

# Failure Is an Option: The Experimental Archives Project Puts Archival Innovation to the Test

Mary O. Murphy, Laura Peimer, Genna Duplisea,  
and Jaimie Fritz

## ABSTRACT

This case study examines the mission and work of the Experimental Archives Project at the Schlesinger Library, Radcliffe Institute, Harvard University. The project developed from the first Radcliffe Workshop on Technology and Archival Processing in 2011, where a group of archivists and technologists brainstormed about innovative ways to scale up manuscript processing and improve access for researchers. The Experimental Archives Project, staffed by a team of Schlesinger archivists, tested direct-to-digital processing with no arrangement of physical materials, database-driven access to content, and systematic electronic redaction, all with the goal of improving processing rates and providing technologically enhanced access to researchers. The team completed five experiments. This case study explores in detail the challenges, failures, and successes of two of the projects: Redaction Redux with the Addenda Papers of Elizabeth Winship and Traditional-to-Tagging with the Records of *That Takes Ovaries*.

© Mary O. Murphy, Laura Peimer, Genna Duplisea, and Jaimie Fritz.



## KEY WORDS

Direct-to-digital processing, Electronic redaction, Arrangement  
and description

The Experimental Archives Project at the Schlesinger Library (Radcliffe Institute for Advanced Study, Harvard University) was a sandbox for archival processing experimentation. The Experimental Archives Project considered and tested ideas from outside of the archival world, as it looked to answer a primary question: if technology could allow archivists to speed up manuscript processing and do anything with archival materials, what would our users want us to do? One idea that helped guide the project's vision is the "push vs. pull theory," an established concept in the worlds of business and journalism. Applying this theory to archival processing and access means that "push" is traditional processing, during which an archivist or gatekeeper determines finding aid content and how information is described. "Pull" processing is user-driven content, where the user's demand determines how digital collections are arranged and delivered online. The project team sought to apply "pull processing" to archival work, where the user's demand drives the arrangement, placement, and importance of historical content online.

The Experimental Archives Project purposefully broke traditional processing rules and questioned archival standards to be creative—all with the goal of developing, implementing, and then scaling up archival innovations into everyday processing workflows. At its core, the project aimed to innovate archival processing by moving beyond theory, encouraging project archivists to act while stressing the importance of play, failure, and trial and error toward the development of successful innovation breakthroughs.

### Project History: Where Collaborative Creativity Came to Life

In the spring of 2010, the dean of the Radcliffe Institute, Barbara Grosz (now Higgins Professor of Natural Sciences at Harvard's School of Engineering and Applied Science) and Marilyn Dunn (executive director of the Schlesinger Library and librarian of the Radcliffe Institute) developed an idea for a multidisciplinary workshop that would encourage new thinking about archival backlogs, processing, and access. Dunn wanted to host a workshop that would be both collaborative and creative, from the planning stages to the event itself, where she hoped to bring together archivists and computer scientists to brainstorm about new ways to use technology to provide maximum access to archival collections and to create sustainable archival methods.

As a result of this idea, a carefully chosen, cross-disciplinary group of forty-five professionals attended the first Radcliffe Workshop on Technology and Archival Processing.<sup>1</sup> Attendees contributed expertise from various backgrounds, ranging from academia to industry. As an example, Richard Pearce-Moses, director, Master of Archival Studies Program, Clayton State University, provided his expertise on archival theory and practice, while Jon Kolko, founder and director

of the Austin Center for Design, provided advice on human-computer interaction. While all of the participants were chosen primarily because of their research specialties, they were also chosen based on their potential willingness to brainstorm creatively with other professionals and with the guarantee that they would not say, “yes, but let me tell you why this can’t be done.” At the Radcliffe Workshop, open-minded, innovative thinking was valued most of all.

The Radcliffe Workshop challenged technologists to think about how recent advances in automation and visualization could assist in the description process and archivists to re-envision their own practices. At the close of the workshop, a number of participants called for continued conversation on the topic and the development of pilot projects to experiment with strong, consensus ideas that emerged from the event.

Launched in autumn 2011, the Experimental Archives Project developed from the Radcliffe Workshop as a team of four archivists and two digitization assistants who worked on a rotating slate of processing experiments. Over the course of three years, the team completed in-depth work on five experiments, while Experimental Archives innovations are now being routinized across the Schlesinger Library.

For the project’s first experiment, Direct-to-Digital Processing with the Addenda Papers of Ida Pruitt and Marjorie King, the team tested the concept and procedures of processing. Instead of using traditional methods of arranging and describing the collection prior to digitization, the group flipped the process, imaging the unprocessed collection in-house prior to arranging it. The team then provided complete access to the collection online through a simple bibliographic record and the photo-sharing site Flickr.<sup>2</sup>

The second and third experiments explored alternative processing methods with two series from the recently processed papers of Shere Hite—feminist, researcher in sexuality, and author of the landmark *Hite Report on Female Sexuality* in 1976.

