively Improving Access to Health Collections," offered perspectives of both historians/researchers and archivists on the importance of making a wide variety of records that contain health information about individuals discoverable. The session was moderated by Susan Lawrence of The Ohio State University and included a presentation by John Harley Warner of Yale University, "Why Patient Records Matter to the Historian."

As a result of these efforts, it became clear that because of the complexities related to applying for access to records containing health information about individuals—whether closed to comply with HIPAA, state law, or institutional policy—different kinds of descriptors were necessary. When survey participants were asked whether or not they could apply to a review board to obtain access, only 56.14 percent (32) had access to a review board, and of the 32 individuals who did have access, only 56.25 percent (18) actually went through the process. Respondents said it took too much time—especially when they learned about the restrictions only when they were already on-site, were not convinced they would actually get access if they went through the process, or believed that they lacked support or guidance. Respondents stated that the most significant barrier to using records containing confidential/protected health information held by special collections, archives, and museums was lack of information about the records themselves and access procedures, stating, "I see records that look interesting in catalogs or collection guides, but I can't tell if they will be useful" and "The process takes too long."

Researchers simply needed more information to help them make decisions. As professionals, we cannot control what the Internal Review Board or Access Board applications require or how frequently they meet, but we can provide more description. Offering samples of variables found in the records, such as patient diagnosis or condition; the age, gender, and race of patients; and the types of medications or procedures, can make a difference. Better indicating the kinds of records in a collection (patient histories vs. case files vs. medical records, rather than just "patient records") and providing more explicit statements at the folder level indicating why the folder is restricted, and for how long, can also help. For example, all of the Center's finding aids authored for the grant include descriptive information about the type of restrictions found in the collections, why they were imposed, and how to obtain access. At the folder level, the Center provided a transcription of the full folder title (redacting patient names), the year the records will open to the public, and a qualifying description in the form of a folder-level scope note to convey the intellectual contents of the folder without revealing protected information. Because the Center is a non-HIPAA covered entity, access restrictions were determined by sampling the content of records in each folder. Hopkins, a HIPAA-covered entity, has a similar descriptive approach, except Hopkins staff must screen documents at the item level for protected health information. Readers are encouraged to review the paper "The Practice of Privacy" by Emily Novak Gustainis and Phoebe Evans Letocha, in this volume, for more detailed information about the descriptive process.

So how do we balance the needs of users who are seemingly satisfied with a structured box and folder list on one side of the continuum and those that want the kind of information that can only be gleaned by labor-intensive sampling? Perhaps for public health and other to-be-determined collections with little or no patient records, we should scale back on descriptive processes, letting the collections do the talking while we ramp up efforts to describe records that do not easily present. Archivists can also become more comfortable describing boxes of access-protected records only at the container or series level, focusing on the records that comprise a series more holistically. Flexible and appropriately applied approaches to processing can be built only through concerted efforts to understand the needs of our multiple constituencies. The Center is now testing the best practices developed for the description of collections containing health information about individuals on the Dwight E. Harken papers, 1930s-1990s, and should be completed concurrent to the symposium. We look forward to sharing the finding aid with members of the community for feedback.

References

Gustainis, Emily. 2015. Now Available! Recommended Practices for Enabling Access to Manuscript and Archival Collections Containing Health Information About Individuals (Feb. 12). Available at http://www.medicalheritage.org/2015/02/now-available-recommended-practices-for-enabling-access-to-manuscript-and-archival-collections-containing-health-information-about-individuals/.

APPENDIX A: Example User Study: "Foundations of Public Health Policy" (2008)

Center for the History of Medicine, User Study, 2009–2010

As part of its Foundations in Public Health Policy grant work, the Center for the History of Medicine, Countway Library, is conducting a study to document and analyze how researchers use and respond to different access tools developed for delivering information about its collections to the public. This grant, as funded by the Andrew W. Mellon Foundation and administered by the Council on Library and Information Resources (CLIR), has enabled the Center to experiment with ways to make collections available to researchers over shorter periods of time. Your participation in this study will help us determine the efficacy and utility of these tools.

The study consists of two parts:

1. An exercise consisting of both multiple choice survey questions and detailed questions designed to assess how participants use two different versions of a spreadsheet to answer questions about a collection. *We estimate that it will take about one hour to complete the study exercise.*

2. An interview (either at the Center or via conference call) that will focus on content and usability issues related to how you used the spreadsheets to complete the exercise. *We estimate the interview to take between thirty minutes and one hour.*

We sincerely appreciate your taking the time to participate and provide feedback. We are happy to answer any questions you might have about the project. Please contact Michael Dello Iacono, Project Archivist, MPD13@hms.harvard.edu, and Emily R. Novak Gustainis, Collections Services Archivist, ERN6@hms.harvard.edu.

A. Introductory Questions

1. How would you identify yourself?

_____ A new/inexperienced researcher who has not used many archival or manuscript collections

_____ An experienced researcher who has visited a number of archives and used a number of archival or manuscript collections

_____ A person in the Public Health Field new to archival research

_____ A person in the Public Health Field who is an experienced researcher

_____ An information professional (librarian, archivist, metadata specialist, etc.)

2. How comfortable are you conducting research online and using electronic resources?

1	2	3	4	5
Very comfortable				Very uncomfortable

3. How many archives have you visited in the last three years to conduct primary research?

a) 0 repositories b) 1–5 repositories

c) 6–10 repositories d) More than 10

4. Have you been unable to use a collection at a repository because it was unavailable for research use (or was "unprocessed")? (Please circle)

Yes No

5. If so, how many times during the last three years? _____

6. Which statement best reflects how you feel about the amount of time an archives or special library takes to provide public access to a new collection you are interested in using?

a) I would rather have a repository provide access to a collection (or part of a collection) that has been minimally reviewed for research use (for example, only has a box list) so long as the collection is made available to the public as soon as possible.

b) I would rather wait until a collection has been well organized and thoroughly documented for optimal research use, even though it may take longer for the repository to make it available to the public.

7. Which is more important to you?

a) Being able to spend less time at a repository because I have very detailed information about a collection and know exactly what to have pulled for me, even if it means I might not be able to access a collection in the immediate future while this type of information is being collected.

b) Having access to a collection, even if it means I may need to spend a lot of time at a repository searching for the information I want.

8A. If you could only have ONE of the following discovery tools, which would you rather have available to you online and in ADVANCE of your visiting the Center for the History of Medicine?

a) A detailed summary of a collection's content, biographical/institutional information about the creator(s) of the collection, and information about what kinds of documents are in the collection, including date spans for all materials.

b) A spreadsheet containing a list of every folder "title" in every box of the collection as it was originally labeled by the person who created or assembled the collection.

8B: Why did you pick A or B?

9. Which is more important to you?

a) Being able to simultaneously search for subjects or people across many finding aids in order to discover which collections at a particular repository may help me with my research.

b) Being able to print an inventory for, or guide to, a collection for personal reference use from a list of collections posted on a repository's website.

B. Spreadsheet-Specific Questions

Part B-I

To complete Part B-I, please copy and paste the following location into your web browser: http://repository.countway.harvard.edu/xmlui/handle/10473/3600

Under "Sample Submissions," select the file: CLIR_baumgartner_boxlist_A.xls

Under "Associated Files," click on "View/Open" for the file: CLIR_baumgartner_boxlist_A.xls

Please take a few minutes to look at the spreadsheet and then answer the following questions. Please feel free to cut and paste answers from the spreadsheet into this Word document.

1. What are the five major groups of records found in the collection?

2. Pick two of the groups you identified in question 1. What kinds of information would you expect to be in these groups of records?
3. What types of materials would you expect to find in the collection?
4. Why?
5. In what year did Baumgartner travel to Russia?
6. Please name two of Baumgartner's published articles from the 1950s.
7. What steps did you take to answer to questions 5 and 6?
8. How much material is there in the collection related to speeches given by Baumgartner?
9. What steps did you take to answer question 8?
10. If you were looking for letters between Baumgartner and her family members, where would you expect to find them?
11. Please list three individuals Baumgartner corresponded with.
12. From 1954-1962, Leona Baumgartner served as Commissioner of Public Health for the city of New York. Where would you look for records in the collection that relate to Baumgartner's public appearances as Commissioner?
13. You are researching the activities of the American Public Health Association (APHA) in the 1950s and 1960s. Which boxes would you ask to see?
14. How would you find out if there were items from the 1970s?
15. How would you determine how much material from the 1970s is in Series 4?
16. Did you need to print the spreadsheet in order to answer the above questions?
17. Did you save the spreadsheet to your desktop or local drive before working on the questions?

Part B-II

To complete Part B-II, please copy and paste the following location into your web browser:
http://repository.countway.harvard.edu/xmlui/handle/10473/3600

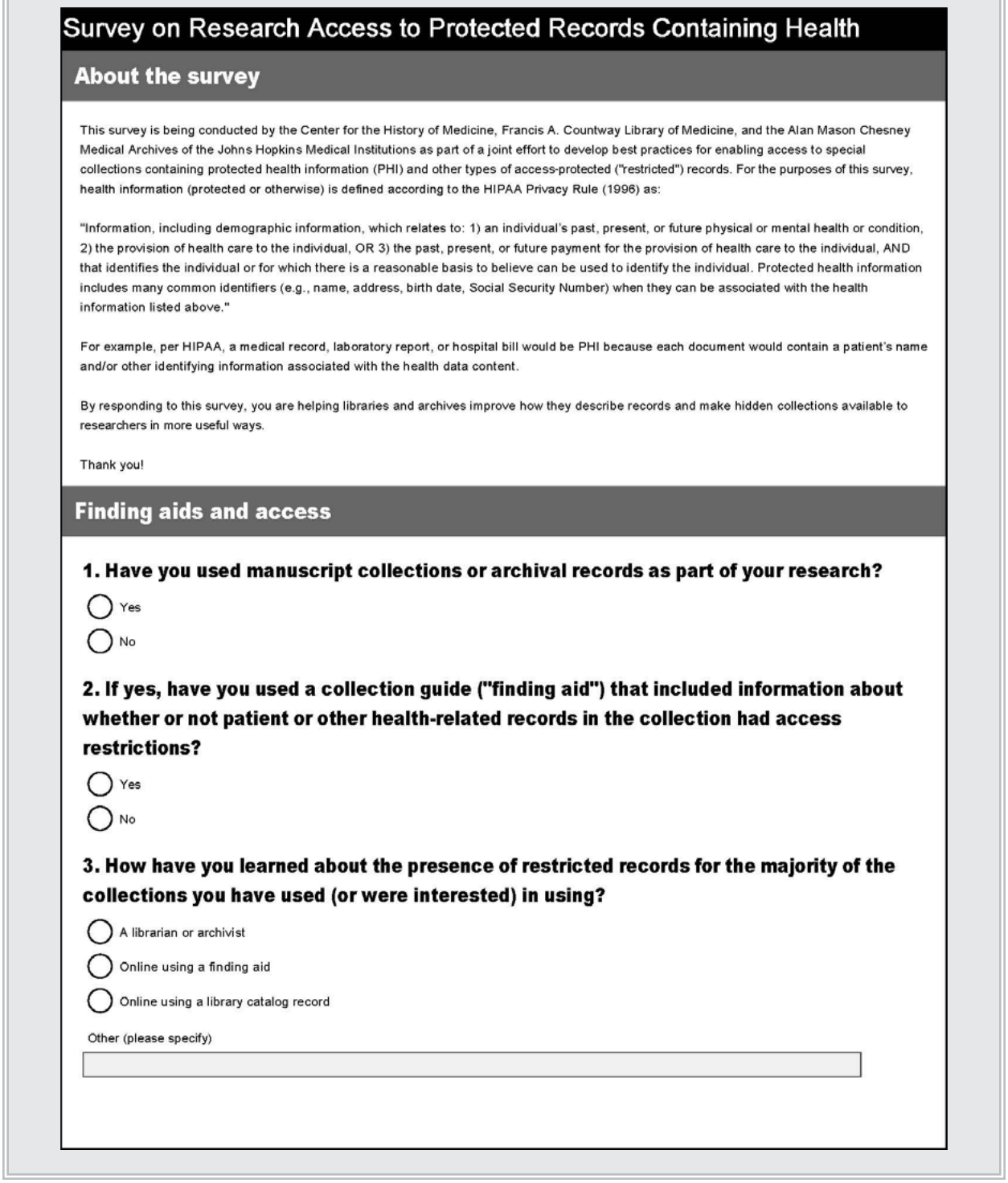
Under "Sample Submissions," select the file: CLIR_baumgartner_boxlist_B.xls

Under "Associated Files," click on "View/Open" for the file CLIR_baumgartner_boxlist_B.xls

Please take a few minutes to look at the spreadsheet and then answer the following questions. Please feel free to cut and paste answers from the spreadsheet into this Word document

1. Summarize the kind of information provided in the "Notes" column.
2. What types of materials would you expect to find in the collection?
3. Baumgartner delivered many speeches and lectures throughout her career. Please name two people who delivered a talk along with Dr. Baumgartner.
4. Please explain how you found the above answer.
5. Do folders with notes in the "Notes" column contain more important documents?
6. Please list the question numbers in section B-I that you would now answer differently having seen the "Notes" column.

APPENDIX B: Private Practices, Public Health Survey (2012)



Survey on Research Access to Protected Records Containing Health

About the survey

This survey is being conducted by the Center for the History of Medicine, Francis A. Countway Library of Medicine, and the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions as part of a joint effort to develop best practices for enabling access to special collections containing protected health information (PHI) and other types of access-protected ("restricted") records. For the purposes of this survey, health information (protected or otherwise) is defined according to the HIPAA Privacy Rule (1996) as:

"Information, including demographic information, which relates to: 1) an individual's past, present, or future physical or mental health or condition, 2) the provision of health care to the individual, OR 3) the past, present, or future payment for the provision of health care to the individual, AND that identifies the individual or for which there is a reasonable basis to believe can be used to identify the individual. Protected health information includes many common identifiers (e.g., name, address, birth date, Social Security Number) when they can be associated with the health information listed above."

For example, per HIPAA, a medical record, laboratory report, or hospital bill would be PHI because each document would contain a patient's name and/or other identifying information associated with the health data content.

By responding to this survey, you are helping libraries and archives improve how they describe records and make hidden collections available to researchers in more useful ways.

Thank you!

Finding aids and access

1. Have you used manuscript collections or archival records as part of your research?

- ◯ Yes
- ◯ No

2. If yes, have you used a collection guide ("finding aid") that included information about whether or not patient or other health-related records in the collection had access restrictions?

- ◯ Yes
- ◯ No

3. How have you learned about the presence of restricted records for the majority of the collections you have used (or were interested) in using?

- ◯ A librarian or archivist
- ◯ Online using a finding aid
- ◯ Online using a library catalog record

Other (please specify)

Survey on Research Access to Protected Records Containing Health

4. What are the kinds of records you were interested in using, but were restricted? Check all that most often apply:

- ☐ Medical records and indices (whether patient, diagnostic, or other) maintained by a healthcare provider, such as a hospital or medical practice
- ☐ Medical imaging records, such as x-rays
- ☐ Photographs of patients
- ☐ Psychiatric or other mental health-related records, such as psychotherapy notes
- ☐ Research records (such as datasets, human subject research information, etc.) that contain personally identifiable information ("personal identifiers"), such as names, addresses, phone numbers, medical records numbers, etc.
- ☐ Other (please specify)

Use of IRB

5. Was submitting a request to an Internal Review Board (IRB) to use the records a possibility?

- ◯ Yes
- ◯ No

6. If yes, did you end up submitting an IRB to access the records you were interested in using?

- ◯ Yes
- ◯ No

7. If no, why not?

Barriers to use

Survey on Research Access to Protected Records Containing Health

8. What do you think is the most significant barrier to your use of records containing confidential/protected health information held by special collections, archives, and museums?

- ◯ The IRB process takes too long
- ◯ Too much paperwork is required to get access to restricted records
- ◯ I see records that look interesting in catalogs or collection guides, but I can't tell if they will be useful
- ◯ Nothing is digitized
- ◯ Repositories aren't open when I have time to do my research
- ◯ I don't think I'll be able to quote, reference, or use the records in publications
- ◯ No one will tell me if I can use the records or not
- ◯ This does not apply to me. There are no barriers to my use of records

Other (please specify)

9. What descriptive information do you think is missing from library catalog records or collection guides (such as those for a manuscript collection)? What information would be most useful to you in deciding whether or not a collection has information relevant to your research?

Ways to Maximize Access

Survey on Research Access to Protected Records Containing Health

10. How useful would having the following descriptive information be in determining whether or not you would submit an IRB to use restricted records containing protected health information?

	Not very useful	Somewhat useful	Very useful	Does not apply to my research
Average age of patients at time of treatment	◯	◯	◯	◯
Date span of records	◯	◯	◯	◯
Diagnosis/condition	◯	◯	◯	◯
Duration of treatments	◯	◯	◯	◯
Genetic information	◯	◯	◯	◯
Geographic region covered	◯	◯	◯	◯
Names of medical devices used	◯	◯	◯	◯
Prescribed medications	◯	◯	◯	◯
Race of patients	◯	◯	◯	◯
Sex of patients	◯	◯	◯	◯
Surgical procedures	◯	◯	◯	◯
Treating physician/surgeon	◯	◯	◯	◯

Other (please specify)

Survey on Research Access to Protected Records Containing Health

11. How useful would knowing that the following record formats were in a group of restricted records be to your determining whether or not to submit an IRB?

	Not very useful	Somewhat useful	Very useful	Does not apply to my research
Admission/registration records	○	○	○	○
Autopsy records	○	○	○	○
Billing information	○	○	○	○
Case files	○	○	○	○
Consultation files	○	○	○	○
Correspondence	○	○	○	○
Diagnostic indices	○	○	○	○
Family medical histories	○	○	○	○
Genetic testing records	○	○	○	○
Graphs and charts	○	○	○	○
Hospital policies	○	○	○	○
Immunization records	○	○	○	○
Insurance claims	○	○	○	○
Informed consent records	○	○	○	○
Lab notebooks	○	○	○	○
Microscope slides/specimens	○	○	○	○
Photographs/medical imaging	○	○	○	○
Patient histories	○	○	○	○
Patient questionnaires	○	○	○	○
Patient summaries	○	○	○	○
Prescription books/logs	○	○	○	○
Research protocols	○	○	○	○
Surgical logbooks	○	○	○	○

Other (please specify)

Comments and Demographics

12. Do you have comments regarding access to special collections containing health records?

Survey on Research Access to Protected Records Containing Health

***13. How do you identify yourself?**

- ◯ A student (any field)
- ◯ A professor/instructor of history, the history of medicine, or the history of science
- ◯ A professor/instructor of another Humanities sub-discipline or a different discipline
- ◯ A researcher (no academic affiliation)
- ◯ A physician or healthcare provider with an interest in the history of medicine or science
- ◯ A librarian or archivist

14. How long have you identified as the above?

- ◯ 1-5 years
- ◯ 6-10 year
- ◯ 11-15 years
- ◯ Over 15 years

15. If we have questions about your responses to the survey, can we contact you to follow-up?

- ◯ Yes
- ◯ No

16. If so, please provide your name, affiliation, and email adresss:

THANK YOU FOR YOUR PARTICIPATION!