In the first experiment, Rethinking Redaction with the Hite Sexuality Questionnaires, the team studied a series of sex surveys that Hite circulated to women and men beginning in the 1970s. Because of the sensitive nature of the material and the fact that 80 percent of the responses were handwritten, the series would otherwise need to be closed without experimental intervention. Thus, the goal of the first project was to turn obstacles presented by these original materials into opportunities for new dimensions of access to digital surrogates. To complete the work, the team’s digitization assistants imaged the 4,416 questionnaires in-house, reading each of them and collecting related metadata in a Filemaker Pro database. They also redacted personal identifying information on the questionnaires through Adobe Acrobat as they imaged them. As a result of this study, the Experimental Archives team was able to provide

researchers with enhanced, searchable access to all of the questionnaires via a secure reading room laptop (the Schlesinger Library is also exploring ways to make this data available online).

In the second Shere Hite experiment, *New Methods for Newspapers*, the team re-evaluated the handling of newspaper clippings within collections. As with the other experiments, the group focused on imaging and metadata collection, this time applying OCR to the clippings for improved searchability.

The fourth and fifth experiments, *Redaction Redux* with the *Addenda Papers of Elizabeth Winship* and *Traditional-to-Tagging* with the *Records of That Takes Ovaries* are described in further detail below. The team documented all of its work through a Google-hosted wiki. Analysis and results from these final key experiments are proving to be valuable for informing future work.

### Innovation and Experimentation in the Archival Field: A Literature Review

Many archivists are developing or applying innovative approaches to different aspects of archival work, rethinking the relationship between collections and users, and enhancing their online presence. For instance, many archives are now using social networking sites, bookmarking and tagging, and uploading images to photo-sharing sites to deliver collections to users more effectively and collaboratively.<sup>3</sup> Similarly, multiple projects re-imagine the traditional archival finding aid in an effort to provide more dynamic access to collections. One example is the Next Generation Finding Aid Project, which came out of the University of Michigan's School of Information. The goals of the project were to move away from the flat, linear finding aid through a more robust application of EAD and by fully applying Web 2.0 technology to better connect users to archival content.<sup>4</sup> In addition, much research is available that relates to the core work of the Experimental Archives Project—specifically mass-digitization and digital access, testing digitization workflows, and rethinking physical arrangement in light of digital methods.

The research shows that researchers increasingly expect digital access to archival collections, and digitization will only grow as part of the archival landscape in the future.<sup>5</sup> Dennis Meissner and Mark A. Greene argued that “instead of dismissing researchers who want to see more of our collections on the Web, we must acknowledge that these expectations will be an increasing reality.”<sup>6</sup> According to Mats Dahlström, Joacim Hansson, and Ulrika Kjellman, digitization of archival materials poses a set of challenges to archival institutions, requiring them to “take on a much more explicit role of producing and shaping the cultural heritage in addition to its accustomed role of preserving it and making it available.”<sup>7</sup> Archivists and institutions respond to these challenges

by creating complex strategies that adapt their workflows and reconceptualize how to arrange, describe, and promote access to intangible, decontextualized materials, as demonstrated by several experimental digitization projects throughout the United States.

Developing digital archival processes requires consideration of how a collection's value and access goals will translate to the electronic environment. The sustainability of digital preservation programs depends on an infrastructure that can support management of multiple types of content and user experience of that content and that works to eliminate duplication of tasks.<sup>8</sup> However, longitudinal studies of digital preservation programs as described by National Archives and Records Administration lead security management and program analyst Shelby Sanett have shown insufficient management to sustain complex, long-term, and expensive efforts.<sup>9</sup> According to Ricky Erway, senior program officer in OCLC Research, archives should be developing workflows for ongoing programs, which means sustained project management, not special projects.<sup>10</sup>

Successful digitization workflows point to increasing research by users. Access is also the primary goal of digitization for Erway, who argued that "by increasing access we increase the perceived value of our collections. If we fail to make our collections better known, we may no longer have sufficient funds to, or even be employed to, continue collecting and preserving originals for our collections."<sup>11</sup> Erway and OCLC program officer Jennifer Schaffner pointed out that archival institutions can always make more preservation copies or turn to the originals as resources allow.<sup>12</sup> In this view, digitizing solely for preservation is a less valuable activity because it does not further access.