The Practice of Privacy

Emily R. Novak Gustainis, Head, Collections Services, Center for the History of Medicine, Francis A. Countway Library of Medicine

Phoebe Evans Letocha, Collections Management Archivist, the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions

Abstract

Drawing on our experiences managing CLIR-funded processing initiatives for public health collections containing protected health information as part of the grant "Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections," we offer insight into how we developed "Recommended Practices for Enabling Access to Manuscript and Archival Collections Containing Health Information About Individuals," which appears at the end of this report, and we make recommendations for enabling access.

Whether privacy is legally mandated, as with HIPAA (the Health Insurance Portability and Accountability Act) and FERPA (the Family Educational Rights and Privacy Act); governed by institutional, state, or federal records schedules; or applied per local practice, the result is that many repositories maintain records that pose significant challenges to access. Yet as researchers continue to focus their attention on histories of medicine, public health, science and technology, disability studies, and patient care—and, increasingly, investigate explanations for more recent developments—they seek out these very collections. How can archivists promote the use of records that inform social and medical histories through the lens of patient care and help researchers decide if an archival collection is useful and worth their time?

By putting these questions up for discussion, and suggesting new answers, we encourage partnerships between archivists and researchers in the area of health collections, advocating for these collections' importance despite access anxiety and the very real challenges of preserving, screening, and making available records of a potentially sensitive nature. Drawing on our experiences managing CLIR-funded processing initiatives for public health collections containing protected health information as part of the grant "Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections," we review best practices developed during the project and make recommendations for enabling access.

The project "Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections" was proposed by the Center for the History of Medicine of the Francis A. Countway Library on behalf of the Medical Heritage Library (MHL). It received funding from The Andrew W. Mellon Foundation as administered by the Council on Library and Information Resources (CLIR) in 2012; project work began in April 2013. The grant enabled the Center for the History of Medicine and its partner, the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions—both MHL principal contributors—to open currently inaccessible public health collections to researchers.

The collections opened as a product of this grant include the Oliver Cope papers, 1891-1992 (Countway); the William George Hardy and Miriam Pauls Hardy Collection, 1875, 1930-2008 (Hopkins); the Harvard School of Public Health's Department of Biostatistics records, 1981-2009 (Countway); the Stephen W. Lagakos papers, 1979-2009 (Countway); the Erich Lindemann papers, 1885-1991 (Countway); the Elmer V. McCollum and Harry G. Day Collection, 1881-2003 (Hopkins); the B. Frank Polk Collection, 1972-1990 (Hopkins); the Arnold S. Relman papers, 1953-2011 (Countway); and the Barbara Starfield Collection, 1948-2011 (Hopkins).

What is HIPAA? What Does it Mean to be a Covered Versus Non-covered Entity?

The adoption of the Privacy Rule under HIPAA, which went into effect on April 14, 2003, has had a major impact on archivists responsible for collections documenting the health sciences and on the researchers who want to use these collections. HIPAA was the first comprehensive federal law on access to and use of health information; the first general federal medical privacy law to extend rights of privacy beyond the file unit of the medical record to individually identifiable health information in all types of file systems, documents, formats, and media; and the first federal law to extend rights of privacy beyond health information of living individuals to health information of the deceased. The Privacy Rule applies only to archives designated as part of HIPAA-covered entities and their business associates, and does not apply to archives not part of covered entities that also hold medical records and other related health information. Archival repositories subject to HIPAA are subject to serious penalties for breaches.

Archives work with their legal counsel to determine whether they are subject to HIPAA. The extension of the HIPAA privacy and security requirements to business associates (as a result of the 2013 changes to the Privacy Rule brought about by the HITECH Act) may bring many more archival repositories under HIPAA regulation. These repositories have turned for guidance to the policies and procedures of archival repositories that have been operating under HIPAA since its inception. There is no list of archival repositories that identifies each of their status under HIPAA. The Chesney Medical Archives is part of the Johns Hopkins HIPAA- covered entity. As the official archival repository for the Johns Hopkins Hospital, its holdings include medical records from the hospital. The Center for the History of Medicine, Francis A. Countway Library of Medicine, is not part of a HIPAA-covered entity because Harvard Medical School does not own the teaching hospitals. The History of Medicine Division of the National Library of Medicine is not a covered entity, although it adopted some HIPAA-like policies for access to some collections such as hospital records. Some repositories may close collections based on the assumption that they are covered by HIPAA when they may not be.

Repositories within HIPAA-covered and non-covered entities must also comply with state laws applying to medical records and health information in holdings, comply with the Federal Common Rule for Protection of Human Subjects (for institutions that accept federal research funds), adhere to institutional requirements for protection of health information, and observe donor agreements for protecting health privacy.

One area of possible confusion may be the differences between the state definition of medical record and the HIPAA definition of protected health information. HIPAA defines protected

health information as "individually identifiable health information transmitted or maintained in any form or medium (electronic, oral, or paper) by a covered entity or its business associates, excluding certain educational and employment records and excluding information on those individuals who have been deceased for longer than 50 years." The definition of what is considered a medical record may vary by context and purpose of creation, as well as state law. HIPAA does not define the term *medical record*. It is a term defined more by state law. State medical records laws vary by state, and institutions may interpret and apply the state definition according to local circumstances and systems. State laws may not all have caught up with the definition of protected health information under HIPAA. Records and information related to individuals who have been deceased for more than 50 years may still be protected by state medical records statutes and other state privacy laws.

The variations in whether and how repositories are covered by HIPAA and differences in state laws result in much confusion for researchers wanting to access and use collections containing health information. They may encounter a different set of access policies at each repository they want to use. Not all archives have the resources to support access, such as privacy boards, institutional review boards, or informed legal counsel.

As archivists at two of the country's leading medical archives, we developed a set of recommended best practices in an effort to enable access to manuscript and archival collections containing protected health information and other types of access-protected records containing health information about individuals. We intend this document to inform our colleagues at other medical archives as well as archivists who encounter these collections at archives that do not have a medical focus. We also want historians of medicine and other researchers to familiarize themselves with the issues we raise so they, too, can advocate for the preservation of these materials.

How Did We Engage Researchers, Historians, and Archivists?

For Countway and Hopkins to develop best practices for archivists facing confusion and challenges in making available health-related records about individuals in their collections, it was essential that we understand the informational needs of researchers seeking to use restricted records. We needed to hear their in-the-trenches experiences trying to access such records, and we needed to elicit information about the descriptive content they considered most valuable to discovery. Such an exploration meant evaluating how language used in finding aids and catalog records correlated with the perceived potential of a collection to satisfy a research need. It also meant seeking feedback on the process for applying for access to collections containing protected records.

To do so, we led discussion sessions, launched an online survey, presented at professional conferences, and then distributed our "Recommended Practices for Enabling Access to Manuscript and Archival Collections Containing Health Information about Individuals" to research and professional communities for feedback. Our first action was to distribute an online survey on access to health records to the Medical Heritage Library governance committee and to a number of professional and discipline-directed listservs. In total, 63 people responded. It was this data that helped spur our conversations between archivists and historians.[1]

As part of the survey, we asked respondents to indicate the types of records they were interested

1 Data are available at http://www.medicalheritage.org/wp-content/uploads/2010/06/Data_All_140424_nocontacts.pdf.

in using for their research that required permission from an access or privacy board. Overwhelmingly, these were medical records and indices (patient, diagnostic, or other) created or maintained by a health care provider such as a hospital or medical practice. The second most cited records desired were psychiatric or other mental health-related records such as psychotherapy notes. When asked whether or not they could apply to a review board to obtain access to sought-after records, only 56 percent (32) of the respondents reported having access to a review board. Of the 32 respondents who did have access, only 56 percent (18) actually went through the process of applying. Why was this last number so low?

Respondents said the application process took too much time, especially in cases where they did not find out about the restrictions until they were already on site. Respondents also said they were not convinced they would actually get access if they went through the process, or they lacked the support or guidance to do so. When asked what the most significant barrier was to using records containing confidential and/or protected health information held by special collections, archives, and museums, the top answers were, "I see records that look interesting in catalogs or collection guides, but I can't tell if they will be useful," and "The process takes too long." What can archivists do to combat these challenges? We found that making the process more transparent, enhancing descriptions for these types of records, educating researchers about the process, and advocacy could help combat frustration.

After the survey closed, we sought opportunities to interact with researchers and members of our profession. We held a workshop for Harvard University's History of Medicine Working Group, comprising graduate students and faculty from the History of Science Department. This gathering helped inform us of the needs of emerging scholars. We also held a lunch workshop at the 2014 annual meeting of the American Association of the History of Medicine, called "Negotiating Access to Patient-Related Materials: A Conversation between Archivists and Historians." At this meeting Evans Letocha explained HIPAA and helped eliminate some of the misconceptions about it. She also illustrated how the law affects those trying to use HIPAA-covered records, while Novak Gustainis presented our initial survey findings and discussed the potential impact of findings on processing practices. As part of the session, we were extremely fortunate to have historians Janet Golden of Rutgers University and Cynthia Connolly of the University of Pennsylvania share with attendees their successes and challenges using patient records to inform their own work. This helped attendees learn more about the difficulties and determine whether it is worth applying for access to an internal review or access board to use a collection. One priority that emerged from the discussion was to improve the user experience, particularly through potential partnerships between professional organizations of historians and those of archivists.[2]

Our capstone presentation was made at the 2014 annual meeting of the Society of American Archivists in Washington, D.C. (Gustainis 2014, Evans Letocha 2014). The session, "Partners in Practice: Archivists and Researchers Collaboratively Improving Access to Health Collections," offered the perspectives of both historians-researchers and archivists on the importance of making discoverable a wide variety of records that contain health information about individuals. The session, moderated by Susan Lawrence of The Ohio State

2 Slides are available at http://www.medicalheritage.org/announcements-and-articles/ under "Presentations."

University, included a talk by John Harley Warner of Yale University on why patient records matter to historians. Lawrence works at the intersection of history and research ethics.[3] Warner focuses on the transnational history of medicine and science and is currently working on a study of the transformation of the hospital patient chart from 1801 to the present.[4]

What We Learned about Processing and Description Practices

One of the things we wanted to accomplish with the grant was to understand what, if any, differences existed between processing collections in HIPAA-covered versus non-HIPAA-covered environments. At Countway, we used a time- and labor-tracking database (called MD) that we developed as part of our first CLIR grant (Foundations of Public Health Policy). With this database, we tracked time spent performing activities specific to a collection. These activities include processing tasks such as rehousing, box and folder listing, and encoding and descriptive work related to creating finding aids. In advance of the grant's project start date, we customized a copy of MD for Hopkins, collaboratively determining how discrete processing activities relative to applying restrictions should be recorded and mapped so that we could compare time spent on specific processing actions, including restrictions reviews.

As a result of creating timing analyses for processing all grant-funded collections, we drew six broad conclusions.

- **It is paramount that archivists and collections managers educate researchers about the different types of restrictions in place at repositories.** After looking at the portion of our collections that are access-protected, we discovered that a far greater percentage processed for the Private Practices grant contained records that were closed because they were created by Harvard University and the Johns Hopkins Medical Institutions as a product of operations. It is the variety of records restrictions (including those for students and personnel) in place at our repositories—not just the presence of health information about individuals—that has resulted in large portions of closed collections. As archivists, we assumed that most of the records that had to be access-protected were patient-related. If we were under this impression, then our researchers must be, too.
- **Processing workflows that are systems-dependent requires further evaluation.** Both Countway and Hopkins have very similar processing approaches (e.g., similar series/records groupings, listing and transcribing practices at the folder level, staffing in place to audit description), though Countway's workflows are less sequenced than those of Hopkins because of the collection management system Hopkins uses. It will be important for Countway, when it moves to ArchivesSpace, to monitor adjustments in workflow, timing data, and outputs. Most processing analyses (including those authored by Novak Gustainis) focus on activities independent of systems. Activities articulated in conjunction with more widely used open source systems merit evaluation.
- **Average processing costs per box for a HIPAA-covered and non-HIPAA-covered entity are virtually the same.** Excluding project oversight costs (costs for Novak Gustainis and Evans Letocha), Countway spent $659.83/cubic foot by start volume and

3 See Lawrence 2007.

4 The session description is available at http://archives2014.sched.org/event/4513e77709f236c8a7721c45787e612d#.VgAxW9NVhBd.

$800.90/cubic foot by end volume. Hopkins spent $661.24/cubic foot by start volume and $786.50/cubic foot by end volume. Averaging the two institutions' costs, a reasonable figure for planning purposes would be $660 to $794 per cubic foot under the following conditions: staffing models, compensation rates, and workflows are similar to those of Countway or Hopkins; collections are predominantly analog in format; and they originate in the twentieth century.

- **Screening for restrictions takes longer in a HIPAA-covered environment.** Countway applies restrictions at the folder level through sampling. To account for sampling, Center for the History of Medicine researchers are required to sign a waiver requiring them not to reveal any personally identifying information should they encounter something missed. Hopkins conducts item-level reviews for restrictions and then does a second-pass audit on restricted folders, unless it is obvious that an entire series or subseries will need to be restricted. Researchers at Hopkins can use only what is absolutely confirmed not to contain protected health information, unless they have a waiver of authorization from its Privacy Board or complete an application form for another route of access allowed under HIPAA.[5] The two institutions' hourly rates for restrictions differed greatly despite a similar volume of records—9.76 to 10.61 hours per cubic foot for Countway, and 21.4 to 46.21 hours per cubic foot for Hopkins. Item-level screening, which requires a minimum of two passes—one by the processor and one by the collections services archivist—only partly accounts for Hopkins's higher rate. The number of people involved with processing greatly affects outputs.
- **As the number of people involved with processing goes up, processing outputs go down.** While cost per cubic foot does not vary much between institutions, the speed at which collections were processed was very different because of staffing models used. Countway used a dedicated project or staff archivist, generally with one processing assistant working 17 hours a week per collection. Hopkins used a project archivist and five or six student employees per collection, which required multiple trainings, more project oversight, and greater efforts to standardize descriptive outputs. Having more skilled and experienced processors is more efficient but costs more. Because of the grant's time restraints, Hopkins had its project archivist manage students on collections concurrently. Normally a project archivist handles one collection at a time, working on multiple collections mainly at the beginning and end of projects. When Hopkins does not have a project archivist, outputs are further reduced. Students get assigned one project at a time, often over multiple semesters, with numerous breaks in between. Opening hidden collections and making dents in backlog require stable, professional staffing.
- **Researchers whose work is supported by using records containing health information about individuals need more robust descriptive information to inform their decision making.** One approach we can take as archivists is to introduce in our description more of the variables researchers are looking for. In our online survey, we provided a list of elements archivists could incorporate into collection descriptions, asking respondents

5 Application forms for access to individually identifiable health information are available on the Chesney Medical Archives website.

to rank them not very useful, somewhat useful, very useful, or does not apply. Fifty-one respondents answered, with some surprising results. For example, respondents ranked providing the date span of records most useful. Given that *Describing Archives: A Content Standard* requires a date statement for minimum-level description, we, as archivists, already provide what respondents rank as the most useful information, or perhaps a surprising number of record descriptions still do not include dates. Other descriptions respondents ranked as useful (after date span) were: patient diagnosis or condition, geographic region, patient gender, patient race, duration of treatment, names of procedures and prescribed medication, names of medical devices used, average age of patients when treated, names of treating physicians or surgeons, and presence of genetic information.

Similarly, we asked what kinds of records researchers would be most interested in knowing were in a collection. Patient histories and case files ranked the highest, with informed consent and autopsy records next, and insurance (and by extension billing and coverage records) ranking last. As archivists, familiarizing ourselves with these kinds of records so that we can better identify them in our finding aids would benefit researchers.

Finally, respondents found folder-level scope notes indicating the types of restrictions that applied to a particular folder to be highly useful. Prior to the grant, Countway used series-level restriction statements to explain why folders were access-protected, and provided opening dates. However, because of the multiple types of restrictions that can be encountered in a series, for the collections processed for the Private Practices project, Countway included a statement explaining the reason for the restriction in a folder-level scope note for each restricted folder. Researchers specifically interested in health information can then better target folders of research interest. Not only is this useful to researchers, it has proved useful to the Center's Public Services staff, who no longer need to recall folders to figure out which restrictions apply. Subsequently, Countway adopted this practice for all collections processing. As a HIPAA-covered entity, Hopkins's approach had to be more granular. Staff screened documents at the item level for protected health information and other confidential material such as student and personnel information. They then redacted such information from descriptions, and identified documents that contain such information in the finding aid.

As a result of this work, Countway is testing the best practices for describing collections containing health information about individuals on the Dwight E. Harken papers, 1911–1993.[6] Countway will share the finding aid with members of the Medical Heritage Library and other constituents who provided comments on the recommendations. At Hopkins, the project highlighted the inefficiencies of using undergraduate student assistants. These included student turnover, schedule changes, limited hours during the spring and fall semesters, and the labor-intensive training and supervision that more senior staff must provide for student employees. Hopkins is reviewing its staffing model and will compare metrics on future processing projects that use different staffing models such as relying on undergraduate students for more limited tasks.

6 The finding aid for the Dwight E. Harken papers, 1911–1993 (inclusive), 1940–1975 (bulk), B MS c118, is now available online at http://oasis.lib.harvard.edu/oasis/deliver/deepLink?_collection=oasis&uniqueId=med00207.

Next Steps to Enable Access

We conclude that three steps are needed to enable access to the collections in question.

1. We need to raise awareness among both archivists and researchers that collections documenting health are hidden and endangered.