With these and other considerations in mind, institutions must begin to make practical decisions about how they will manage the complex task of digitization. For many, this starts with the question of processing. Larisa K. Miller has argued for eschewing physical processing in favor of creating a descriptive accession record that can be used as a basic search tool for digitized collections.<sup>13</sup> Meissner and Greene similarly stated that the need for initial physical processing is "a fallacy," saying that "it is simply not that difficult to find items if the description of series or files is done well."<sup>14</sup>

Eliminating lengthy physical processing provides an opportunity to devote more resources to digitization, allowing for "vast quantities of digitized primary materials" as opposed to "a few superbly crafted special collections," as Erway and Schaffner remind us.<sup>15</sup> However, many suggest that digital processing will not become more widely applicable unless archivists are allowed to determine collection needs on a case-by-case basis. In her discussion of OCR in digital collections, Oya Rieger, associate librarian at Cornell University, posed the option of allowing mistakes to pass, returning to correct them only when users or extra time dictate.<sup>16</sup> The need for item-level metadata is similarly debated, with many

echoing Sarah Sutton's assertion that collection- or file-level metadata should suffice unless collections see heavy use, or a unique opportunity for metadata creation exists that will not impact any additional digitization efforts.<sup>17</sup>

Finally, decisions must be made at the institutional level about what kinds of technology and workflows will be used to support these projects. In the spirit of quantity over "boutique" quality, several institutions have favored the kind of "prosumer" technology described by Ricky Erway, designed to allow for high levels of production while providing adaptability in terms of time and image quality.<sup>18</sup>

The goal of streamlining workflows drives the Digital Southern Historical Collection project at the University of North Carolina, which is digitizing sixteen million items from 4,600 manuscript collections. Ricky Erway wrote that "a significant feature of this program is the care with which filenames are determined in advance; metadata is extracted from the finding aid to form folder-level metadata to describe all the scans for that folder. The finding aid provides description and enables discovery and links to the images."<sup>19</sup> Organized digitization, supported by detailed workflows, has the potential to add descriptive value to collections without the weight of item-level description.

The Augmented Processing Table (APT) project, based out of the University of Texas, is an innovation with the goal of expediting archival work, clearing backlogs, and making collections available to users much sooner and online. A collaboration between archival science and human-computer interaction, APT restructures the processing workflow of born-digital and digitized materials into a single stream using a large multitouch table.<sup>20</sup>

In their "bare essentials" version of MPLP (More Product, Less Process), Meissner and Greene described many of the aspects of digitization that archivists may need to consider. Establishing a "minimum level of work" and providing "the most material available in a usable form in the briefest time" may seem anathema to the traditional responsibilities of processing, but they are proving to be essential in successful digitization experiments.<sup>21</sup> Additionally, many of the existing case studies support the need for flexibility and case-by-case decision-making. As more institutions begin to support high-throughput digitization, we hope additional literature will guide archivists toward a set of general, adaptable standards for the creation and management of digitization programs.

### Experimental Archives at the Schlesinger Library: Project Evaluations

Research and case studies certainly guided the Experimental Archives Project, shaping both the team's thinking and workflows, especially during the planning phase and outset of each experiment. With that said, it is the team's hope that the following detailed descriptions of two experiments—Redaction Redux with the Addenda Papers of Elizabeth Winship and Traditional-to-Tagging

with the Records of *That Takes Ovaries*—will in turn guide other archivists in their experimental projects by illustrating challenges, failures, and successes.

## Redaction Redux with the Addenda Papers of Elizabeth Winship

### BACKGROUND

From 1963 to her retirement in 1998, Elizabeth Winship wrote an advice column for teenagers, called “Ask Beth.” She was a popular columnist, due mainly to her sensible and thoughtful approach to teen questions and partly from the lack of other advice outlets for teens on sex and relationships, particularly during the early years of her column. The column started in the *Boston Globe*, and, in 1970, the *Los Angeles Times* Syndicate picked it up. At its peak, “Ask Beth” had seventy subscribing newspapers. Starting in the 1980s, Winship’s daughter Peg assisted her in writing responses for the column. A family therapist, Peg Winship then signed on as coauthor in 1993 and continued the column on her own from Beth’s retirement in 1998 until 2007.

In 2008, an archivist from the Schlesinger Library processed the original collection of Elizabeth Winship papers, which consist of nearly five linear feet of letters to the “Ask Beth” column. At that time, she arranged the collection using traditional processing methods. The archivist sorted the letters manually into a chronological arrangement, physically refoldered materials, and reviewed all of the letters, redacting the names and addresses from particularly sensitive ones.

Peg Winship then donated an addendum of Winship papers in 2009, and it is this collection that became the focus of Redaction Redux. The addendum consists of two linear feet of papers and contains letters submitted to the “Ask Beth” column; copies of messages sent to the column from readers through the Internet service provider Prodigy in the 1990s; as well as research materials, notes, and some professional letters from organizations.

### GOALS AND CHALLENGES

The goals of the Winship project were to use digitization to simplify processing, to provide complete digital access to a recent accession of an archival collection, and to more efficiently remove obstacles to access due to the presence of sensitive materials. Finding ways to complete the project in a quick and efficient way was another important goal which informed much of the process and was the reason for rethinking workflows when confronted with unforeseen complications and time-consuming tasks.

Redaction Redux combined two workflows from previous projects: digital redaction of sensitive materials, along with the recording of item-level metadata,

which the Experimental Archives team first accomplished in the Shere Hite sex surveys project, and direct-to-digital processing of a collection that the team had applied to the Ida Pruitt and Marjorie King papers two years prior. As with the sex surveys, most of the Winship letters are handwritten, so optical character recognition was not an option to help in searching the digitized documents. Direct-to-digital processing of a collection meant no refolding, sorting, file renaming, or any conservation work, such as photocopying acidic materials or flattening oversize materials.

However, the Winship letters posed new challenges in their various forms and content, with each letter providing different details to be mined for metadata. The team adapted workflows for imaging, redacting, and recording metadata several times throughout this experiment, capturing and recording over six thousand files into a searchable database—comprising over seven thousand individual images or “pages.”

#### PLANNING AND IMAGING

Peg Winship donated the addendum in folders and in good condition, so the collection required little preparation, outside of an initial review prior to imaging. The Experimental Archives team chose to use straightforward, out-of-the-box technology—a digital camera and camera stand, off-the-shelf software like Adobe Photoshop (for cropping or editing images, if necessary), Adobe Acrobat Pro (for digital redaction), Microsoft Excel, and eventually Filemaker Pro (for the capture of metadata).

The team’s digitization assistant then digitized the materials in batches. This stage in the process proved to be the least time consuming because the team had previously developed specific settings for the camera and methods of image cleanup in Adobe Bridge and Adobe Photoshop.

However, this collection posed some new challenges, particularly with capturing content written in pencil or printed onto glossy fax paper, which had faded over time. Under the powerful photographic lights, already pale or faded text washed out. For these materials, the team’s digitization assistant needed to significantly edit the digital surrogate in Photoshop to correct contrast and tone. This image correction to improve legibility privileged content over preservation of the appearance of the physical item. The team’s decision to value access over authenticity of the original is evident in many of the Experimental Archives’ projects, but it has also been an ongoing topic of discussion.

Imaging the printouts of Prodigy emails was also a time challenge. The printouts consisted of multiple reader email messages on a single page. The team determined that researchers would find it most helpful if we created separate images with associated database records for each email and

not per page—allowing for more accurate metadata and searchability. This resulted, however, in the digitization assistant having to take many images of each page, which added a significant amount of time to imaging and data recording.

## METADATA

Initially, the team determined that capturing robust metadata would aid in searching, but after extensive data gathering proved complicated, inconsistent, and slow, the team greatly modified its data collection method midway through the project.

The metadata that the team recorded included a unique ID for each item and whether or not items had been redacted; physical location (box and folder) of each item; subject terms; location such as city/state; gender; age of writer (if available); and a general description/comments field (if needed). The database and the metadata allow for browsing the materials in a similar order to their original arrangement, as well as for some searching and sorting for more efficient and customizable research.

It is important to note that the metadata is incomplete due to the lack of information on the original documents. For instance, 2,235 records do not have dates because the dates are not noted on the original letters. Particularly regarding age and gender, the team had to be careful not to assume gender or age information from the letters if it was not explicit.

Subject terms also proved problematic. Because many of the folders that contained letters had vague titles, such as “Letters Done,” nothing indicated the scope of topics that might be represented in the letters. The team found it nearly impossible to create a detailed list of subject terms (from folder headings) prior to reading the letters, so the digitization assistants created an authority list as they imaged. The lead assistant found it necessary to record up to three subject terms per item, due to the complexity of some of the subject matters. As a result, within the first few weeks of the project, the list of controlled vocabulary terms quickly grew to over two hundred subjects. This effort to create and apply subject terms required the assistants to constantly compare and revise the vocabulary, which was time consuming.

The process was also subjective. Subject areas that were not straightforward in the letters were sometimes defined differently depending on who recorded the information. For instance, a letter from a child complaining about his parents spanking him may have been identified by the assistants as “abuse” or “strict parenting” depending on how they interpreted the letter.

This is a good example of how the process for one project does not always translate well to another. For instance, the Experimental Archives team collected

metadata for the Hite surveys from what was already written on the documents either by the respondent or by Hite. In the Winship project, the team found that subjective analysis was not helpful to working quickly through the project and potentially not helpful to researchers due to inconsistency in the application of terms.

To remedy the difficulties related to interpretation and to make the final product available more quickly, the team decided to stop collecting subject terms, thereby limiting access points to a level more closely aligned with the original Winship papers that were physically processed back in 2008. While the digitization assistants still captured some metadata when they were clear and readily available, such as folder title or age, gender, and date, they eliminated the time spent thinking about and applying subject terms to each new record.

## REDACTION

The Experimental Archives team determined that it was important to redact the digital collection extensively. Assuming that the digitally copied letters will eventually be made available globally through a Web-based delivery system, the team paid close attention to privacy concerns. Thus, the digitization assistants redacted the majority of names even if they deemed the subject being expressed as not highly sensitive. They also redacted information such as addresses, school names, phone numbers, and Prodigy IDs.