Because of access anxiety on the part of archivists and their repositories, these collections remain hidden and at risk of destruction. Many repositories refuse to deliberately collect patient-related materials because they do not have the capacity to manage access. Local libraries and historical societies are reluctant to accession collections that may be subject to privacy concerns. They may not have adequate staffing or training to handle requests for access to restricted records. The penalties for HIPAA breaches may pose an unacceptable risk for repositories. It is easier to say no than to invest the resources necessary to make these collections accessible to researchers. Even repositories at large academic health centers whose mandate is to document the history of medicine may not have the resources to accept large runs of medical records after they are no longer required for active patient care activities. Medical records generated by centralized hospital medical records divisions are massive. They constitute millions of records, with significant storage costs. We cannot expect that every medical record can be preserved. Repositories need to have access policies in place that enable research use to justify preservation costs. Scholars need to overcome their access anxiety, and push to gain access to these collections to justify the need to collect these materials. Both archivists and scholars will need to make a commitment to advocate for preserving and using collections documenting the health of our populace.

2. Archivists need to make descriptions available so that researchers can request and use these collections.

Archivists should be advocates for both the collections and the researchers who produce valuable scholarly work using them. Archivists may not even be aware of all the patient-related records they may already have in their own repository, if these materials are unprocessed and hidden. Hidden patient records may not show up in catalog searches, so public services archivists do not know to refer interested researchers to them. Rather than branding these collections restricted and off limits or remaining ignorant to their existence, archivists need to discover these materials and then commit to facilitating access to them. Archivists must increase their awareness of what HIPAA and state medical records laws do and do not allow. HIPAA includes provisions for access to protected health information for research purposes when it is necessary for research and the researcher has a plan to protect it. Archivists need to better familiarize themselves with the research provisions of HIPAA and state medical records laws. They should then insist on becoming part of the review process by serving as members of privacy boards and institutional review boards that can offer waivers of authorization to allow researchers regulated access to protected records in compliance with HIPAA, state laws, and institutional policies.

Once they become aware of the existence of health collections, archivists need to work with historians and other researchers to appraise the research value of these records and advocate for preserving the most significant collections. This will enable the creation of new knowledge, using historic medical records as primary source materials. Before researchers can produce scholarly

works using health information, they need to know these collections exist. Our best practices document (Appendix A) offers guidelines to help archivists describe these holdings in a privacy-aware manner that provides researchers with the qualitative information they need to make decisions about pursuing research activities with these records. More repositories need to overcome their access anxiety and describe these holdings. Protecting privacy is a responsibility shared by archivists and researchers.

Parallel to the CLIR project, Hopkins conducted a citation study of the scholarly output of the 243 researchers who, between April 2003 and July 2014, applied for Privacy Board waivers to use HIPAA-protected Hopkins holdings. We hope this study's findings demonstrate measurable data on what scholarship can be produced if repositories develop the infrastructure to enable access to restricted records in a privacy-aware and HIPAA-compliant environment.

3. Archivists and researchers need to work together to advocate for changes in federal and state law that balance individual privacy protections with the need for scholarly access to create new knowledge in the history of medicine.

In August 2014, the Society of American Archivists (SAA) adopted a HIPAA issue brief that outlines advocacy efforts that the society endorses at the federal, state, and institutional levels. Phoebe Evans Letocha and Lisa Mix worked with SAA's Committee on Advocacy and Public Policy and the Science, Technology and Healthcare Roundtable to present this issue brief to the SAA Council (SAA 2014).

The issue brief outlines a series of recommended changes in HIPAA at the federal level, in state medical record laws, and in practices at the institutional and professional level. The 2013 changes in the HIPAA Privacy Rule enacted because of the passage of the HITECH Act include a change in the definition of protected health information to exclude information about individuals deceased for more than 50 years. In 2005, archivists Nancy McCall and Steve Novak testified in favor of this change before the National Committee on Vital and Health Statistics, and SAA endorsed this new definition during the 2010 comment period to change the Privacy Rule.

This was a welcome advocacy accomplishment, but more changes are needed: a) to provide a date from record creation at which records would no longer be protected in cases where the death date of an individual is unknown; b) to allow easier access to protected health information for family members conducting medical genealogy research; c) to clarify the extent to which archival repositories that are not part of covered entities and that have health-care-related holdings are subject to business associate agreements; and d) to make it clear that individually identifiable information and photographs that appeared in publications or other public venues are not considered protected under the Privacy Rule.

At the state level, medical record statutes need to be aligned with federal regulations to allow for standardization. Archivists and historians also need to turn their attention to advocacy efforts to propose changes in state laws that would enable research using medical records. At the institutional level, through our professional organizations including SAA, Archivists and Librarians in the History of the Health Sciences, and the American Association for the History of Medicine, archivists and researchers should communicate and collaborate to develop best practices and promote a common research agenda that

makes these collections available for scholarly use. Collaborations such as this one, between Hopkins and Countway through the Medical Heritage Library-sponsored CLIR project, enable the creation and promotion of best practices for processing description, and research use of these collections.

References

Alan Mason Chesney Medical Archives, HIPAA compliant application forms. Available at http://www.medicalarchives.jhmi.edu/hipaaform.html.

Evans Letocha, Phoebe. 2014. "How Privacy-Aware Descriptive Practices Can Promote Access and Use of Protected Information." Presentation at Society of American Archivists Annual meeting. August 14, 2014. Available under "Presentations" at http://www.medical-heritage.org/announcements-and-articles/.

Lawrence, Susan. 2007. Access Anxiety: HIPAA and Historical Research. *Journal of the History of Medicine and Allied Sciences* 62 (4): 422-60.

Novak Gustainis, Emily. 2014. "Processing Collections with Protected Information: A Comparison of HIPAA and Non-HIPAA Covered Entities." Presentation at Society of American Archivists Annual meeting, August 14, 2014. Available under "Presentations" at: http://www.medicalheritage.org/announcements-and-articles/.

SAA (Society of American Archivists). 2014. Issue Brief: Health Information Portability and Accountability Act. Approved by the SAA Council August 2014. Available at: http://www2.archivists.org/statements/issue-brief-health-information-portability-and-accountability-act.

Collections cited:

Arnold S. Relman papers, 1953-2011: http://oasis.lib.harvard.edu/oasis/deliver/deepLink?_collection=oasis&uniqueId=med00188.

Barbara Starfield Collection, 1948-2011: http://www.medicalarchives.jhmi.edu/finding_aids/barbara_starfield/barbara_starfieldd.html.

B. Frank Polk Collection, 1972-1990: http://www.medicalarchives.jhmi.edu/finding_aids/frank_polk/frank_polkd.html.

Elmer V. McCollum and Harry G. Day Collection, 1881-2003: http://www.medicalarchives.jhmi.edu/finding_aids/elmer_mccollum/elmer_mccollumd.html.

Erich Lindemann papers, 1885-1991: http://oasis.lib.harvard.edu/oasis/deliver/deepLink?_collection=oasis&uniqueId=med00191.

Harvard School of Public Health's Department of Biostatistics records, 1981-2009: http://oasis.lib.harvard.edu/oasis/deliver/deepLink?_collection=oasis&uniqueId=med00187.

Oliver Cope papers, 1891-1992: http://oasis.lib.harvard.edu/oasis/deliver/deepLink?_collection=oasis&uniqueId=med00189.

Stephen W. Lagakos papers, 1979-2009: http://oasis.lib.harvard.edu/oasis/deliver/deepLink?_collection=oasis&uniqueId=med00185.

William George Hardy and Miriam Pauls Hardy Collection, 1875, 1930-2008: http://www.medicalarchives.jhmi.edu/finding_aids/william_hardy/william_hardyd.html.

APPENDIX A: RECOMMENDED PRACTICES FOR ENABLING ACCESS TO MANUSCRIPT AND ARCHIVAL COLLECTIONS CONTAINING HEALTH INFORMATION ABOUT INDIVIDUALS

Phoebe Evans Letocha, Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions

Emily R. Novak Gustainis, Center for the History of Medicine, Francis A. Countway Library of Medicine

1 February 2015

About

The following recommendations were developed by the Alan Mason Chesney Medical Archives of the Johns Hopkins Medical Institutions and the Center for the History of Medicine at the Francis A. Countway Library of Medicine in an effort to enable access to manuscript and archival collections containing protected health information (PHI) and other types of access-protected records containing health information about individuals. The recommendations serve to underscore the importance of making available primary sources for health information about individuals to historians and other researchers to inform the history of American medicine and serve as a foundation for evidence for policy-shaping works. When these collections remain hidden and inadequately described, they are at greater risk for destruction, thus impeding future archival research that furthers our collective understanding of health and disease. Facilitating access involves striking a balance between the privacy concerns of living individuals and the greater public good that can be accomplished by scholarship.

This work was made possible through the generous funding of the Council for Library and Information Resources' *Cataloging Hidden Special Collections and Archives* program (2012: *Private Practices, Public Health: Privacy-Aware Processing to Maximize Access to Health Collections*).

The recommendations need not be pursued in sequential order. Repositories are encouraged to pursue some policy recommendations concurrently or to test one of the many descriptive enhancements. It is the hope of the authors that these recommendations will help alleviate many of the concerns repositories have related to collecting and preserving health services records, especially those that are not affiliated with hospitals or medical schools.

Determining an Institution's Status and Policy Needs

- Repositories should train staff to recognize individually identifiable health information, regardless of whether or not they are entities covered by the Health Insurance Portability and Accountability Act (HIPAA). Repositories that are HIPAA-covered should provide training to familiarize staff with legal requirements.
- Repositories should survey their holdings to determine the extent to which they include individually identifiable health information that may be protected by federal or state laws.
- Repositories should consult with their administration and legal counsel to determine their status under HIPAA; the Federal

Common Rule for the Protection of Human Subjects; and their state's medical records laws.

- Repositories should document their status under such rules and statutes and determine their institution's risk tolerance, since (1) laws such as HIPAA allow institutions to be more restrictive than the law requires, and (2) some donor agreements may require restrictions beyond that which is covered by HIPAA.
- Repositories should create intra-organizational partnerships to align policies, for example, among special collections repositories at the same institution, medical records/health information management departments in hospitals, and/or institutional records management offices.
- Repositories holding records of outside institutions that contain individually identifiable health information should consult with the depositing institution and with their own legal counsel to determine whether housing the records would make the repository subject to HIPAA business associate agreements.
- Repositories should review the types of requests that they receive for access to individually identifiable health information and develop access review processes relevant to the type of use requested, such as medical genealogy, biography, and research as defined by HIPAA and the Common Rule.

Implementing Policy and Fostering Process Transparency

- Repositories, to the extent possible, may want to create an impartial Access Board or Privacy Board or consult with an Institutional Review Board (IRB) to review applications for access to protected health information and medical records in their holdings. An archivist with knowledge of the holdings should be designated to be part of the review process, either as an advisor to or as a member of the review board. If no Access Board is possible, repositories should be prepared to explain why access can be granted to some users and not others.
- Repositories should document their decision-making processes and policies and apply them consistently. Decision trees may be helpful tools to review access decisions (see Johns Hopkins examples).
- Repositories should publish their access and use policies on their websites and should provide copies of any application forms online; researchers should be reminded that publishers may also have their own privacy requirements as a condition of accepting a manuscript for publication.
- Repositories should clearly articulate the steps a researcher or other user would need to take to apply for access and the application workflow, so that users know how far in advance they will need to make an application before they may be granted access.
- Repositories may wish to provide model applications or a process by which applicants can ask questions or seek guidance on the application process so that they can successfully complete the application.
- Repositories should create a user agreement for patrons to sign that communicates personal liability for the misuse or distribution of health information about individuals.

Communicating the Nature of Restrictions

- Repositories should provide non-technical information on their websites about the kinds of access restrictions their users will encounter when considering the use of records, regardless of whether restrictions are imposed by: Federal law (HIPAA, FERPA); U.S. government records laws; state law; gift agreement; deposit agreement; or institutional policy.
- Repositories should provide at least one example of each of the restrictions found in their collections using a published or otherwise publicly available finding aid or catalog record to illustrate the restrictions.
- Repositories should explain where users can find information about access restrictions, such as publicly accessible catalog records, online finding aids, or published inventories. Repositories should provide information about the gaps in systems where information is generally provided (such as restrictions only being noted in catalog records for collections that have been processed), as well as overtly state when information about access restrictions is only available through consultation with Public Services staff.
- Repositories should embed information regarding the presence of access restrictions at all levels of hierarchical description. Collection-level access descriptions may alert users to the presence of restrictions, but it is series, subseries, and folder-level notices regarding access status that enable users to understand which restrictions apply to records of interest.
- Repositories should clearly articulate their policies regarding citation. Access Board and IRB applications should clearly indicate if citation is permitted, and if so, repositories should have specific examples for citing records in collections that are not accessible without access approval and, if the collection is unprocessed, whose physical organization may change in the future.
- Repositories may want to allow and encourage users to deposit a code key to medical records and other protected records that cannot be cited by identifiers, such as patient name or medical record number, without authorization. Repositories should clearly state in finding aids when records have been redacted or removed from the collection.

Describing Records to Best Enable Discovery and Access

The following recommendations are intended to illustrate the rich descriptive information that archivists can offer without revealing patient names or other identifiers. When selecting descriptive approaches, processors should balance the needs of their research communities with local processing practices to determine which of the following descriptive enhancements could improve discoverability and use of their collections.

- When describing collections containing health information, communicate the specific record formats in which health information is found. A developing list of different kinds of records containing health information and their scope may be found here. Examples include: admission records; autopsy records; case files; diagnostic indices; doctor-patient correspondence; medical records; patient histories; prescription logs; surgical logbooks; and specimens. If you are not sure of the kind of record you have, try to create a redacted copy of the record (or a page or two from a

volume) and consult an archivist or librarian who more routinely encounters these types of records.

- Descriptions should overtly state if a collection is a part of a much larger, original group of records, as well as inform users as to what happened to the rest of the records or where they may be found. (For example, when a collection consists of 20 boxes transferred to the archives as a representative sample from an original 100 boxes of records, indicate that the remaining 80 boxes were destroyed per institutional policy.) Specimens related to a collection that are housed elsewhere should be indicated, regardless of whether or not they can be accessed.
- Processors should identify when records were created for a specific research study or when doctors assembled sets of patient records as source material for specific publications.
- Processors should record types of commonly collected information about patients in the records, such as diagnoses, names, dates of birth/death, and ages at time of treatment. As time or expertise permits, processors should sample the records and incorporate into the description patient-related information, such as marital status, number of children, race, ethnicity, occupation, and place of residence or employment; and treatment-related information, such as the names of frequently mentioned doctors, surgeons, midwives, mental health professionals, and/or dentists encountered, pharmaceuticals, types of medical treatments and procedures, and instrumentation and devices used. A developing list of variables may be found here.
- Because processing methodologies vary from repository to repository, processing information in finding aids should include how record descriptions were created, such as through a percentage of records sampled per container or per alphabetical or numeric run.
- Repositories should enable opportunities for user enhancement of collection descriptions, particularly for unprocessed or infrequently used collections. A survey instrument or quick conversation with a researcher may help contextualize records, add to lists of procedures or treatments employed, or enrich collection-level descriptions of holdings. Users may also provide examples of "the patient's own words" that can be included anonymously in finding aids to help characterize records. Similarly, health care providers familiar with the creation of specific categories of patient record types can help contextualize records based on their clinical experience of how records are used. Health care providers may also be able to decipher medical shorthand or abbreviation unfamiliar to archivists who don't have specialized medical training or familiarity with local institutional terms.
- Repositories should consider digitally imaging redacted versions of records and embedding them in finding aids in order to visually communicate how information is organized in the records. Repositories can also consider embedding blank versions of survey instruments, commonly found forms in medical records, pages from codebooks, and protocols.

Grinnell to GUIDs: Connecting Natural Science Archives and Specimens

Christina Fidler, Museum Archivist, Museum of Vertebrate Zoology, University of California, Berkeley

Barbara Mathé, Museum Archivist and Head of Library Special Collections, American Museum of Natural History

Rusty Russell, Program Director for Collections & Informatics, United States National Herbarium, Smithsonian Institution

Russell D. "Tim" White, Director of Collections and Operations, Peabody Museum of Natural History, Yale University

Abstract

Four recipients of Council on Library and Information Resources (CLIR) Hidden Collections grants in natural science museums—two archivists and two scientists—discuss their view of and approach to access for resource discovery, the effects of digitization, integrated access in aggregators, and the issues of long-term digital preservation. The long-term goal is to integrate and link access to digital data and images across natural science institutions, whether libraries, archives, or museums.

> "It is quite probable that the facts of distribution, life history, and economic status may finally prove to be of more far-reaching value, than whatever information is obtainable exclusively from the specimens themselves."
>
> —Joseph Grinnell (1910, 166)

Joseph Grinnell, the first director of the Museum of Vertebrate Zoology at the University of California, Berkeley, foreshadowed—by more than a century—the growing recognition within today's natural science community of the value of all information found and gathered during a collecting event.

In 1908, Grinnell developed and implemented a detailed protocol for recording field observations. These integral materials, gathered along with specimens, contain extensive information that may not appear on labels attached to or stored with collections objects. They may include detailed accounts of individual species' behaviors; annotated topographic maps; photographs of collecting sites; observations that did not result in specimen collection; interactions with local or indigenous populations; and other data, such as weather conditions, vegetation types, vocalizations, and various evidence of animal presence in a given locale.

Integrated Digitized Biocollections (iDigBio), part of the National Resource for Advancing Digitization of Biodiversity Collections (ADBC) funded by the National Science Foundation, is an aggregator that will allow data and images for millions of biological specimens to be made available electronically. In March 2014, iDigBio, Yale University, and the Field Book Project sponsored a workshop focused on digitizing original source materials associated with scientific collections. This gathering reflected a growing momentum toward providing access to all types of resources related to natural science, not only specimens and species publications but field recordings—in

the form of books, notes, sketches, correspondence, and audiovisual materials along with records of the research conducted using these things. Although there is still a crying need for the time-consuming work of processing archival collections, the long-term goal is integrated access to natural science information—data and images—wherever they are held. As the development of linked data applications using globally unique identifiers (GUIDs) proceeds, this goal can become a reality.

▶ Christina Fidler, Museum Archivist, Museum of Vertebrate Zoology, University of California, Berkeley

Project title: Cataloging Hidden Archives of the Museum of Vertebrate Zoology: Increasing Integration and Accessibility for Interdisciplinary Research (2011)

Access (resource discovery): Why and how have you made it possible for your archival collections to be found by those looking for them (or perhaps by those who did not know they existed)?