Some redaction issues were unforeseen. Unlike the body of creators in previous experiments, the writers of letters to the “Ask Beth” column interacted with each other somewhat, whether responding to previously published letters or responding directly through the Prodigy service. The Prodigy messages in particular presented a loss of usable information upon the redaction of names. Our lead digitization assistant on the project, Genna Duplisea, discussed this on the project wiki. She wrote:

Prodigy emails pose an interesting problem of authorship. It is often evident that the same person wrote to Winship multiple times. [It is also apparent that at times,] at least some of their problems and identities may have been invented. I have been trying to note in the database items that share a Prodigy sender ID while still redacting.<sup>22</sup>

Thus, while the team worked to retain the useful parts of the redacted information, once again it took additional time to find those connections and record them.

## DELIVERY

The redacted Winship images are currently available via a laptop computer in the Schlesinger Library reading room, although it is a goal to make them available through an online delivery system in the future. In addition to creating a user-friendly layout in Filemaker Pro, the team created a user guide for the database to help researchers navigate the digital collection. The team also designed a feedback form through the Springshare product, LibSurveys, which users can access by clicking the Feedback Form button available on every record in the database. One question on the form asks users if they would be interested in tagging subject terms or adding data to other fields in the records as they use them. A crowdsourcing functionality in any delivery system would be a potential help in freeing up the data-gathering time of the digitization assistants. The team recently made the digital product available to the public, so we do not yet have the data needed to fully assess the project and how we could improve it.

## ANALYSIS

Among the lessons learned from Redaction Redux are three points described below. They represent just a few of the takeaways from this work, but they will surely help inform future innovation projects.

1. While it is important to try to stay on track with the original project goals, flexibility is key. Re-evaluating certain steps midway through a project if the work is getting bogged down, or trying to determine what is working and what is not at any point, is useful. The Winship metadata journey was a great learning experience related to this.
2. It is important not to assume that processes that work well for one collection will work as well for another. Every collection is different and knowing that at the beginning can help with reducing some of the troubleshooting work that may need to occur.
3. To move as quickly as possible through a processing project, whether it is digital or not, less process (such as recording less metadata) will indeed translate to a project completed more quickly. This is an obvious point, but it is often hard to limit oneself with digital projects where item-level metadata and, therefore, improved access, can be very doable. How valuable is the loss of extra access points and more searchable information? We will find out when users start working with the Winship addendum.

## Traditional-to-Tagging with the Records of *That Takes Ovaries*

### BACKGROUND

*That Takes Ovaries* is a book, a play, an “open-mic” movement, and a non-profit organization dedicated to issues of women’s equality and empowerment. Donated by founder Rivka Solomon in 2011, this relatively small collection (two cartons) consists primarily of materials generated to promote *That Takes Ovaries* events, such as programs, posters/fliers, and reviews. It also contains a significant number of newsclippings, as well as correspondence, notes, and photographs. Also included are personal stories and release forms from participants at *That Takes Ovaries* events.

By the time the Experimental Archives team selected the records of *That Takes Ovaries* as an upcoming project, the team had already dealt with a number of unique challenges in digitizing collections. They had worked with fragile materials, ameliorated serious redaction concerns through new workflows, and tested a number of delivery methods. However, the team had yet to work with a collection that arrived in a completely disorganized state. As a result, the team determined that the lack of original order found in *That Takes Ovaries* was both a challenge and a welcome opportunity to get creative with “flipped” processing methods. It also allowed the team to question the way researchers gain access to the physical contents of digitized materials.

### GOALS AND CHALLENGES

Because *That Takes Ovaries* is a relatively small collection that contains a wide variety of materials and subjects as well as no original order, the Experimental Archives team considered it to be an interesting candidate to challenge traditional processing and arrangement methods. The goal of the project was to flip traditional methods by taking the unprocessed collection and digitizing it first, before arranging it or imposing any organizational structure. This challenged the concept of processing and moved it from the physical realm to the digital, with the arrangement and description of digital surrogates only.

### PLANNING AND IMAGING

Instead of following traditional processing procedures of surveying the collection and then physically moving files into an arrangement, team archivists asked the digitization assistant to photograph all of the items in the collection from front to back while not imposing any prior physical order. Rather, the plan was to leave the material as is and instead, only photograph the collection and

then arrange the digital surrogates into rough groupings using desktop folders. Once sorted, the assistant was to upload the groupings, or “sets,” into a new collection page within the Experimental Archives Flickr Pro account and tag each item with descriptive metadata.

The team was able to follow the original plan throughout the experiment. When preparing for digitization, the Experimental Archives assistant did separate a large chunk of newspapers and magazines from all of the other content to minimize the need to adjust the camera settings repeatedly throughout the imaging process. But otherwise, she let the collection remain in its original state. And, as the imaging process began, the team continued to plan and brainstorm by investigating options for content delivery, to conceptually build a tag library, and to consider how to address any sensitive materials that might be hidden in the collection. The team also discussed whether or not researchers would want to go back to the original materials once they were digitized and, if so, whether they would be willing to dig through a carton-sized container to do it.

The digitization assistant completed the initial imaging of this collection in a relatively short amount of time—photographing and editing over a thousand items in the course of two days. Because the team had already developed an efficient imaging workflow with earlier experiments, the assistant was able to spend less time on imaging and more time on the issues of tagging and arrangement unique to this collection.

Some items did present new challenges for the team, such as a significant amount of memorabilia like buttons, clothing, and even a perishable chocolate “ovary award.” While these items ended up requiring more attention in the physical realm—as they had to be cataloged and stored separately from the paper materials in the collection—the team decided to treat all of the materials similarly during the imaging process by simply capturing photographs of every item.

The team also knew that Flickr—the project’s method of content delivery—would impose some restrictions on file formats and that only single-page images could be presented on its site. Knowing this, the team chose to save separate copies of multipage images as single-page, Flickr-friendly JPGs and as multipage PDFs locally. Being able to quickly create copies of an object in multiple formats has been one of the greatest benefits of digital processing.

## METADATA AND ORGANIZATION

Knowing that the product of this type of digitization would be a mass of raw materials, the team faced many questions regarding how to transfer the traditional processing tasks of arrangement and description into the digital world. Though similar experiments had given the team confidence that it would be possible to make the collections accessible and easily navigable, the hope was

that the ease and relative speediness of digital processing might provide added value without expending too much additional effort.

The first question, given that the materials did not arrive in any original order, was how to group the images and present them as an arranged collection. The priority was to ensure that the digital collection could be easily browsed, so that users with little or no knowledge of the collection would still be able to find interesting materials. The team considered traditional processing methods at this point, reflecting on procedures that require an archivist to predetermine a method of arrangement and then moving the physical materials around to reflect that decision. If that subsequent arrangement did not suit the needs of the library or its users, the materials would have to be physically and conceptually reorganized.

The Experimental Archives team found that imaging the records before attempting any arrangement was liberating to the process. Once digitized, the thousand-plus items, including memorabilia and posters, could simply be moved around on a desktop to test the effectiveness of various arrangements. This ease of movement—and the loss of pressure to have the perfect structural outline in place before moving objects—freed the group to experiment with different types of organization to find the one that worked best. In a matter of hours, the digitization assistant arranged the digitized collection into series, including correspondence, events, and publicity, with subseries presented as nested folders.

The opportunity to add value to the collection became most apparent when discussing options for describing the materials. Typically, descriptive efforts would result in the creation of a finding aid, including scope-and-content notes and container lists. However, given the design of a Flickr Pro account, which features robust tagging functionality and the ability to frame digital images into a collection with descriptive text blocks placed along the top and sides of each main collection page, the team decided to use that as its primary descriptive method. It was also extremely easy to add direct links back to the traditional Harvard catalog record within this same area of the collection's homepage.

Flickr also allowed the Experimental Archives team to do the scope note differently. Instead of providing a detailed written description about the collection, the team chose to imbed a short, to-the-point introductory video by the archivist. This allowed her to simply talk directly to users about the collection. In this way, the team leveraged creativity by modeling this section not from the world of archival management, but instead from airline safety videos, which are meant to quickly welcome passengers on board while communicating just the must-know information.

## REDACTION AND DELIVERY

The Experimental Archives team was less concerned with digital redaction when processing this collection. The few items that required redaction included participant release forms from *That Takes Ovaries* events, as well as stories submitted by individuals either for publication in the *That Takes Ovaries* book, or for readings at open-mic nights. Although these participants gave the information willingly, the team felt the contemporaneous nature of the collection necessitated protection of their personal information. Furthermore, some contributors were minors who submitted their personal stories for afterschool projects and workshops. In these cases, the team saw the need for redaction of any and all personal information.

Delivery was as simple as uploading the digital content to Flickr. The team then used the platform's batch editing tools to sort content into sets and tag images with relevant metadata, including locations. With those tags, Flickr's mapping feature also proved to be beneficial as it presents tagged items from the collection on a searchable map, showing users the global impact of *That Takes Ovaries*.

Within the system, tags can be searched using a typed query, or by clicking the tag on an image in the collection. Sets can be browsed individually, and all images are available at multiple resolutions. In all, Flickr proved to be an extremely useful tool in facilitating quick, simple access to the digital collection. Working with a site that is so widely used, and whose format is becoming ubiquitous for digital presentation of images, also allowed the team to capitalize on user familiarity. While many less-seasoned researchers have difficulty navigating traditional content management systems, Flickr's interface should be relatively familiar to the average Internet user.