Beginning in 2003, the Museum of Vertebrate Zoology (MVZ) at the University of California, Berkeley, executed the Grinnell Resurvey Project, a large-scale resurvey of ecological transects in California. These transects define different specified areas where the abundance and distribution of faunal populations can be measured. To conduct this extensive resurvey, researchers relied primarily on the archival field notes that document the floral and faunal conditions of the same transects roughly 100 years ago. It was an astonishing testament to the importance of these historical data. The resurvey was a catalyst for the museum's CLIR Hidden Collections grant, which began in 2012. I am the archivist hired on the CLIR grant, which is dedicated to establishing a formal repository and cataloging the archival field note collections. I work under the supervision of the grant's co-principal investigators, Carla Cicero and Michelle Koo.

In 1910, Joseph Grinnell, the museum's first director, wrote with alarming accuracy, "At this point I wish to emphasize what I believe will ultimately prove to be the greatest value of our museum. This value will not, however, be realized until the lapse of many years, possibly a century, assuming that our material is safely preserved. And this is that the student of the future will have access to the original record of faunal conditions in California and the west, wherever we now work" (Grinnell 1910, 166).

As mentioned earlier, Grinnell developed a means for capturing data while collecting and observing in the field. The Grinnell method of taking field notes has been practiced over the last century and is still being practiced much as it was originally at the MVZ. This has resulted in more than 700 volumes of field research materials. The field notes are traditionally bound and made available for research in the museum's library. Through the MVZ CLIR grant, I had the responsibility of formalizing the MVZ's archival repository and implementing a program to identify, process/describe/catalog, integrate, and make available the field notes and the many other archival collections housed at the MVZ.

I quickly learned that traditional archival best practices in access and description were neither sufficient nor appropriate in the natural history museum context. Although archives professionals are critical to the curation of historical materials, there are some nuances in the needs of natural history museums. Scientific data do not fit categorically into the traditional archival concepts of active and inactive records. Scientific

data are treated and accessed much like active records, and this presents many challenges with security and preservation.

Additionally, traditional specimen cataloging is both granular and dynamic with structured relationships across individual objects. This is a departure from the current shift in archiving toward the approaches described in Mark Greene and Dennis Meissner's 2005 groundbreaking paper "More Product, Less Process." The MVZ Archives address these inherent conflicts while also using traditional archival description and dissemination. The result is mutually beneficial. Reference requests have increased exponentially since the MVZ Archives began publishing finding aids in the Online Archive of California (OAC), and our finding aids are more robust and dynamic to fit the needs of researchers.

To achieve our goals in the CLIR grant, the MVZ Archives executed the following strategy:

1. Survey all museum materials not actively cataloged by staff curators. Converse with personnel about papers that they may have inherited from former staff and faculty.
2. Install and implement Archivists' Toolkit.
3. Prioritize collections for processing and preservation.
4. Create collection records for all identified archival collections-collection building.
5. Process. Process. Process.
6. Establish portals and disseminate information.

As we worked, we developed administrative documentation surrounding archival policies. These include a mission statement, collection policy, use policy, image use and permission forms, accession forms, processing manual, imaging protocols, and numerous other administrative resources.

Initially, the materials were housed throughout the museum. Personal papers were traditionally passed down to incoming faculty and staff. Field notebooks were stored in three different locations, depending on their binding status. After an extensive survey of the museum, the Archives staff began building the collection and bringing the materials together. We identified approximately 600 linear feet of archival materials. The museum had once held a large reprint collection in a dedicated room, and the Archives staff processed the reprint collection and made room for the incoming archival collections. The Reprint Room is now a dedicated room in the Archives.

Once we established the foundation for the Archives, we explored how best to integrate archival description with specimen data. The MVZ catalogs its specimens, observations, and other biodiversity records in Arctos, a multi-institutional collection management system for natural science collections. This collaborative solution contains more than 3 million natural history museum records. In close collaboration with our staff curators, we developed a process by which to catalog localities and specimen records, and create external relationships to Arctos. I detail this process later in this paper.

The MVZ Archives has published 42 finding aids on the OAC, 35 of which are fully processed collections with encoded archive description (EAD) finding aids. Although this is only 10 percent of the collections, we anticipate that we will publish collection MARC records for the remaining 90 percent of collections by the end of the grant period.

We also provide numerous work opportunities to students. We offer internships for processing collections to graduate students attending the San Jose School of Information. We also offer exposure to archival practices, including rehousing,

indexing, exhibit planning and development, and processing of small collections to our university's Undergraduate Research Apprentice Program; these students write about their experiences on the MVZ Archives Blog.

As we process collections and publish finding aids to the OAC, our reference questions are increasing exponentially. In the first year of our CLIR grant, we documented 8 reference requests; last year, there were 31, a 300 percent increase. Six months into the new fiscal year, we had 24 reference requests. Our reference questions vary across disciplines. In addition to the traditional users from the biological sciences, there are a growing number of requests from those in the humanities.

The CLIR grant allowed the MVZ to establish intellectual control over its valued archival collections and develop innovative approaches for integrating the archival collection with its specimen collection while facilitating discovery and access. We are motivated to promote and create opportunities for interdisciplinary research. We recently hired John Deck, a programmer whose previous projects include the Moorea Biocode Project, Biological Science Collection (BiSciCol) Tracker tools, and BerkeleyMapper. John's experience in informatics standards and the semantic web to integrate metadata across domain-specific databases and data workflows is ideal for creating a robust method for the online search and delivery of archival content that will be the foundation for new methods for archival data visualizations and discovery.

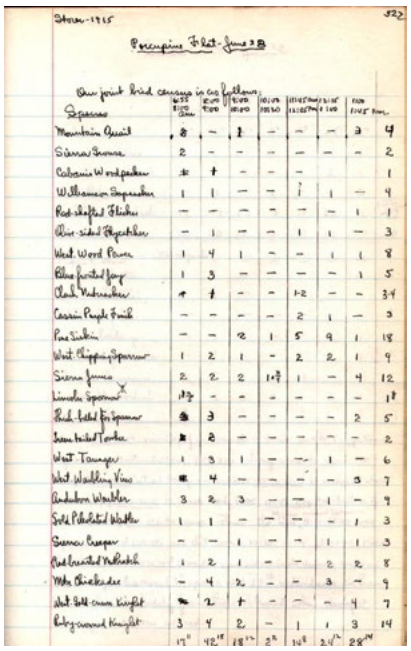
Digitization: Why should we digitize? How can digital versions of these materials enhance access to the material? What are the issues involved in digitization (including crowd sourcing facilitated by digital versions of the resources)?

There is an urgent need to make the data contained within field notes available for assessing the impacts of rapid and ongoing environmental change. As more MVZ field notes collections are processed, researchers from across California and the country are requesting access to digitized copies. It is somewhat challenging to meet this need with just under 50 percent of our field notes scanned. Each volume contains an average of 200 pages, and as we noted earlier, there are approximately 700 bound volumes.

Furthermore, field notes contain critical observational data that our scientists can quantify and catalog. Typically, the number of observations recorded in field notes at a given locality is larger than the number of specimens collected and cataloged in Arctos. Quantifying observational data and creating observational records in Arctos would give researchers a more comprehensive understanding of changes in faunal conditions across the twentieth century.

In the example shown on the next page, three times as many bird observations are made in the field notes as there are specimens for the same location.

Field notes are handwritten and are not generally good candidates for current OCR (optical character recognition) tools. However, there is a strong citizen science community that could be engaged for transcribing field notes. With transcriptions, MVZ personnel would be able to extract data; catalog them; and relate them to specimens, localities, and other primary sources. We are continuing to explore and test different mechanisms for field note transcription.



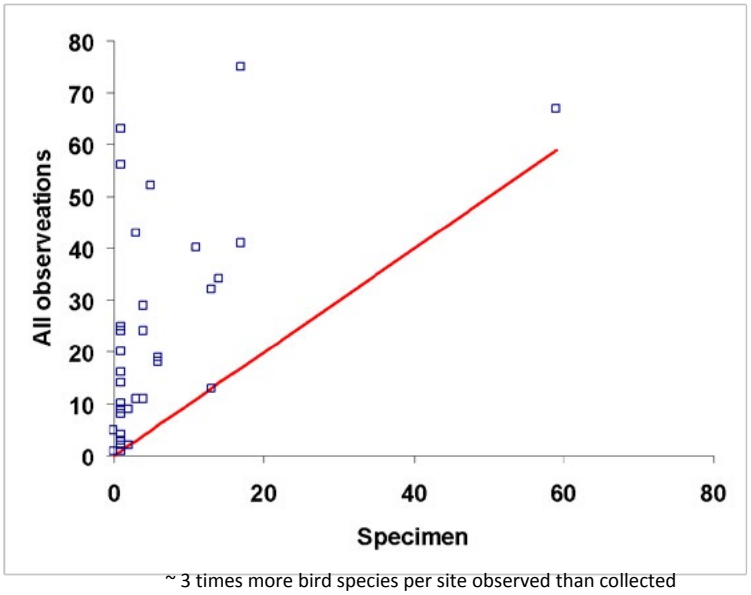


Fig. 1: Comparison of observed bird species as recorded in field notes (left) with number of specimens collected.

Why is integrated access to data and images across scientific disciplines, institutions, and type of resource (e.g., via aggregators like iDigBio and the Biodiversity Heritage Library) important? How can this be accomplished?

Each specimen record in Arctos contains metadata describing the collecting event of the specimen, including who collected it, when, and where. The field notes contain the same data, but they include more context for the conditions of the collecting event.

Because MVZ researchers use the Grinnell method of taking field notes, the MVZ field note volumes are structured in a predictable and consistent way. Field notes typically consist of three sections: specimen catalog, journal, and species accounts. Each section has its own specified format, including placement of names, localities, and dates. Because the materials are consistent in their formats, we could reliably expect to extract data surrounding collecting events:

- *Localities.* Arctos has its own geographic name structure. The MVZ staff use this structure as the name authority, which will support future georeferencing of the materials.
- *Specimens.* We use a Related Materials Note in our finding aids to connect researchers to specimens described in a given field notebook section.

These connections allow researchers to view the inherent relationships between the field notes data and the specimens.

We are also moving toward making these connections within Arctos. This would expose the archival collections to three distributed data networks for vertebrates: MaNIS, HerpNet, and ORNIS. It would also make the materials available to the Global Biodiversity Information Facility (GBIF).

Digital management and preservation of this work, the digital data, and images: What is the current status of digital data management and preservation in your institution?

Unquestionably, digital data management and preservation are major challenges facing researchers and archivists alike. In addition to an internal image vault, the MVZ stores its digitized archival content offsite. The Texas Advanced Computing Center (TACC) provides storage of our digitized archival materials and audio, including all TIFF preservation files and all access copies.

The MVZ Archives is experiencing an influx of deposited floppy disks with data created on software that is now obsolete. And there is growing concern in general over the management of data in the scientific community. The National Science Foundation now requires that all grant proposals provide data management plans that include outlines for sharing and archiving data. Specifically, plans address file formats for long-term access, documentation for data interpretation, and copyright status. Although software and file type obsolescence is happening, data management is a proactive approach to address these problems. This is an area where archives can lend best practices in preserving data integrity, storage, and preservation. At the University of California, Berkeley, the Bancroft Library is presently implementing a digital forensics system to secure born-digital collections. This program is in its infancy, but we expect to continue exploring and improving mechanisms for digital preservation and access to born-digital collections.

▶ Barbara Mathé, Museum Archivist, American Museum of Natural History

Project titles: For the People, for Education, for Science: the American Museum of Natural History Archives (2010); and Hidden Connections: Linking Museum Expeditions, Scientists and Collections at the American Museum of Natural History (2012)

Access (resource discovery): Why and how have you made it possible for your archival collections to be found by those looking for them (or perhaps by those who did not know they existed)?

Historically and to this day, fieldwork is the source of natural science research collections. After first surveying and creating records for the scientific archives in the American Museum of Natural History (AMNH) with funding from CLIR and the Institute of Museum and Library Services (IMLS), we are using our second Hidden Collections grant to focus on AMNH archives from museum expeditions. We want to better document and relate collecting events, noting particularly the people who were on the expeditions at the same time, place, and circumstances when the collections were made. As Grinnell pointed out, this information is essential to the full significance of the specimen collected, whether for the systematic study of species or for use in ecological research. AMNH scientists have not used Grinnell's data fields on a consistent basis, although currently there is a growing interest in doing so. For example, a scientist returning from Papua, New Guinea, on a recent Explore 21 field trip was just extolling the virtues of Grinnell-formatted field notes.

Founded in 1869, the AMNH will soon celebrate its 150th year. I am the museum archivist, based in the research library and charged with coordinating efforts with others in the museum who

manage archival collections. I report to the library director, who reports to the provost, who is senior vice president for science. There are five divisions in the administrative unit of Science, of which the library is also an administrative unit. The divisions are Anthropology, Physical Sciences, Invertebrate Zoology, Vertebrate Zoology, and Paleontology (both vertebrate and invertebrate). The Division of Anthropology has a professional archivist, and her time is split between archives and managing traveling loans. Other divisions either have retired scientific assistants or curatorial associates, whose chief responsibility is the management of scientific collections, overseeing their archives.

The specimens and objects in the scientific collections now number more than 33 million. We can estimate 25,000 linear feet of archives across the museum, in the science departments, and in the research library. This estimate was the result of the timely confluence of our first CLIR grant in 2010 to create minimal catalog records for the archives in the library, with an IMLS risk assessment grant that allowed us to survey the archives in the science departments. These two projects resulted in the creation of some 3,275 minimal-level catalog records for the archival collections in AMNH Science, 1,400 in the library and 1,875 in the science divisions. We call these minimal records "skinnies." Although they are mapped to MARC and to EAD, and published online so that people might discover their existence, they are hardly full finding aids. Our first CLIR grant also resulted in 21 fully descriptive finding aids.

Many of these archival collections, especially the field books and notes, relate directly to the scientific collections. This is a gold mine of related data of all kinds, but the knowledge of the existence of, and particularly the relationships across, the materials is still largely dependent upon the accumulated institutional knowledge of the staff. One key to integrating access across archives and scientific collections, that are described only minimally might be found in the relationships between expeditions and the many scientists who traveled on them. For example, a researcher looking at the circumstances surrounding the collection of a bird could find that a herpetologist on the same expedition had documented the environment at the same place and time the bird was collected. With a minimal EAD collection record of that herpetologist's field book, listing only her name, the information would be hidden. But if a record for the expedition listed both the name of the herpetologist and the name of the ornithologist who collected the bird, each assigned an individual identifier and related by the identifier of the expedition, then the connection can be made. With such an enormous collection, approaching access through minimal records that can be linked and then enhanced over time seems the best approach. We just cannot describe every record one by one and expect to get results in a reasonable period of time. When we began the second CLIR project, we anticipated the development of linked data applications and are looking at a number that have recently been launched.[1]

Using Encoded Archival Context–Corporate Bodies, Persons, and Families (EAC–CPF) to create records that can normalize the names of the expeditions and those associated with them will allow us to integrate access to collections (whether archival or specimen), even if they are described on a minimal level. We came to this approach because an earlier library project had

1 As of early November 2015, the AMNH team will be working to develop a proof of concept for linking data across different collections across the scientific and archival collections in the museum with a triplestore.

documented close to 1,000 museum expeditions. We updated that work with the information found in vertical files for expeditions and for people associated with the museum. With the help of many volunteers over the years, the library had created spreadsheets that captured basic data. Happily, the data fields were a close match to fields in the EAC–CPF data standard, and we set out to verify and use this data as a starting point for the project. Incidentally, we have similar data for museum departments, halls, and temporary exhibitions.

So we embarked, once again, on our effort to create minimal level records, this time for entities, specifically AMNH expeditions and personal names. We have also created templates and protocols for our interns to use to create fully developed EAC–CPF records, recognizing that the biographical or historical note for the entity can then be repurposed for any collection that has an EAD record, using the EAC–CPF biographical/historical note related collections. For example, there are five archival collections from the Whitney South Sea Expedition, one in the library, one in the Department of Herpetology, and three in the Department of Ornithology. The EAC–CPF record for the entity, "Whitney South Sea Expedition," is being repurposed for the EAD finding aid records for all five collections.

The result is a large number of minimal-level EAD and EAC records to manage, as well as a number of records more fully describing collections and entities. Anticipating this result, we committed to develop a cyberinfrastructure for our archival records in our CLIR Hidden Collections grant proposal in 2012.

The team includes Becca Morgan, project archivist, and Iris Lee, metadata analyst (both veterans of the first CLIR/IMLS project), along with Nick Krabbenhoeft and Bill Levay, who have added their technical insight and expertise as interns and subsequent volunteers. Neither CLIR project would have been possible without our interns and the many volunteers who are listed on our Hidden Collections blog.

After developing our functional requirements for managing the records, we found that there is no archival management system that fulfills our requirements for EAC records, so we are using xEAC. For EAD records, it was a close contest between ArchivesSpace and Atom, but the decision was made to use ArchivesSpace (Lee 2015).

In response to the extensive work we did developing EAC–CPF records, as well as our process methodology and documentation during this project, we were invited to join SNAC (Social Networks and Archival Context), a collaboration between the U.S. National Archives and Records Administration, the Institute for Advanced Technology in the Humanities at the University of Virginia, the California Digital Library, and the School for Information Science at the University of California, Berkeley. According to the SNAC home page, the project "demonstrates the feasibility of separating the description of persons, families, and organizations—including their socio-historical contexts—from the description of the historical resources that are the primary evidence of their lives and work."

We also just received funding from the Leon Levy Foundation to begin cataloging the 3,224 AMNH Field Books identified across the museum. Before the advent of our funding from CLIR and the IMLS, we would not have known how many we have or where they might be found.

Digitization: Why should we digitize? How can digital versions of these materials enhance access to the material? What are the issues involved in digitization (including crowd sourcing facilitated by digital versions of the resources)?

Digitization is the next step in establishing wider integrated access to museum collections of all kinds. Because of the scale of our collections, we are releasing minimal records. If we can attach scans to those records, we will be in a position to engage the citizen scientists in crowd sourcing the transcription of these materials, revealing more detailed information that can be added (and linked to and from) those records.

Not only can making scanned materials available to researchers allow them to work remotely, but also they will not have to wait for collections to be fully processed before they begin their research. Obviously, there are privacy and other issues relating to the content of the collections, but archivists will learn how to manage them.

Why is integrated access to data and images across scientific disciplines, institutions, and type of resource (e.g., via aggregators like iDigBio and the Biodiversity Heritage Library) important? How can this be accomplished?