## ANALYSIS

As with work on Redaction Redux, the Experimental Archives team learned a number of lessons from Traditional-to-Tagging with *That Takes Ovaries*. These lessons represent just a few of the many takeaways gathered from this experiment, but they continue to inform future projects.

1. When thinking about digitization, consider letting go of physical processing for certain collections. Online arrangement and description can be sufficient. And there is value to embracing the flexibility of this more customizable, item-level approach to access. With many delivery systems available today, researchers now have the capability to re-sort and create new groupings as well as to add tags and descriptions of their own to digital images.
2. When asked how to gain access to the original materials, do not be reluctant to tell users that they may have to dig. For those who still

need to see the physical object, they can do so because digitized images carry item IDs that tie the objects to specific collection and container numbers. However, with direct-to-digitally processed collections, users may be asked to search through one box of materials. This simply acknowledges a shift in focus, as more researchers seek full access to collection content online.

3. When faced with lack of resources or institutional resistance to new methods, seek creative solutions. Being a commercial product, Flickr may not be the ideal system to present collections, but it is available, it is inexpensive, and it is used by approximately eighty million people. At the end of the day, fast, efficient access should be paramount.
4. At the same time, a powerful, noncommercial and customizable delivery system may be preferable to a tool like Flickr. In 2013, the Experimental Archives Project began working with computer programmers to create a new platform for Schlesinger digital collections that meets the library's specific needs.<sup>23</sup>
5. As previously noted, obtaining user feedback is both desirable and useful. Library professionals need to know what researchers want and what they can accept. For instance, some questions that researchers can help to answer include what types of metadata best facilitates access? Are digital delivery platforms intuitive and are there functionalities that would improve the researcher experience? And finally, are users willing to dig for a physical document when necessary to gain item-level access more quickly and online?

## Conclusion

The time results of these and other work of the Experimental Archives Project vary widely. For instance, Redaction Redux with the Addenda Papers of Elizabeth Winship was processed (imaging, database record creation, and item redaction) over twenty-eight working days. The project took more time than expected, possibly due to the changes in workflows and other related challenges. Traditional-to-Tagging, on the other hand, was digitally photographed and uploaded into Flickr in approximately five days. The final product is more easily accessible to researchers with nearly comparable processing times to traditional physical processing workflows. The library is now moving forward by applying things that worked well during the experiments and integrating them into routine processing practices, such as digital redaction on a larger scale. Experimental work also continues in the areas of born-digital processing, pattern-recognition technologies for handwritten material, collecting and repurposing user-generated images of archival materials, and developing and testing

various delivery platforms—all with the goal of discovering the best tools available to meet researchers' needs. The library is now experiencing a new comfort level with emerging technologies. Staff continue to be committed to testing new, technology-driven processing methods and to a philosophy of exploration and collaboration with colleagues at Harvard University, from across the United States, and around the world. A few years ago, a researcher who used the digitized version of the Shere Hite questionnaires commented that digital access to manuscript collections is a fundamental step in the right direction, but that there is a world of possibilities for archivists still to explore.