Natural science collections span libraries, archives, and museums. A species is "named" upon the first publication of its description by the person who found it. For this reason, rare books are active documents in natural science research. The Biodiversity Heritage Library is a consortium of libraries that are digitizing pre-1923 natural science publications. They are also beginning to include digitized field books. It is not unreasonable to foresee integrated access in aggregators like the Biodiversity Heritage Library and iDigBio, as well as other aggregators like VertNet and MorphoBank, as a long-term goal to make it possible for researchers to seek information across institutions, whether libraries, archives or museums. And by using globally unique identifiers and linked data applications, it may soon be possible to search across the aggregators.

Digital management and preservation of this work, the digital data and images: What is the current status of digital data management and preservation in your institution?

The AMNH is fortunate to be hosting a National Digital Stewardship Resident who has been interviewing the museum's scientific curators and staff about the state of digital preservation and management of their scientific research data. The data include enormous files, such as genomic sequences; three-dimensional CT scans of specimens; and high-definition film as well as collection data. Her report will form the basis of our effort to address the very complex solutions involved in establishing and staffing a permanent program for digital preservation and management at the AMNH. In time, we hope it will result in a trustworthy digital repository for the museum's scientific research and collection data in digital form.

Addendum: As a result of the ongoing work done by on the CLIR projects, followed by that done by the AMNH NDSR resident, we are beginning to move onto the next phase of our work to produce a cyberinfrastructure at the AMNH. A poster that we designed and presented at the recent Hydra Connect meeting illustrates a conceptual model of our current information landscape as compared to an sketch of how we imagine organizing our assets into a cyber-infrastructure. The model is based upon the OAIS model and can be seen on the process tab on our AMNH Hidden Collections site: http://images.library.amnh.org/hiddencollections/process/

▶ Russell D. "Tim" White, Director of Collections and Operations, Peabody Museum of Natural History, Yale University

Project title: From DNA to Dinosaurs: The Globalization of Science in America and the Development of a University Natural History Museum (2010)

Access (resource discovery): Why and how have you made it possible for your archival collections to be found by those looking for them (or perhaps by those who did not know they existed)?

Through seven generations of faculty and students at Yale University, the Peabody Museum of Natural History has acquired a treasure trove of historical and scientific documentation that relates to some of the earliest scientists in America (e.g., Benjamin Silliman) and some of the earliest organized (and unorganized) natural history collections. The ability to keep these materials in an organized, accessible, and retrievable fashion has bolstered the value and use of Peabody's specimens and artifacts, and it has led to many new discoveries in the collections and in the field. Moreover, it is important in documenting the history of scientific inquiry.

At the Peabody, archival materials from previous faculty and students have been "managed" and housed within the 10 scientific collection divisions. In some divisions, archival collections are well organized, properly housed in archival quality materials, and cataloged; in others, collections have been at best physically located in a designated area with little or no arrangement. While faculty members are active, their field notes, maps, etc. remain in their possession, generally kept in their offices and labs, and are only turned over to the museum after retirement.

For too long, archival collections at the Peabody have been seen as supporting documentation for specimens and artifacts; they have not been recognized as significant historical documentation on their own. However, field notes, maps, photographs, diaries, and other materials often document the details of a collecting event and scientific discovery. As part of the CLIR project, archival collections within each curatorial division were assembled and assessed for preservation; diversity of materials; relationship to Yale faculty and students; and documentation of specimens, artifact discovery, and use. Access to the cataloged archives and special collections of the Peabody Museum has opened up new avenues for research beyond the bounds of natural science, including the history of early science in America and the exploration of the American West.

Our CLIR-funded project, From DNA to Dinosaurs, allowed us to assess and inventory all of our archival holdings in the 10 curatorial units plus the museum archives, appropriately rehouse materials, and catalog our collections using the EAD standard in a manner that allows for maximum accessibility. In fall 2014, the Peabody launched 78 finding aids that are available on the Yale Finding Aid Database, and another 20 are in production.

With collections from A to Z at the Peabody Museum, including anthropology, botany, geology, paleontology, and zoology, the discipline-specific historical practices need to be evaluated for methods of documentation of the collecting event, the geospatial occurrence, and the geologic age. By using original documentation, we have developed workflows in the curatorial divisions to catalog our specimen and artifact collections to better define best practices for this type of collection, which has led to improved cataloging and informatics, as well as increased productivity.

In 2016, the Peabody will be celebrating the 150th anniversary of its establishment as a Yale University museum with a gift from George Peabody. One of the celebratory events is a book by Richard Conniff, who is writing about the 15 faculty curators and significant events in our history. Use of the Peabody finding aids and unfettered access to the well-organized archives has made discovery for this project feasible and possible—a situation that did not exist prior to this Hidden Collections project.

Digitization: Why should we digitize? How can digital versions of these materials enhance access to the material? What are the issues involved in digitization (including crowd sourcing facilitated by digital versions of the resources)?

Access to digitized versions of Peabody's archival materials enhances our ability to develop methods for better management and dissemination of specimen-level information. Also, it offers a better understanding of the people involved in these collections. The breadth of work of Benjamin Silliman, Benjamin Silliman, Jr., and James Dana in the early nineteenth century helped to define Yale's nascent programs in science education for undergraduate students. Their archival materials provide insights into the development of a teaching program for science at Yale and other insights into early geological investigations and classifications, such as *Dana's System of Mineralogy* (first published in 1837). In the late nineteenth century, O. C. Marsh and A. E. Verrill defined advance training for biological and paleontological students, and the establishment of Yale as a leading center for paleontology and marine biology in North America (Conniff, in press).

Why is integrated access to data and images across scientific disciplines, institutions, and type of resource (e.g., via aggregators like iDigBio and the Biodiversity Library) important? How can this be accomplished?

Within the natural history community, there are well-established tools for disseminating specimen and collecting event information through the Global Biodiversity Information Facility, iDigBio, VertNet, and other aggregators that specialize in collating morphological information through images (e.g., Morphbank, MorphoBank). Through the efforts of the Biodiversity Heritage Library, some specimen- and collection-level documentation is available, but to date, there is no one aggregator that pulls information from specimen, collecting events, and historical documentation into a central repository.

Some institutions have developed search interfaces to look across collections from disparate sources. At Yale University, the Cross Collections Discovery (CCD) aggregates materials from the Peabody Museum, Yale art galleries, and libraries in one search portal, allowing for discovery of vastly different material in one location.

Digital management and preservation of this work, the digital data and images: What is the current status of digital data management and preservation in your institution?

At the Peabody Museum, archives and special collections are cataloged using KE EMu, the same database management application that is used for cataloging and tracking Peabody's specimen and collecting event information. Specimen information is distributed to numerous biodiversity portals, and archival information is published through the Yale Finding Aid Database. With both types of collections being cataloged using KE EMu, we can cross-reference different types

of associated information in a consistent and structured manner. Over the past three years, the Peabody has digitized more than 1,000 cataloged field notes, ledgers, and other forms of original documentation using a robotic book scanner and linked these digitized documents to specimens and collecting events in KE EMu.

One challenge for the Peabody is to disseminate these linked data sets of archival information and specimen or artifact information to the broader community. Another challenge is to extract information from these types of digitized legacy documents (e.g., institutional catalogs) and original documentation (e.g., field notes). Transcribing targeted information such as people's names, geographic places, taxonomic names, geological ages and rock units, and dates is a goal. Crowd sourcing this kind of activity also seems highly desirable.

▶ Rusty Russell, Program Director for Collections and Informatics, United States National Herbarium, Smithsonian Institution

Project title: Exposing Biodiversity Fieldbooks and Original Expedition Journals at the Smithsonian Institution (aka The Field Book Project) (2009)

Access (resource discovery): Why and how have you made it possible for your archival collections to be found by those looking for them (or perhaps by those who did not know they existed)?

This question gets at what was precisely the inspiration for The Field Book Project (FBP). My frustrating attempt to locate field notes from the U.S. Exploring Expedition (1838–1842) led me to more than a dozen institutions. Although there were some electronic services that provided information about these materials, most content was found through inferential searching or simple guesswork. And the record is still incomplete. Similarly, over many years, I have received innumerable inquiries from colleagues looking for original sources (primarily field notes) of information documenting botanists' field activities. Sometimes they knew we had it, sometimes they thought we might have it, and sometimes they were simply casting a wide net. In hindsight, the idea that we should address field books in the same way that we catalog books or inventory natural history specimens seems painfully obvious. And yet, except for materials that were deposited in archival repositories, most of these items were "hidden" because of the lack of a local archives; libraries' inclinations not to claim them; or a well-intentioned, but ill-advised, tradition of biologists "caring" for their predecessor's scholarly work in support of biodiversity research.

My co-principal investigator Anne Van Camp, director of the Smithsonian Institution Archives (SIA), had also identified the need for better descriptions of field books. The SIA contains thousands of field books but, as is the standard in archives, descriptions of individual items within field book collections were minimal. Recognizing our mutual goal of improving access to field books, we envisioned a Field Book Registry that would serve as a global source for locating and describing information within field notes and other field research materials.

Biodiversity research begins with field books, primary source documents that record information about what was observed, discovered, and collected in nature. Because of their integral relationship to specimen collections, field books are often consulted by researchers for a variety of scholarly inquiries. Yet it is their close relationship to the specimen collections, as well as their nature as primary source documents, that makes their categorization as objects ambiguous. Biologists frequently consult field books when they are reconstructing field activities that resulted in collected specimens. Are field books museum objects that

should physically reside close to specimen collections and be described and made available in a similar way? Are they archival materials that should be stored and preserved in archives and described in finding aids? Or should these book-like resources be considered library objects to be cataloged as individual volumes in an online catalog? Perhaps all of these aspects are valid.

The Field Book Project began as a joint initiative of the National Museum of Natural History (NMNH), and the SIA. Together, they applied for and received funding from CLIR to uncover the hidden collections of field books at the Smithsonian. The early years of the Field Book Project consisted of two phases. During the initial phase, the Field Book Project focused on locating field notes throughout the Smithsonian's many research facilities and creating catalog records for these field books within a local database. Phase two saw these records being migrated to a robust, XML-driven pilot registry that will be opened to other institutions so that they can add catalog records for field books in their repositories.

Field books at the Smithsonian Institution are maintained by multiple departments or units: the NMNH science departments, the Smithsonian Institution Libraries, the National Zoo, the Smithsonian Tropical Research Institute (STRI), and the SIA. Some departments, like the Department of Botany, had previously created inventory lists with basic information to help staff members and patrons find the field books they needed. In other NMNH departments, the SIA had created basic finding aids in the 1970s and 1980s, and had done some light processing, but the collections remained in the custody of their respective departments. Earlier SIA finding aids had been very helpful as a foundation for description, but are now outdated in many cases.

Field books are held in the stewardship of museums, archives, and libraries, and therefore benefit from a flexible, yet consistent, standards that combine descriptive practices from all three fields. Our project has drawn from metadata and encoding schemas and content standards across the disciplines to create a hybrid descriptive framework that bridges collection- and item-level description.

Collection-level metadata is a hallmark of archival description. Archival finding aids describe materials as collections, rarely becoming more granular than brief folder-level descriptions. In archival description, providing the context in which materials were created is as important as describing the materials themselves. Unlike published works, which are self-justifying and stand-alone objects, archival documents are like pieces of a puzzle; although they are useful on their own, they need other documents within the collection to provide context and tell a full story.

Libraries have long been innovators for item-level metadata, which have traditionally emphasized access points such as authors or creators, subjects (e.g., topics, locations, names), and titles. The access points of geographic and topical coverage are incredibly important for addressing the bulk of known research needs for primary source field notes. Such access points allow researchers to more easily pinpoint desired volumes. Prolific scientists may create more than 100 field books over the course of their career, spanning multiple collecting events across the globe. Distinguishing one volume from the next based on content, therefore, becomes important for meeting information needs. And because libraries are precedent setters for controlled vocabulary within access points, we follow library descriptive practices for cataloging at the item level.

We define a "collection" as any group of field books with a unifying relationship. Field book collections can be assembled in many ways; our collections, however, are usually grouped by collector or expedition. For example, a collection grouped by the collector Alexander Wetmore would consist of field books created or owned by Wetmore. Alternatively, a collection grouped by the "United States Exploring Expedition" might consist of field books created by various individuals who participated in that expedition. Less frequently, our collections are assembled by an organization as a creator. However they are grouped, the way the field books were physically organized, with respect to the provenance and order in which they were received and maintained prior to our involvement, determines the collections, in accordance with archival practice.

Digitization: Why should we digitize? How can digital versions of these materials enhance access to the material? What are the issues involved in digitization (including crowd sourcing facilitated by digital versions of the resources)?

The initial CLIR grant provided generous funding for our cataloging efforts in the Field Book Project. Early in our workflow designs, we recognized that many field books needed remedial conservation to survive the necessary handling required of this effort. Conservation tasks needed to be a concurrent feature in the workflow with discovery and cataloging. We leveraged the CLIR grant to obtain funding from the National Park Service's Save America's Treasures program to support our conservation efforts. This allowed us to stabilize objects for cataloging and for future digital imaging. In addition, the Smithsonian Women's Committee awarded funds to continue conservation and begin the task of digitally imaging field books at the page level.

What does it mean to digitize something? Technically, it is simply converting data from analog to digital form. Digitization can include producing a digital image of a field book page. It also describes the task of converting (transcribing) text from specimen labels, correspondence, or pages in a field book. However, the term *digitization* has become synonymous with digital imaging, and the term *transcription* is now used to describe the digitization of text. For the purposes of this account, I will refer to digital imaging and transcription.

Digital imaging does not improve access to content. Digital indices, keywords, and web linkages provide that service. That is why it was important to design a catalog record for field books that provided enough metadata to facilitate the process of searching for field book content. The catalog records that we created employ multiple existing standards for managing objects; collections of objects; and inherent connections to people, institutions, and defined collecting events. Once found, a digital image can provide a wealth of information, either explicit or interpreted. In the case of field book pages, it also provides an unambiguous account of a field collector's activities and experiences. It is unambiguous because it is original. There has been no modification, transcription, or transfer of information.

Digital imaging at the page level was a critical piece of our vision for the Field Book Project. Therefore, the design of the Field Book Registry was originally seen as supporting page-level navigation of digitally imaged field books. More importantly, page-level imaging was understood to be the necessary bridge to fully word-searchable original source materials. Only when every word in a field book is essentially an index term can the incredible, yet latent, power of field books be

fully realized. For example, we could deliver every mention of a native use of *Artocarpus altilis* from Samoa prior to 1950.

The task of transcribing every word on every page in every field book, however, will be the most complicated and time-consuming step toward this goal. Currently, the Field Book Project is using the Smithsonian Institution Transcription Center and its army of volunteers to perform transcription services. Our project competes with other Smithsonian Institution projects for the attention of transcribers, and progress is slow. Other transcription services must be considered if the field books are going to become a critical and productive resource for biodiversity research.

Why is integrated access to data and images across scientific disciplines, institutions, and type of resource (e.g., via aggregators like iDigBio and the Biodiversity Heritage Library) important? How can this be accomplished?

Although the need to aggregate related content across institutions, people, and events is what inspired the Field Book Project, the Holy Grail is being able to recognize and deliver inherent relationships between field books, specimens, and published citations, or what is now routinely called "connecting content." Aggregators of institutional content, disciplinary content, and object types serve an important role in our need to see and interpret data as never before. But although these efforts have been highly productive, we have only scratched the surface in our plan to marshal technology and informatics toward seamless access to related content.

Recently, the Field Book Project has joined forces with the Biodiversity Heritage Library in a collaboration that will combine BHL content with field books from more than a dozen major natural history partners.

Digital management and preservation of this work, the digital data and images: What is the current status of digital data management and preservation in your institution?

The Field Book Project is an unusual example of cross-bureau collaboration within the Smithsonian (i.e., the National Museum of Natural History and the SIA). From the beginning, the Office of the Chief Information Officer was engaged to help oversee the technical development and to manage resulting content.

References

Conniff, Richard. In press. *House of Lost Worlds: Dinosaurs, Dynasties, and the Story of Life on Earth.* Yale University Press.

Gaines, Richard V., H. Catherine W. Skinner, Eugene E. Foord, Brian Mason, and Abraham Rosenzweig. 1997. *Dana's New Mineralogy,* 8th ed. New York: John Wiley & Sons, Inc.

Greene, Mark A. and Dennis Meissner. 2005. More Product, Less Process: Revamping Traditional Archival Processing. *The American Archives* 68 (Fall/Winter): 208-263.

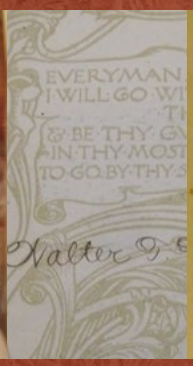
Grinnell, J. 1910. The Methods and Uses of a Research Museum. *Popular Science Monthly* 77 (January 1): 163–169.

Lee, Iris. 2015. Making Decisions…Archives Space or AtoM? Blog post, July 27, 2015, in "Hidden Collections: Stories From the Archive." Available at http://images.library.amnh.org/hiddencollections/2015/07/making-decisions/#more-4847

MVZ Archives Blog: https://mvzarchives.wordpress.com/

SNAC (Social Networks and Archival Context) website: http://socialarchive.iath.virginia.edu/

Outreach


EVERYMAN
I WILL GO WI
TH
& BE THY GV
IN THY MOST
TO GO BY THY S
Walter G.

Opening up the Urban Archive: Digital Outreach to Urban Studies Scholars

Morgan Gieringer, Head of Special Collections, University of North Texas

Jaime Janda, Collections Processing Manager, University of North Texas

Abstract

"Post-War Industry and Development of the Southwest Metroplex" is a project seeking to connect users with previously hidden archival collections at the University of North Texas. Electronic finding aids are the primary access point to these collections; however, this project also seeks to connect individuals with archival resources through digital outreach methods. Digital outreach is an important component of this project because of social scientists' perceived difficulty in accessing and using archival collections, and finding grey literature within archival finding aids. To meet this need, project staff are using a blog and social media as outreach tools to highlight the diverse nature of these collections and their potential for use by historians, social scientists, journalists, and scholars in other urban studies-related disciplines. Social media strategies employed by project staff include use of hashtags and connecting blog posts to current events to increase the likelihood of being shared. Plans for future social media outreach include expanding the scope of the blog to include a greater geographic area, and using video content. Through these efforts, project staff hope to further expose the content of particular collections and attract new users to the archive.

Introduction

Post-War Industry and Development of the Southwest Metroplex is a two-year initiative to process eight archival collections related to the growth of the Dallas/Fort Worth Metroplex following World War II. This project also seeks to connect scholars in a variety of disciplines with archival resources relevant to urban studies through digital outreach methods. Prominent urban studies scholars have been outspoken about the scarcity of archival collections that document the modern history of the Dallas/Fort Worth region. In addition to reaching users currently seeking archival resources, project staff are reaching out to urban studies scholars who may not regularly use archival collections or who might benefit from increased access to grey literature (unpublished papers and reports) found within collections. To this end, project staff are using a blog and social media as outreach tools to highlight the diverse nature of these collections and their potential for use by historians, social scientists, journalists, and scholars in disciplines related to urban studies.

The transformation of the Dallas/Fort Worth area after World War II began with the expansion of the national interstate highway system, which placed Dallas at the intersection of major east-west and north-south highways. In 1969, Dallas and Fort Worth began joint construction of the largest inland port in America—the Dallas/Fort Worth International Airport—further transforming the urban landscape and solidifying the region as a major transportation and industrial

center. Almost five million new residents came to the Dallas/Fort Worth area between 1960 and 2010, making it the largest metropolitan area in the southern United States (Social Science Data Analysis Network 2002). Cities within the Metroplex have raced to accommodate this new population by developing transportation and municipal services to serve the residents and industries fueling the growth.

Demonstrated Need for Access to Urban Studies Resources

The special collections department of the University of North Texas Libraries holds evidence of Dallas/Fort Worth's growth and development in the form of original records, photographs, and collections of personal papers. Eight of these collections were selected to form the 2013 hidden collections proposal to the Council on Library Information Resources (CLIR) for the project Post-War Industry and Development of the Southwest Metroplex. These collections include the papers of Texas Representative Lanny Hall, which document local constituent concerns during years of massive growth in Fort Worth and Arlington as well as political redistricting taking place on the state level (1979-1984). Also included are the archive of *Texas Metro* magazine and the Southwest Federal Regional Council Records (1964-1984), which document the planning and creation of Dallas/Fort Worth International Airport. Other collections reveal how implementation of the federal Model Cities Program (1966-1974) and the Community Development Block Grant programs of the 1970s affected north Texas cities. Topics included in these collections are especially relevant to scholars studying modern urban issues such as affordable housing, urban design, transportation, crime, education, and employment.

Urban studies scholars have been fervent in their support for making these collections more readily available, citing both the relative lack of existing scholarship on the Dallas/Fort Worth area for a metropolitan area of its size and the failure of other local repositories to adequately collect resources in this area. Harvey J. Graff, professor emeritus at The Ohio State University, supported this project, saying, "The development of much needed fundamental research on this important area has been retarded by the lack of serious and sustained library and archival collection development" (Graff 2013). Carl Abbott, professor emeritus of urban studies at Portland State University, noted, "There is great need for these sorts of research materials." He compared important recent work on Sun Belt cities such as Phoenix and Los Angles by saying, "Indeed, Dallas/Fort Worth seems especially underserved by scholarship when compared to greater Los Angeles" (Abbott 2013). Robert Fairbanks of the University of Texas at Arlington provided further support for this project by commenting, "As an urban historian writing on the twentieth century urban Southwest, especially the Dallas-Fort Worth Region, I am acutely aware of the dearth of sources available for the post World War II period, an era of the region's most spectacular growth" (Fairbanks 2013).

Issues Preventing Use of Urban Archives

Although these comments strongly indicate the desire for access to these collections among those who regularly use archival collections, project staff are also focused on outreach to scholars who may not regularly use archival collections. Urban studies of major metropolitan areas is an area of research that traditionally has relied heavily on quantitative research methods such as data collection and statistical analysis. However, developments in the field of contemporary social theory in the past 20 years have led scholars to embrace more qualitative approaches (Jacobs 1993). Ethnographic techniques

such as participant observation and interviews are examples of qualitative methods that are regularly incorporated into urban studies.

Textual analysis, a method of research that examines the content, meaning, structure, and discourse of texts, is another qualitative method that is highly applicable to urban studies and could make good use of archival resources (Lockyer 2008). Textual analysis in urban studies scholarship has used texts such as municipal regulations, public comments, campaign literature, program reports, and policy papers. These last two categories—program reports and policy papers—hold the greatest potential for scholars in the urban studies archive. Unpublished papers and reports, also known as grey literature, are found within many records collections dealing with urban issues (Figures 1 and 2).

Members of the Grey Literature Network Service have identified 72 types of grey literature, according to the 2004 Grey Literature Survey. These include many of the types of documents a researcher would encounter in the urban archive, including bulletins, case studies, feasibility studies, legal documents, policy statements, project proposals, research reports, and white papers (GreyNet International 2004). Specific examples of grey literature found within this project include items such as a brochure for the Southwest Federal Regional Council, proposals and feasibility studies for a new drinking water reservoir, and Dallas/Fort Worth's proposal to bring a major scientific center to North Texas, "A Look into Tomorrow: The Plan to Bring the Superconducting Super Collider to Texas." One particularly noteworthy example is a 1968 report summarizing the initial plans for construction of the airport. This report, titled *Dallas/Fort Worth Airport Planning and Development*, shown in Figure 3, describes the location, size, terminal

Fig 1: Trinity Route of the Dallas-Fort Worth Turnpike: Turnpike Exploratory Investigation, *1971*

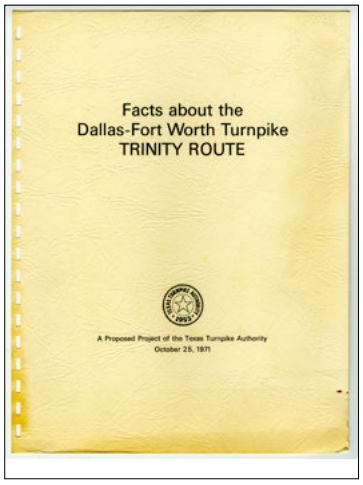


Fig. 2: Facts about the Dallas-Fort Worth Turnpike Trinity Route, *1971*

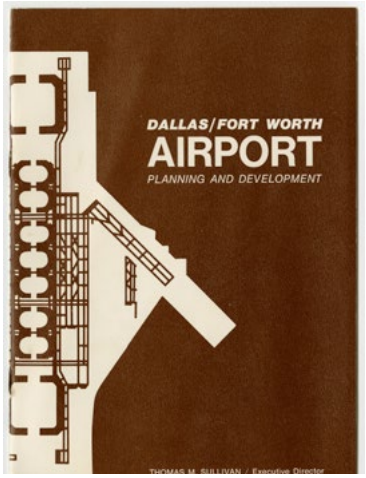


Fig. 3: Dallas/Fort Worth Airport Planning and Development *is an example of grey literature found in the collection.*

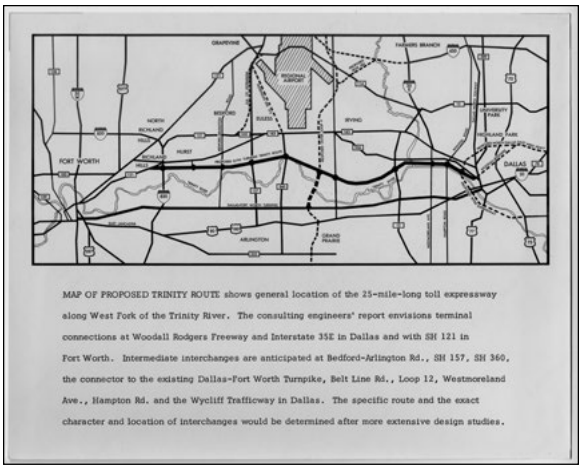
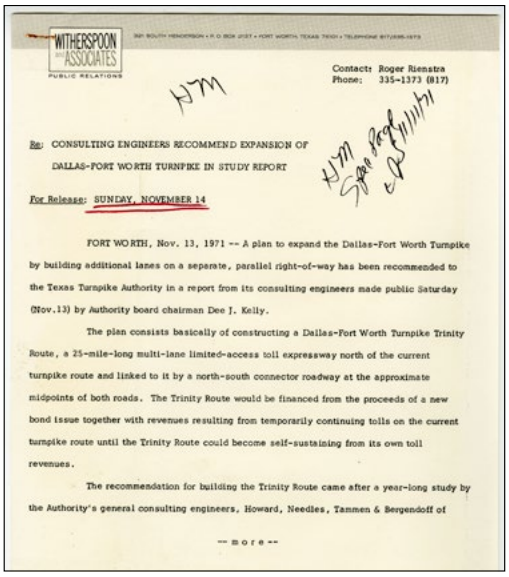


Fig. 4: Map of Proposed Trinity Route. Lester Strother Collection (AR0327), University of North Texas Special Collections.

facilities, ground transportation plans, airspace requirements, and financing for what would become the world's largest airport. (See Figures 4, 5, and 6 for more examples of grey literature).

Library literature supports the potential for grey literature's contributions to scholarship as well as the need for archivists to use outreach methods to connect users with these valuable hidden resources. A 2013 article in the journal *Library Management* cited growing use of grey literature because of increasing opportunities for electronic publishing and digitization. "Many special collections units have increasingly made their deposited non-circulating resources available as digital assets and those collections, once


WITHERSPOON and ASSOCIATES
PUBLIC RELATIONS

Contact: Roger Rienstra
Phone: 335-1373 (817)

Re: CONSULTING ENGINEERS RECOMMEND EXPANSION OF DALLAS-FORT WORTH TURNPIKE IN STUDY REPORT

For Release: SUNDAY, NOVEMBER 14

FORT WORTH, Nov. 13, 1971 -- A plan to expand the Dallas-Fort Worth Turnpike by building additional lanes on a separate, parallel right-of-way has been recommended to the Texas Turnpike Authority in a report from its consulting engineers made public Saturday (Nov.13) by Authority board chairman Dee J. Kelly.

The plan consists basically of constructing a Dallas-Fort Worth Turnpike Trinity Route, a 25-mile-long multi-lane limited-access toll expressway north of the current turnpike route and linked to it by a north-south connector roadway at the approximate midpoints of both roads. The Trinity Route would be financed from the proceeds of a new bond issue together with revenues resulting from temporarily continuing tolls on the current turnpike route until the Trinity Route could become self-sustaining from its own toll revenues.

The recommendation for building the Trinity Route came after a year-long study by the Authority's general consulting engineers, Howard, Needles, Tammen & Bergendoff of

-- more --

Fig. 5: Witherspoon and Associates 1971. Consulting Engineers Recommend Expansion of Dallas-Fort Worth Turnpike in Study Report, November 13 1971. Lester Strother Collection (AR0327), University of North Texas Special Collections.

considered visibly grey, become less grey. The encoded archivists' description tool kits provide resources for librarians and archivists to create user guides to such collections" (Gelfand and Lin 2013, 541). In this instance, the authors suggest that access points, in the form of electronic finding aids, perform as the outreach strategy to increase use of grey literature. Often, however, the size of an archival collection precludes item-level description, as is the case with the collections that make up this project. Project staff have decided

Fig. 6: Rural Trinity Route. Lester Strother Collection (AR0327), University of North Texas Special Collections.

instead to use selective digitization and description, combined with a social media strategy, to provide increased access to grey literature and to visual and other types of under-described textual materials.

Digital Outreach Methods

Making collections available through electronic finding aids and digitization is a major step toward outreach to researchers in the Dallas/Fort Worth area, as well as to other institutions across the country, yet it is still only one step. To make processing, digitization, and finding aids more widely accessible, project archivists chose to conduct outreach through social media and other available services. A blog called "Discovering the Southwest Metroplex," hosted on the University Libraries website, together with the special collections Twitter account, provide the main channels for collection outreach. The special collections Facebook page, university newsletters and listservs, and press releases to the local community constitute secondary methods of outreach.

The project blog is used to promote specific entries through social media outlets, primarily Twitter and Facebook. Between May 2014 and January 2015, project archivists created 14 original blog posts. Of these, about 65 percent have been shared through other media outlets by library staff, special collections staff, and the public. Of the posts shared through social media, 26 percent were tweeted, and 39 percent were shared on Facebook pages (Figures 7 and 8).

In examining the figures related to outreach through social media, project staff thought that the statistics fell short of original expectations. In the second year of the project, staff must ask themselves what more can be done to enhance social media outreach.

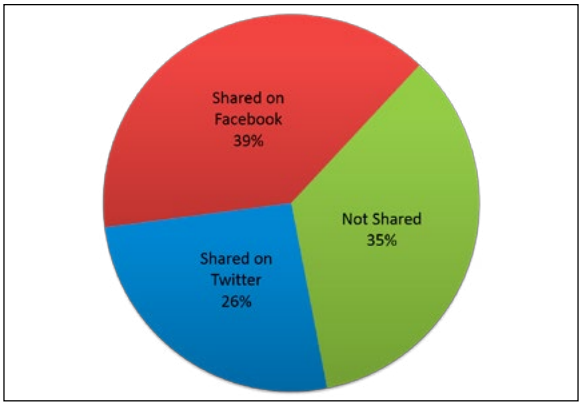


Fig. 7: Statistics of blog sharing

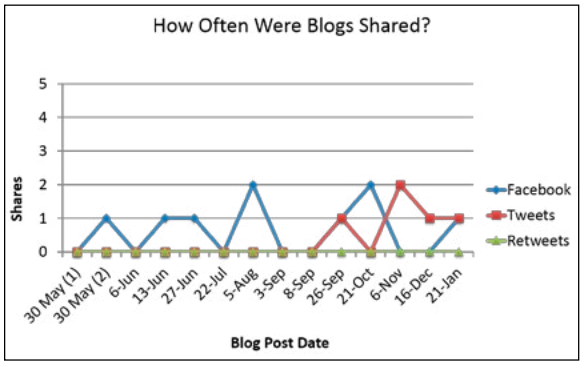


Fig. 8: How often were blogs shared?

Early on in the project, a major problem staff faced was lack of knowledge about collection contents. Collections that qualified for inclusion under this grant were hidden collections and thus lacked any descriptions besides the most rudimentary accession records. Staff selected boxes at random and sifted through them to find general information or specific items to write about in a blog. To overcome this obstacle, staff supplemented blog posts with information from related collections that were not included in the grant, but were connected to the grant collections through larger themes of transportation, housing, population, and tourism. Some examples of blog post topics include the Texas State Fair, the El Chico restaurant chain, the Fort Worth Stockyards, and the construction of the Dallas/Fort Worth

International Airport. Once processing started on the grant collections, staff wrote more specific posts highlighting potential research areas in the hidden collections.

Project staff looked at several ways to improve the reach of blog posts through social media. The first method is to tweet specifically to potential interested parties through established Twitter conventions (Figure 9). Staff used this method in a post on the Texas Rangers baseball team.

When special collections staff tweeted a link to this blog post, they made specific mention to "@ Rangers." This method can also use specific hashtags. Because there is no automated way for the designated tweeter in special collections to be notified of a new blog entry, this method works best when the blog author notifies the Twitter account holder, and makes specific suggestions about people to mention or hashtags to incorporate.

A second method is to connect blog posts with current events. For example, the Rangers Stadium blog post is scheduled to be retweeted closer to the opening of baseball season. Staff also plan to encourage wider outreach to baseball fans by including AT&T Stadium, the Houston Astros, Major League Baseball, and even Six Flags over Texas in the tweet because they also were mentioned in the blog post. This method will cast a wide net that will attract a larger audience and new researchers.

Because the primary audience for this project is urban studies and geography scholars, social scientists, historians, and journalists interested in the development of the Southwest Metroplex, project staff have identified the need to package collection information in a way that is useful and relevant to these audiences. By enlarging the scope of the blog to include closely related cities, project staff anticipate an even larger audience. Future posts will incorporate information not only about

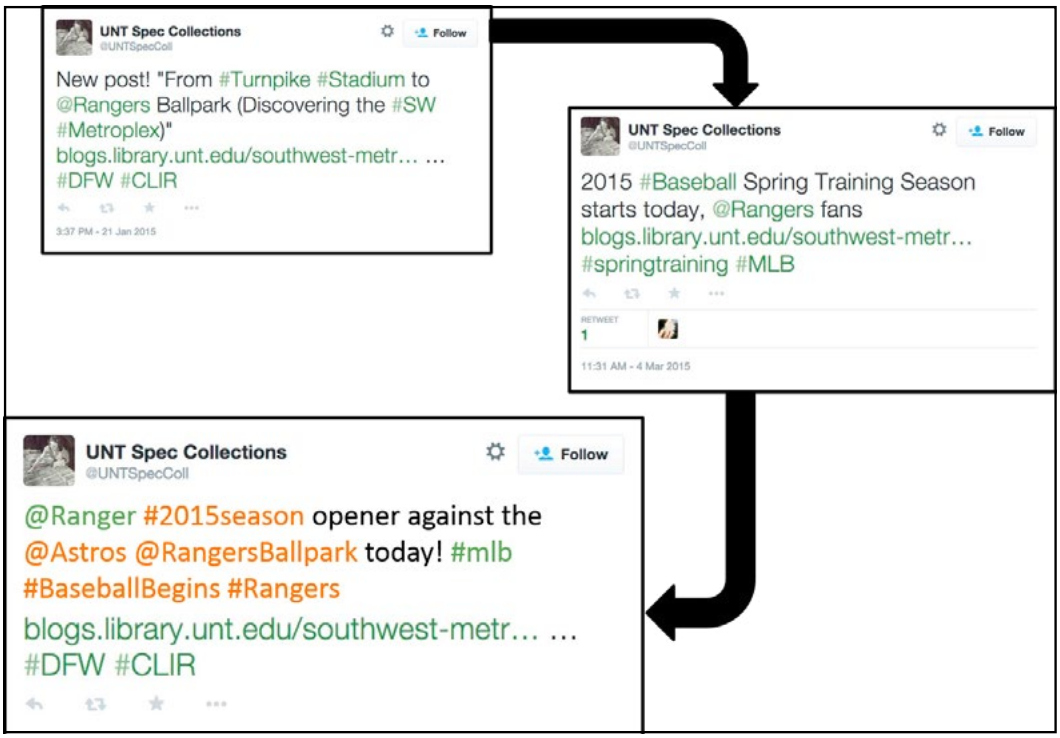


Fig. 9: How blogs should be tweeted

the growth of Dallas/Fort Worth, but also (if available) the growth of other cities in Texas and surrounding states. For example, the records of *Texas Metro* magazine contain photographs, brochures, and articles on many Texas cities outside the Dallas/Fort Worth area that experienced growing tourism in the 1960s and 1970s. Lester Strother, a journalist and later owner and publisher of *Texas Metro*, also traveled to other states and countries, writing articles on hotels and vacationing spots. These topics are not the main concern of this project, but they can be highlighted to appeal to urban studies scholars, historians, journalists, or whoever else might be interested.