## NOTES

- <sup>1</sup> Schlesinger Library, Radcliffe Institute for Advanced Study, Harvard University, May 16–17, 2011.
- <sup>2</sup> Ida Pruitt and Marjorie King Papers, 1891–1994 (MC 701), [https://www.flickr.com/photos/experimental\\_archives/collections/72157629840700987/](https://www.flickr.com/photos/experimental_archives/collections/72157629840700987/).
- <sup>3</sup> Some relevant case studies relating to the application of social media in archives can be found in *A Different Kind of Web: New Connections between Archives and Our Users*, ed. Kate Theimer. (Chicago: Society of American Archivists, 2011).
- <sup>4</sup> Elizabeth Yakel, Seth Shaw, and Polly Reynolds, "Creating the Next Generation of Archival Finding Aids," *D-Lib Magazine* 13 (May/June 2007).
- <sup>5</sup> Amanda Focke, "More Pixels, Less Process: Decision Making for Minimal Processing Digitization" (presented at Society of American Archivists Annual Meeting, San Francisco, 2008), <http://www.archivists.org/conference/sanfrancisco2008/docs/session701-focke.pdf>.
- <sup>6</sup> Dennis Meissner and Mark A. Greene, "More Application while Less Appreciation: The Adopters and Antagonists of MPLP," *Journal of Archival Organization* 8, nos. 3–4 (2010): 195–96, doi:10.1080/15332748.2010.554069.
- <sup>7</sup> Mats Dahlström, Joacim Hansson, and Ulrika Kjellman, "'As We May Digitize'—Institutions and Documents Reconfigured," *Liber Quarterly: The Journal of European Research Libraries* 21 (April 2012): 456.
- <sup>8</sup> Bradley J. Daigle, "The Digital Transformation of Special Collections," *Journal of Library Administration* 52 (April 2012): 258, doi:10.1080/01930826.2012.684504.
- <sup>9</sup> Shelby Sanett, "Archival Digital Preservation Programs: Staffing, Costs, and Policy," *Preservation, Digital Technology and Culture* 42, no. 3 (2013): 140.
- <sup>10</sup> Ricky Erway, "Supply and Demand: Special Collections and Digitisation," *Liber Quarterly* 18 (November 17, 2008): 324–36.
- <sup>11</sup> Erway, "Supply and Demand," 326.
- <sup>12</sup> Ricky Erway and Jennifer Schaffner, "Shifting Gears: Gearing Up to Get into the Flow," OCLC Online Computer Library, Inc., 2007.
- <sup>13</sup> Larisa Miller, "All Text Considered: A Perspective on Mass Digitizing and Archival Processing," *The American Archivist* 76 (Fall/Winter 2013): 521–41.
- <sup>14</sup> Meissner and Greene, "More Application while Less Appreciation," 195.
- <sup>15</sup> Erway and Schaffner, "Shifting Gears," 6.
- <sup>16</sup> Oya Y. Rieger, "Enduring Access to Special Collections: Challenges and Opportunities for Large-Scale Digitization Initiatives," *RBM: A Journal of Rare Books, Manuscripts and Cultural Heritage* 11 (March 20, 2010): 11–22.
- <sup>17</sup> Sarah C. Sutton, "Balancing Boutique-Level Quality and Large-Scale Production: The Impact of 'More Product, Less Process' on Digitization in Archives and Special Collections," *RBM: A Journal of Rare Books, Manuscripts and Cultural Heritage* 13, (March 20, 2012): 50–63.
- <sup>18</sup> Erway, "Supply and Demand," 327.

- <sup>19</sup> Ricky Erway, "Rapid Capture: Faster Throughput in Digitization of Special Collections," OCLC Online Computer Library, Inc., 2011: 16.
- <sup>20</sup> Jeff Crow, Luis Francisco-Revilla, April Norris, Shilpa Shukla, and Ciaran B. Trace, "A Unique Arrangement: Organizing Collections for Digital Archives and Libraries," *Theory and Practice of Digital Libraries (TPDL) 2012: Proceedings of the Second International Conference*, ed. P. Zaphiris, G. Buchanan, E. Rasmussen, and F. Loizides (Berlin: Springer-Verlag, 2012), 335–44.
- <sup>21</sup> Meissner and Greene, "More Application while Less Appreciation," 175.
- <sup>22</sup> Project wiki: <https://sites.google.com/site/experimentalarchives/home>.
- <sup>23</sup> The Blackwell Family, Charlotte Perkins Gilman, and Dorothy West Collections now available through the Schlesinger Library Digital Collections Suite.  
<http://schlesinger.radcliffe.harvard.edu/onlinecollections/blackwell/>  
<http://schlesinger.radcliffe.harvard.edu/onlinecollections/gilman/>  
<http://schlesinger.radcliffe.harvard.edu/onlinecollections/west/>

## ABOUT THE AUTHORS

---



**Mary O. Murphy** is the administrative librarian at the Schlesinger Library, Radcliffe Institute, Harvard University. She also served as the project manager for the Experimental Archives Project and the Radcliffe Workshop on Technology and Archival Processing. Before moving into library management, Murphy began her career as an archivist working with organizational records and personal papers relating to the history of women in America. Prior to joining the Schlesinger Library in 2008, Murphy served as archivist and records manager at the Boston Redevelopment Authority, where she confronted challenges associated with the management and preservation of backlog and electronic records.



**Laura Peimer** is an archivist and a former team member of the Experimental Archives Project at the Schlesinger Library, Radcliffe Institute, Harvard University. She has worked at the Schlesinger Library since 2008 where her primary responsibilities include processing and public services. Prior to this, she worked as an archivist at Harvard Business School and the Brooklyn Museum, and as the photo archivist at the American Jewish Historical Society. She received a master's in history with a certificate in archival management from New York University.



**Genna Duplisea**, a former digitization assistant at the Schlesinger Library's Experimental Archives Project is the archivist and special collections librarian at Salve Regina University. She earned her master's degrees in archives management and history from Simmons College in 2013 and previously earned her bachelor's degree in English from Bowdoin College and a certificate of graduate study in nonprofit management at the Muskie School for Public Service at the University of Southern Maine. In addition to her archival work and interests in pedagogy and digital humanities, she pursues historical research on gender, empire, and knowledge production.



**Jaimie Fritz**, a former digitization assistant at the Schlesinger Library's Experimental Archives Project, holds a BA in English literature from Smith College and an MS in library and information science, with a concentration in archives management, from Simmons College. She currently serves as project archivist at the Bentley University Library. Her major research interests include digital preservation life-cycle management and the expansion of access to archival collections.