Forging connections with scholars interested in the Dallas/Fort Worth area as well as other Texas metro areas such as Austin, Houston, and San Antonio opens up the possibility of partnering with university and city archives in places that share similar collections. For example, in the Lanny Hall Collection are materials related to many universities and towns that project staff could connect with outside of the Dallas/Fort Worth Metroplex. Lanny Hall not only was a Texas state representative during the 66th, 67th, and 68th legislative sessions, but he also studied at the University of Texas and at Hardin-Simmons University, where he later became president. In addition, he has professional ties with Howard Payne University, Baylor University, and Wayland Baptist University, among others. Even though the collection deals mainly with his time as a state representative in the Fort Worth area, researchers may be interested in all aspects of his career, including his speeches and other publications, his involvement in major projects around Dallas/Fort Worth, and legislation he authored or helped pass. The benefit of connecting with scholars interested in what these collections have to offer is their communication with fellow scholars in their disciplines.

Another area of outreach staff identified was partnerships with local libraries, museums, and universities to host events about the growth of the Dallas/Fort Worth area. Events can be used to promote the eight grant-funded collections and highlight others that may be of interest to the local community. Videos can be taken during the event and used to create a specific YouTube channel for special collections and then easily shared through established social media networks. Other ideas for video content include short videos of staff processing parts of the collections. In these videos, staff plan to highlight maps and blueprints that show the expansion of road systems, neighborhoods, lakes, and the Superconducting Super Collider. These videos will encourage use and raise awareness of the department's collections and services. They can benefit the department by showing the effort put forth to make collections available to the public, which may also encourage collections donations from local businesses or personal papers from well-known community members.

For collection outreach, staff are now evaluating many types of social media outlets. Facebook, Twitter, and the project blog are the top choices; others are just as good but are currently overlooked, such as Wikipedia, YouTube, Instagram, Reddit, Google+, LinkedIn, and even the long-lost MySpace, which is becoming popular again with artists, writers, musicians, and others in similar industries. These outlets will let us reach out locally and eventually, we hope, worldwide, while staff engage locally in more personal and one-on-one conversations.

Conclusion

The success of the Post-War Industry and Development of the Southwest Metroplex project hinges on the ability to connect users with

previously hidden collections. Electronic finding aids are the primary point of access to these collections; however, this project also seeks to connect scholars in a variety of disciplines with archival resources relevant to urban studies through digital outreach methods. Digital outreach is an important component of this project because of social scientists' perceived difficulty in accessing and using archival collections, and finding grey literature within archival finding aids. To meet this need, project staff are using a blog and social media as outreach tools to highlight the diverse nature of these collections and their potential for use by historians, social scientists, journalists, and scholars in other urban studies-related disciplines. Social media strategies employed by project staff include use of hashtags and connecting blog posts to current events to increase the likelihood of being shared. Plans for future social media outreach include expanding the scope of the blog to include a greater geographic area, and using video content. Through these efforts, project staff hope to further expose the content of particular collections and attract new users to the archive.

References

Abbott, C. 2013. Letter of Support for University of North Texas Proposal to the Council on Library and Information Resources.

Fairbanks, R. 2013. Letter of Support for University of North Texas Proposal to the Council on Library and Information Resources.

Gelfand, J., and A. Lin. 2013. Grey Literature: Format Agnostic Yet Gaining Recognition in Library Collections. *Library Management* 35(6): 538-550. http://www.emeraldinsight.com/doi/abs/10.1108/LM-03-2013-0022.

Graff, H. J. 2013. Letter of Support for University of North Texas Proposal to the Council on Library and Information Resources.

GreyNet International. 2004. *Document Types in Grey Literature*. Available at http://www.greynet.org/greysourceindex/documenttypes.html.

Howard, Needles, Tammen and Bergendoff Consulting Engineers. 1971. "Trinity Route of the Dallas-Fort Worth Turnpike: Turnpike Exploratory Investigation." November. Lester Strother Collection (AR0327), University of North Texas Special Collections.

Jacobs, J. M. 1993. The City Unbound: Qualitative Approaches to the City. *Urban Studies* 30: 827-848.

Jacobs, K. 2001. Historical Perspectives and Methodologies: Their Relevance for Housing Studies? *Housing, Theory and Society* 18 (3/4): 127-135. http://www.tandfonline.com/doi/abs/10.1080/14036090152770492.

Lockyer, S. 2008. *Textual Analysis*. In L. M. Given, ed., *The SAGE Encyclopedia of Qualitative Research Methods*. Available at http://dx.doi.org/10.4135/9781412963909.n449.

Map of Proposed Trinity Route. Lester Strother Collection (AR0327), undated. University of North Texas Special Collections.

Rural Trinity Route. Lester Strother Collection (AR0327), undated. University of North Texas Special Collections.

Social Science Data Analysis Network. 2002. *Dallas, TX Population Growth 1960-2000*. Available at http://www.censusscope.org/us/m1920/chart_popl.html.

Sullivan, Thomas M. "Dallas/Fort Worth Airport Planning and Development." Lester Strother Collection (AR0327), undated. University of North Texas Special Collections.

Texas Turnpike Authority. 1971. "Facts about the Dallas-Fort Worth Turnpike TRINITY ROUTE," October 25. Lester Strother Collection (AR0327), University of North Texas Special Collections.

Witherspoon and Associates. 1971. “Consulting Engineers Recommend Expansion of Dallas Fort Worth Turnpike in Study Report,” November 13. Lester Strother Collection (AR0327), University of North Texas Special Collections.

Success Beyond Access: CLIR-ing the Way

Harlan Greene, Head of Special Collections, College of Charleston

Dale Rosengarten, Curator of the Jewish Heritage Collection, College of Charleston

Amy Lazarus, Processing Archivist, College of Charleston

Abstract

The College of Charleston Libraries have used three successive Council on Library and Information Resources (CLIR) grants to process about 1,000 linear feet of collections documenting African-American and Jewish history and culture. Project staff built on relationships not only with these communities, but also with students, researchers, and local, national, and international cultural heritage institutions, to provide constituents a level of access going far beyond traditional finding aids. Beyond processing collections, we made them available to our stakeholders, researchers, and even non-traditional users such as tourists, in ways previously unimaginable. Researchers can now access our collections in our reading room, or by searching the Lowcountry Digital Library for manuscript and iconographic materials and oral histories, or by viewing virtual exhibits focused on our collections. Project staff regularly give class presentations and supervise student internships and honors theses. In collaboration with like-minded institutions, we produce exhibitions, publications, and online educational resources. We welcome tour groups from across the country that come to see materials processed with funding from CLIR. Our home institution recruits students and attracts donors by showcasing these collections. Remarkably, even people who never set foot in the reading room profit from the collections' accessibility.

Below we offer three perspectives. Harlan Greene discusses the projects at the Avery Research Center for African American History and Culture and Jewish Heritage Collection, and examines how accessible archives affect a wide range of audiences. Dale Rosengarten describes the persistent efforts that have contributed to the Jewish Heritage Collection's remarkable success. Amy Lazarus provides examples of incorporating technology, academic engagement, and outreach into her work on our latest CLIR grant.

▸ Sowing the Seeds

Harlan Greene, Head of Special Collections, College of Charleston

Since 2008, the College of Charleston has been fortunate to receive three Hidden Collections CLIR grants. There is some irony in this, for Charleston is known, and markets itself, as "America's Most Historic City." Tourists come from around the globe to see, tour, touch, and now even taste the revealed, preserved, and accessible past. The city is regarded as a world attraction, and a significant part of its economy derives from the allure of its historic assets. How could we have hidden history?

Easy. The fact is, the collections that received CLIR funding document groups written out of history or underrepresented because their stories are painful or difficult to parse in a tourist brochure. Scholars, entrepreneurs, and the chamber of commerce are now recognizing the value of including previously excluded stories in interpretations of the past. Belatedly, the city is embracing and advertising both African American

Fig. 1: Harlan Greene leading a walking tour in Charleston, September 2010

and Jewish histories—the subjects of our newly "unhidden" collections.

While our archivists did not drive this change, it has been reinforced by their good work. "Success beyond access" describes not only our innovative processing strategies but also the impact three CLIR grants have had outside the archives and research room.

The Avery Research Center for African American History and Culture (Avery) received the first CLIR grant in 2008. With collections of materials documenting enslaved Africans, free people of color, civil rights leaders, African American social groups and institutions, and the work of a trailblazing English anthropologist, Avery was well known for its historical building, erected in 1868 as a school for recently freed slaves, and for its cultural programs. Its archives were touted but not used to their fullest extent, as no archives can be until their contents are arranged and described. With its collections now largely cataloged and accessible, use by students, scholars, genealogists, and filmmakers has escalated—with ripple effects across the campus and beyond.

Today, Avery's collection is used as a recruiting tool for a burgeoning African American studies program. As the program expands and hires more staff, candidates are brought to Avery to see the rich primary materials available to faculty to use as teaching aids and in their own research. The College now offers a major and a minor in African American studies—a fact not unrelated to Avery's accessible archives.

But it is not only faculty from that department who use Avery. Jon Hale, an assistant professor in the school of education, reports that "The Foundations of Education course is now a general education (humanities) course because we utilize the collections at Avery for a primary course assignment. This assignment is used for SACS [Southern Association of College and Schools] accreditation." Dr. Hale also consults for Charleston County Public School System—a system, like the city itself, that still suffers from the sins of the past. "While many of the local teachers know Avery," Dr. Hale notes, "I make it a point in my courses that students visit Avery and utilize their collections to create Document Based Questions, which are a requirement of the new Common Core standards. Therefore, as teachers readjust their teaching to new standards, a number are incorporating Avery sources into their pedagogy …. Avery," Hale summarizes, "has helped grow the work of the education department."

Black history is being incorporated—or perhaps the more appropriate word is integrated—into general education courses and grammar school classrooms. Charleston, a historically black city for much of its existence, never included much black history in its self-representation except in a stereotypical, servile manner. But now, partially with the impetus of the historical assets at Avery, that is changing. The city's plans to build an International African American Museum, years in the making, depend in part on access to Avery's well-cataloged collection of artifacts, photographs, and manuscripts.

While Charleston has long been recognized as "the Ellis Island of black America," less well-known is its history as the cultural capital of American Jewry. In 1800, Charleston had the largest Jewish population of any city in North America and remained the Jewish metropolitan center of the new republic until the 1820s. How many people know that South Carolina claims the first professing Jew elected to office in the Western world, the first Jewish patriot to die in the American Revolution, the first Hebrew Benevolent and Orphan Societies, and the first dissidents to attempt to reform Judaism in the United States?

Fig. 2: Conference-goers gather outside Special Collections in Addlestone Library at the College of Charleston to view the exhibits in our display cases. Biennial Scholars' Conference on American Jewish History, held in Charleston, South Carolina, June 2006. PHOTO BY DANA SARDET.

Bountiful yet unprocessed materials on southern Jewish history, the Holocaust in Europe, and a vast collection of iconographic Judaica from around the world earned us two other CLIR grants. Here, too, the broad and lasting effects of revealing hidden collections can be demonstrated. Our experience processing 357 linear feet of the Jewish Heritage Collection, housed in Addlestone Library's Special Collections, about six blocks from the Avery Center, is another case of success beyond access.

Since its inception in 1995, the Jewish Heritage Collection has enjoyed a strong symbiotic relationship with the College's Jewish Studies Program, established in 1984 and expanded from a minor to a major in 2013. Says Student Recruitment Counselor Helen Slucki, "Your collections do help us in recruiting students . . . . When I meet prospective students who are interested in majoring in history or Jewish studies, I definitely go into more detail about . . . what resources we have available here, including the archives." These materials, Ms. Slucki notes, give families of prospective students "a level of comfort" and are a "push factor" in recruiting. For those interested in the Jewish Studies Program, a dedicated Jewish Heritage Collection strengthens the academic resources of the department. Slucki further notes that even for students who never actually use the archives, its presence influences their decision to attend the college. Prospective Jewish Studies donors are also brought to the library to show off the collection and encourage their support. Again, accessible archives have a tangible impact on the growth and development of academic programs and on the parent institution itself.

Beyond the academy, diverse communities have taken note of these collections. When Brith Sholom Beth Israel, an Orthodox Jewish congregation in

Charleston, brought together Jews from across the country for a Shavuothon (a communal observance of the biblical holiday of Shavuot), organizers provided time in the program for these very observant individuals to walk the distance from the synagogue to Addlestone Library, bypass the elevators, and take the stairs to the third floor to see a display of new treasures from the Jewish Heritage Collection. Staff was on hand to guide them through the exhibit and engage in discussions about southern Jewish life, past and present.

As noted earlier, Charleston is a tourist city, and the now unhidden collections have added an asset to the local economy. Jewish museums, congregations, and federations from across the country, seeking educational opportunities, now regularly inquire about the collection. Jewish Heritage Collection Curator Dale Rosengarten provides slide lectures on southern Jewish history and showcases materials from our internationally important Rosenthall Collection. For 20 years, Ms. Rosengarten has worked diligently to secure the funds to process collections and pursue new avenues for public education and outreach.

Fig. 3: A roundtable on the Eldridge Street synagogue restoration project in lower Manhattan was held in the sanctuary at Kahal Kadosh Beth Elohim during the Biennial Scholars' Conference on American Jewish History, held in Charleston, South Carolina, June 2006. PHOTO BY DANA SARDET.

Tilling the Soil

Dale Rosengarten, Curator, Jewish Heritage Collection, College of Charleston

We built it and they came! With persistent effort, we have created opportunities to attract an audience, engage students and scholars, publicize our collections, and reach the general public. As a faculty member, I regularly teach courses with an archival research component. As members of the Special Collections staff, Harlan Greene and I both present classes and workshops on demand, highlighting our collections whenever possible. We produce educational materials and websites, and collaborate with reference librarians to generate research guides for specific syllabi with links to material in the Jewish Heritage Collection, including guides for Southern Jewish History and Holocaust Studies. When students come to Special Collections looking for material for research papers, they receive extended consultations—sometimes with excellent results. A recent honors student advisee of mine, for example, wrote a paper on a Jewish-owned rice plantation on the Ashley River based on the planter's journals, and successfully submitted it for publication in a respected academic journal. Our position as faculty members, archivists, and scholars gives us access to colleagues in allied fields. In 2003 we invited Jeffrey Gurock, a professor at Yeshiva University and an expert on the history of Orthodox Judaism in the United States, to explore our collection of Brith Sholom Beth Israel records and write a history of the congregation, which we published in 2004, in time for BSBI's one-hundred-and-fiftieth anniversary. Last year, when eminent Jewish historian Gary Zola was in residence as the College of Charleston's Arnold Distinguished Visiting Chair, we provided his class with a resource guide and supplied him with archival material for his current research

project. Zola, executive director of the Jacob Rader Marcus Center of the American Jewish Archives and professor of the American Jewish Experience at Hebrew Union College-Jewish Institute of Religion in Cincinnati, agreed to collaborate with us on a website that will provide digital access to manuscript prayer books written by founders of the Reformed Society of Israelites that now reside in his collection and ours.

When a scholar finds a useful source in the Addlestone Library catalog or the Lowcountry Digital Library (LCDL), we frequently suggest additional materials now processed and available, if not digitized. In one mutually profitable exchange, a PhD student from the University of Florida requested imagery of the Jewish ritual of *kapparot*. He saw a postcard from the Rosenthall Collection posted on the LCDL and hoped it was available for use in an article. We sent him that image—and another 15 scans that were not yet online at the time of his request, but are now digitized and available. Realizing there were holdings from which other scholars would benefit, he returned the favor by offering us a *kapparot* image in his possession.

As we raise the profile of the Rosenthall Judaica Collection—the focus of our latest CLIR grant—we are attracting interest from serious collectors, most recently from an expert in New York who specializes in commemorative medals. He was excited to see our newly processed materials and offered to curate an online exhibit in the future. Collectors not only can offer detailed subject knowledge; they can become potential funders, donors, or both.

Efforts over the past 15 years to collect Holocaust-related materials helped persuade a major donor to establish an endowed chair at our Jewish Studies program. Another donor—himself a Holocaust survivor—gave substantial gifts to both Jewish Studies and Special Collections, and subsidized the creation of the Holocaust Quilt Memorial website, based on archival material we proactively collected. Most recently, we provided a rationale for endowing the new Pearlstine/Lipov Center for the Southern Jewish Culture, an initiative in which we are a full partner. Our plan is to develop a consortium of southern Jewish archives and academic programs, build an online portal that will help researchers locate relevant archival materials across the South, offer research fellowships and archival internships, encourage scholarship in regional studies, and promote Jewish heritage tourism.

Fig. 4: A highly successful panel brought together the "Kings of King Street"—ten King Street merchants, past and present. The event, which attracted a standing-room audience, was held in Arnold Hall in the Jewish Studies Center, January 24, 2013. PHOTO BY DANA SARDET.

For the Jewish Historical Society of South Carolina (JHSSC), we produce a biannual newsletter, loaded with no-longer-hidden archival images and documents. We help the society plan its fall and spring meetings, sometimes constructing an archival exhibit on the theme of the conference. This spring, for instance, we created a display of materials on World War II for a meeting held on the seventieth anniversary of VE-Day. JHSSC returns our support with an annual stipend that pays part of the salary of our oral history archivist, who played a vital role in the Jewish Heritage Collection's first CLIR grant.

Special Collections has become a sought-after stop on the itinerary of Charleston's Jewish-focused tours. I am frequently asked to address a visiting Jewish federation or congregational group from places as far flung as Rochester, New York, Dallas, Greensboro, Memphis, Savannah, Baltimore, Chicago, Cleveland, and Denver. These encounters are seldom a one-way street. Tour groups regularly make financial contributions to Special Collections, and offer leads and sometimes actual artifacts to enhance our holdings.

While our mission is mainly academic, we assist in efforts of cultural conservation. The Legacy Project, for example, which strives to help declining congregations plan for an uncertain future, used our collections and images, and interviewed staff on camera to produce a film focusing on Temple Sinai in Sumter, South Carolina, and its emotionally wrenching decline. The materials we presented, processed during our first CLIR grant, had been gifted to the Jewish Heritage Center by the congregation as part of its living will. Closer to home, we revised the docent manual for Congregation Kahal Kadosh Beth Elohim, and provided the first in a series of training sessions.

Evolving archives, too, look to us as a safe haven or a wise counselor. Last year a Jewish collection in a neighboring state considered relocating its material to our library in an effort to find a more activist and professional steward. To counter the risk of losing the collection, the archives' home institution offered to upgrade its services.

Because the Rosenthall Collection and Holocaust Archives are global in scope, we have become a resource for the European theater as well as for American audiences. Through the LCDL, researchers from New York to California and internationally from Austria, England, and Wales have discovered collection images online and requested permission to use them in a variety of media, including articles, books, videos, educational television specials, and exhibits. An organization in Amsterdam inquired about a Dutch survivor whose story is included on the Holocaust Quilt website. (The staff in the Netherlands had recently acquired a photo archives that contained studio shots of the woman and her mother in 1942.) Here was a dramatic demonstration of the presence of the past, an example of what archives are uniquely able to do.

All of our activities, exchanges, and outreach rest on our ability to process and catalog our collections. Without the painstaking work of CLIR-funded professionals like Amy Lazarus, we would not have achieved nearly as much.

▸ Harvesting the Crop

Amy Lazarus, Processing Archivist, College of Charleston

While processing the William A. Rosenthall Judaica Collection, grant-funded archivists succeeded in establishing new audiences on and off campus, increasing the collection's discoverability, and exploring new ways of using the materials. These successes are certainly part of access, but they were accomplished in non-traditional and innovative ways beyond arrangement and description—we used strategies that required us to reach outside the archives and the research room, both physically and virtually.

Specifically, we recruited student interns to digitize materials; established regular contact with students and faculty, informally spreading word about the collection; maximized use of social media; introduced Rosenthall Judaica to Special Collections patrons; and collaborated with a scholar to create an online educational tool. Some strategies were more fruitful than others.

The Rosenthall CLIR-funded grant called for hiring a project archivist and a processing archivist to process and catalog the collection. Both actively recruited undergraduate interns to digitize items during processing, thanks to an existing vehicle for access—the college's Lowcountry Digital Library (LCDL). By the end of the grant period, almost 5,500 collection items were made available online. While the student workers were unable to produce totally reliable metadata on their own, their recruitment provided an excellent opportunity for outreach. Early in the grant cycle, project staff began to interact with various college departments, contacting staff, posting fliers, and asking specific professors to make these internships known to students. Jewish Studies and Anthropology, in particular, saw the potential value of the program.

The student interns, who earned three credits for a semester's work, were the main beneficiaries. Instead of turning them into drones, we put the tasks in context, introducing interns to the basics of archival science, its uses and importance. After they learned about digitization procedures and metadata, and gained an appreciation for the significance of the collection materials, we provided background information on aspects of Jewish history and culture, topics the students knew little about. In this way, we contributed to the larger mission of our parent institution by broadening the worldview of its students. Each student received one-on-one guidance on the cultural facts relevant to the specific task. As a result, several interns expressed interest in a career path related to archives and museums. Besides encouraging a vocational direction, project staff provided reference letters supporting applications for further archival internships and post-graduate jobs. Recruiting interns helped spread news of the collection through academic circles and affected individuals in a meaningful way.

While the interns themselves had many positive experiences, the products of their work were not uniformly successful. Examining the reasons why will help guide future projects. The internship program began during the Jewish Heritage Collection's first CLIR grant. Though dedicated and enthusiastic, students lacked the specialized subject and language skills needed to understand the multilingual and sometimes esoteric materials. Without knowledge of Jewish ritual, synagogue fixtures, key moments in Jewish history, and Bible stories and figures, eager interns often missed the relevance of certain crucial details. The use of an Internet engine to translate languages also led to misidentified items and other inaccuracies.

The college has no master's program in library science to tap. At our small liberal arts institution, finding undergraduates with the combination of skills and education required was difficult. Despite the obvious appeal of making archival materials available in digital format to a worldwide audience, access at any cost is not the way to proceed. Realizing this, and seeing that it often takes more time to correct metadata than to create it, archivists who were hired under the second CLIR grant took care to look for interns with language proficiency, to establish quality control measures, to supplement the students' gaps in knowledge by providing translations and background on Jewish history and ritual, and to increase supervision and scrupulous review of work. Ultimately this proved too time-intensive to be practical. After two semesters, we decided the Rosenthall Collection was not a good candidate for student internships. To continue uploading new items to LCDL, we created metadata for a backlog of scans from the previous grant. With a more suitable collection, training student interns for digitization might be more appropriate.

Fig. 5: Project Archivist Sarah Glover and Processing Archivist Amy Lazarus review Rosenthall Collection materials with consultant Samuel D. Gruber, November 2014. PHOTO BY DALE ROSENGARTEN.

It is never a mistake, however, to reach out to other academic departments and provide individual instruction to a new generation of archival patrons and potential archivists.

Student intern recruitment was not the only way we used the academic environment to create a network of users. I audited Hebrew language courses and took the opportunity to engage with faculty and students. This allowed me to introduce a key constituency to the Rosenthall Collection and describe examples of the materials it contains. Many students had little sense of what an archivist does, and their curiosity spurred additional promotion and outreach.

Toward the end of the project, we had an unexpected opportunity to create a digital exhibit. With CLIR's permission, we hired Samuel D. Gruber, an established scholar with a specialty in synagogue architecture, to choose images from the collection and write text around a central theme highlighting one of the collection's strengths. A digital exhibit, we believed, would not only serve scholars and publicize the collection within the field of Jewish studies, but would also provide an educational tool for a lay audience. Adopting the title "The Life of the Synagogue," we produced an insightful and engaging online exhibit that will interest scholars around the globe and also serve Jewish day schools, high school students, and undergraduates.

In creating this resource, project archivists also functioned as scholars. Building on formal education in Jewish history and culture and the knowledge gained from the research required to process and describe the collection, we guided our consultant, showing him what images were available, suggesting others, and contributing to the exhibit text. Project coordinators Harlan Greene and Dale Rosengarten helped edit texts and ensure the exhibit was accessible to a range of audiences.

To encourage use of the collection, staff of the previous CLIR grant established A Synagogue A Day Tumblr, which gained dedicated followers week by week, beginning in August 2011 and continuing to the present. Each post includes a short blurb about the collection, where it can be found, and links to LCDL, the Rosenthall website, and the collection's Twitter account. Thus, no matter where viewers find the image, it leads them back to the Rosenthall Collection. Through both grants, project archivists maintained a blog with updates, a highlight reel, and information on parts of the collection that were not digitized. The blog prompted several requests for images, specifically from an entry about the German periodical *Allgemeine illustrirte Judenzeitung.*

While many of our ventures were deliberate undertakings requiring planning and implementation, we found opportunities in unlikely situations, such as translating materials for donors and researchers using collections other than Rosenthall. While providing this language assistance to researchers, project staff served as "ambassadors" for the Rosenthall Collection, introducing Special Collections users to new sources available for research.

Through these various means, it became apparent to everyone involved in the project that thinking creatively about access can lead to novel ways of thinking about the collection itself. Even when attempts did not fully succeed, they provided valuable lessons for the future. Our efforts to engage new audiences and make use of emerging technologies already have raised and will continue to raise the visibility of the "unhidden" collections at the College of Charleston.

References

Holocaust Quilt Memorial: http://holocaustarchives.cofc.edu/

"News and Updates," Rabbi William A. Rosenthall website: http://rosenthall.library.cofc.edu/news-and-updates/.

A Synagogue A Day: http://asynagogueaday.tumblr.com/

Concluding Remarks

William Noel, Director, Kislak Center for Special Collections, Rare Books and Manuscripts; and Founding Director, Schoenberg Institute for Manuscript Studies, University of Pennsylvania

I want to make two claims for us. The less ambitious first: With CLIR's help we have shaped history. By redefining what constitutes an archive, by inventing new ways to process the archives, and by creating myriad paths of access to a stunning amount of previously hidden material from our past, we have in the last eight years expanded the well of material upon which so many diverse audiences can draw. This is an enormous achievement and we should thank CLIR for funding our efforts and for providing a compass to orient us. We can also thank each other for brilliantly mapping the terra incognita that so many others will enjoy exploring in their own ways.

The second claim I want to make is that, together with CLIR, we have made history. This is a different point. I don't mean simply that we have organized the past in ways that make it history out of debris—although that might be true. I mean rather that the Hidden Special Collections program has been a landmark program in library history—that it has fundamentally changed the way in which libraries go about their business. We have made archives history, and I applaud Jackie Goldsby's conceptualization of this as the creation of the relational archive.

If we are creating relational archives, what then has the archivist become? What are the characteristics of the relational archivist? Lets take Jackie literally. The archivist now is someone who creates acts of human contact and sociability, who actively engages the user in making the archive visible and functional, whose dialog with the stakeholder becomes the quintessence of his or her practice, and who conceives of their archive as working interoperably with similar materials in geographically distant repositories. These are the archivists that I see every day in the Kislak Center, working collaboratively in a regional environment and in an international one, orchestrating students not just to process collections but to publish on them, and varying the depth of their cataloging with an eye not only to the vast pile of yet-to-be-processed material in their inbox, but also with an eye to their users' needs. And this is a very far cry from the archivists that I trembled before as a student in England 30 years ago.

So much has been achieved by archivists such as these, and the work is not done. There are vast tracts of analog special collections still to be processed, and it is our job to continue to process them without continued funding from CLIR. We still have CLIR's compass, but we don't have CLIR's money. CLIR, which is enormously proud of the achievements of the Hidden Collections program, is as anxious about this as many of you are. But I think the seeds of our future success are in our current practice. We have created digital gateways to our collections in experimental ways. We have created rich digital assets that can now be built on, further explored, and emulated. Even though the archives that we have processed are for the most part analog, I would like to suggest that in another environment we would be called digital humanists: we have certainly provided

the tools and the data for these humanists to use, and, who knows, maybe these same people will help us as we continue to celebrate and discover our past. And the delightful thing is that the academy is catching up with us, and I predict that comparatively well-funded digital humanities student programs in universities across the land will be looking to the archives, and to archivists, to help them explore in a digital environment the still-hidden analog world.

CLIR itself is changing its emphasis to the digitization of hidden special collections, and with good reason. While it has changed the landscape for the cataloging of hidden collections, current practices for the digitization of those collections need the radical rethinking that CLIR's money and brainpower can inspire.

Its pretty clear to see what CLIR is aiming for. If you can just work with your neighbor, you double your money. Half a million instead of $250,000 for those of us who are farsighted enough to see that it is our combined data that is important, rather than our branded proprietary website, and that digital collections can have an increased autonomy, integrity, and power precisely by transcending the physical limitations of analog collections. Collaboration is hard; digital haves need to work with digital have-nots. It necessitates a generosity of spirit, by both the haves and the have-nots, that more than matches the generosity in cash that CLIR promises to the lucky few. But, as Chuck Henry has noted, there is no ambiguity: the future of academic libraries and higher education rests on the ability to reconceive ourselves holistically. We shouldn't be doing this for the money; we should be doing it for the future of humanity. So lets thank Chuck, and all at CLIR, for a fantastic past, and let's steel ourselves, have a drink, and go and make a future together. Thank you.

Epilogue
Hidden Collections for Everyone

Michael Peter Edson, CLIR Presidential Fellow; Open Knowledge Foundation

Years ago, when I was working in my studio as a young, starving artist, I heard a radio interview with the renowned violinist and conductor Itzhak Perlman. At one point, Perlman was asked if he thought the highest levels of creative genius were the result of innate genetic gifts or if genius could be acquired by ordinary people through instruction and study—if genius could be taught.

Perlman's reply was warm and thoughtful, appreciative of both the magic of raw potential and the power of practice and discipline, but I remember feeling that there was something missing from the conversation. It seemed to me, as I stood there at my easel with a brush in my hand and the oblivion of an empty canvas before me, that the most important question wasn't *if genius could be taught*; it was whether we should lead our lives as if we were certain it was true.

Genius comes from unexpected places. When we first started putting collections online 10, 20, and even 30 years ago we made grand statements like "everyone in the world will be able to use these resources." We believed in the potential of "everyone" but based on what we built and shared I think that what we really had in mind was just more of the same kind of researchers, academics, and students we had always served. More of the same isn't bad—you can do a lot of good with more—but it's not everyone. Everyone—a global, connected everyone, full of potential—is something new, and it's a game changer, but most of the time that *everyone* seems more theoretical than real. It's hard to see.

I posted a job offer recently to Freelancer.com, an online marketplace of 16 million freelance workers from around the world. "Tell me what it would mean to you," I wrote, "if the resources of European and North American libraries, museums, and archives were made available to you, for free." I offered $25 to anyone who would send me a few paragraphs of writing or a minute or two of online video.

Almost instantly I was approached by remarkable people from all over the world: a teacher in Mexico, a young mother and weightlifting coach in the Ukraine, a medical student and his friends in Pakistan, a young businesswoman in Indonesia, and university students in Nigeria and Ghana. All of these individuals were experienced writers and researchers; all of them were entrepreneurs, using the web to find and serve clients almost everywhere the Internet can reach; and all of them seemed to respond, immediately and viscerally, to the hidden meaning of my question: can what we do here in the Hidden Collections program—among the libraries, archives, and museums of North America—be important to a wide and disparate audience? Can unhiding collections help people cultivate their own personal talents, regardless of who they are or where they live?

It turns out that if you're a curious, motivated individual in the Swat Valley or Ado-Ekiti, access to knowledge and information means *a lot* to you: it can, in fact, be central to your values, your education, your livelihood, and the way you explore and understand the world. "Knowledge is

a weapon to change the way of people's thinking," Kusmarni told me from Jakarta, in a sentiment that was echoed by everyone who answered my ad. But information, they all reminded me, is only valuable if you can get it, and use it, where you live.

Joseph, from the Ekiti State University in Nigeria and Owusu, from Donkorkrom, Ghana, told me of the insurmountable difficulties they and their classmates face when attempting to use Western collections: paywalls, geoblocking, and membership requirements routinely stand in their way. "You can imagine my frustration when I stumble across a site that has all the information I am seeking only to be informed that my geographic location does not have access to this," wrote Joseph. Owusu's stories echoed these frustrations, but he added hopefully, "To have an online digitized library that is free to be accessed is in my opinion the best thing that can ever happen to me and most of my mates at the University."

It is an article of faith in our community of scholars, institutions, and collections that we work for a noble purpose: that building, preserving, and disseminating knowledge is one of the defining acts of a wise and enlightened society. But our success in these endeavors will be defined by the choices we make—what we choose to catalog and digitize, what we share with the world, and under what conditions we permit or restrict access to our wealth of resources. Remarkable people like Kusmarni, Owusu, and Joseph seem to be everywhere—all over the globe as well as in our local classrooms and communities. Will we in the Hidden Collections program act as if we believe in their genius? We can and we should.

Appendix 1: Unconference Sessions

Data Tools

Workshop Session: Using Open Refine to Clean and Gather Data
Daniel Johnson

Workshop Session: Quick Tools for Visualizing Collections Information Geographically
Mitch Fraas

Discussion Session: GIS Mapping as an Access Tool
John Zarrillo

Metadata

Workshop Session: Description at Scale: Working with DPLA
Amy Rudersdorf

Discussion Session 1: Replicable and Scalable Metadata Methodologies
Jackie Dooley

Discussion Session 2: Sustainability of Web Presence
Karen Cariani and Jack Brighton

Discussion Session 3: The Challenge of Rich Media: Best Practices for User-Contributed Metadata
Sadie Roosa, Mark Cooper, Mark Williams, and Karen Cariani

Sustainability

Discussion Session 1: Collaborating between the Public and Private Sectors
Ramsay Thurber

Discussion Session 3: Sustainability of Collaborations
Margery Sly and Katie Rawson

Community

Workshop Session: Using Special Collections for K-12 Students
Tamar Dougherty

Discussion Session 1: Building an Inclusive Community of Archival Practice
Jack McCarthy and Celia Caust-Ellenbogen

Discussion Session 1: Alt-Archives: Community Based Video Collections
Rachel Mattson, Yvonne Ng, and Rebecca Fraimow

Promoting Collections

Discussion Session 2: Cataloging and Promoting Collections
Jordan Patty

Discussion Session 2: Working with Wikipedia
Annie Johnson

Donors

Discussion Session 2: The Living Donor and Archives: Strategies, Benefits, and Challenge
Stephanie Maxwell

Assessment

Discussion Session 3: Assessment and Evaluation of Projects
Naomi Steinberger and Christa Williford

Appendix 2: Posters

Eleven posters were presented at the symposium. The following are a few examples. The collection can be seen at http://www.clir.org/hiddencollections/2015-symposium-unconference/poster.

Appendix 3: Learning at Work in the Archives: The Impact of Access to Primary Sources on Teaching and Learning

"Learning at Work in the Archives: The Impact of Access to Primary Sources on Teaching and Learning" is part of a larger project, sponsored by the Council on Library and Information Resources, titled Observations on Scholarly Engagement with Hidden Special Collections and Archives. The goal of the study is to better understand how librarians and archivists are structuring and developing relationships with scholars in Cataloging Hidden Special Collections and Archives grant projects. The Scholarly Engagement project focuses on identifying and describing current practices, while also encouraging substantive conversation between librarians, archivists, and expert users about those practices. Among other observations, the Scholarly Engagement project has found that students—as project employees, emerging scholars and professionals, and citizens in their local communities—are playing a significant role in Hidden Collections projects, and that librarians and archivists are increasingly working with teaching faculty to incorporate primary sources into university and college curricula.

To explore these topics in greater depth, authors Kelly Miller and Michelle Morton examined a subset of six Hidden Collections projects focused on civil rights-related materials. These projects were located at a diverse group of institutions, directly connected to local communities and relevant to contemporary issues and curricula:

1. Emory University/Auburn Avenue Research Library on African American Culture and History–Archives from Atlanta, Cradle of the Civil Rights Movement (2008)

2. Robert W. Woodruff Library–Processing the Voter Education Project Collection (2008)

3. Amistad Research Center–Working for Freedom: Documenting Civil Rights Organizations (2008)

4. Arizona State University Libraries–Labor Rights are Civil Rights/Los Derechos de Trabajo son Derechos Civiles (2010)

5. Stanford University–Documenting Mexican American and Latino Civil Rights (2010)

6. University of Chicago/Black Metropolis Research Consortium–The "Color Curtain" Processing Project (2010)

Read the report.

Council on Library and
Information Resources

1707 L Street NW, Suite 650, Washington, DC 20036-4201
Tel: 202.939.4750 • Fax: 202.939.4765 • Web: www.clir.org