Developing a Practical Electronic Records Transfer Workflow for Wisconsin Public Records

By Hannah Wang and Laura Poplett

ABSTRACT: This article chronicles the development of a streamlined workflow for state agencies to follow when transferring electronic public records to the State Archives at the Wisconsin Historical Society. The workflow, which utilizes an open source application and SFTP (Secure Shell File Transfer Protocol) for remote transfers, was developed to establish a secure and standardized process for transfers to replace the *ad hoc* manner of earlier transfers. The authors describe the genesis of the workflow and the significant challenges faced during implementation, including differences in technology usage and organizational culture, and they discuss the lessons learned from the project and plans for future work.

Introduction

Between 2016 and 2018, the Wisconsin Historical Society (WHS), with generous funding support from the National Historical Publications and Records Commission (NHPRC), developed a stronger and more practical electronic records workflow to facilitate the transfer of public records with historical value. Prior to 2016, state agencies transferred electronic records to the WHS via physical media using makeshift workflows. However, the vastly expanded volume of public records in digital formats made the *ad hoc* transfer of electronic records via external hard drives and CDs increasingly unsustainable. Furthermore, no protocols existed for verifying the integrity of the files transferred. A streamlined transfer option was long overdue. Establishing a remote transfer workflow involved identifying a usable and standards-compliant software tool, developing a workflow that identifies the responsibilities of both state agencies and the WHS during the transfer process, and working with partner agency staff to implement the tools and workflows on site.

This article will provide an overview of the workflow development, a discussion of the challenges faced during testing and implementation, and plans for future work.

Background

Establishing a transfer workflow was just one part of a larger grant project funded by the NHPRC, entitled the Wisconsin State Preservation of Electronic Records (WiSPER) Project. The NHPRC's stated goal was to fund projects that "actively engage in activities that preserve electronic record assets and make them available for public discovery" and "bring together multiple parties to plan for and carry out project

goals and outcomes."[1] The WiSPER Project was a collaborative interagency effort that ran from October 2015 through March 2019 and resulted in a more efficient and fully established electronic records preservation and access program at the WHS. The WiSPER Project addressed several overarching goals:

- Increase the percentage of scheduled state agency electronic records with historic value that are transferred to the WHS;
- Document the processes and tools used to identify and schedule electronic records, develop transfer workflows, and train employees in electronic records governance best practices; and
- Improve discovery and access to electronic records in the repository through the development of links from cataloged series and web-based access points to records collections.

The development of a secure and easy-to-use transfer workflow is a part of the second goal and is the focus of this article. This process required collaboration between the WHS and its state agency partners as well as significant research into the technical options available for remote transfer. The WiSPER Project team developed a workflow compliant with relevant archival standards, affordable for all agencies, and easy for state records officers to utilize in their daily work.

Several projects laid the foundation for the WiSPER Project's work. The Library of Congress (LOC) has been a leader in the development of standards and tools for packaging data. Work by the LOC and the California Digital Library led to the development of the BagIt specification, a hierarchical packaging standard widely used by libraries and archives for data storage and transfer. The LOC has since developed a suite of open source tools that implement BagIt, including Bagger, which provides a graphical user interface for creating and transferring BagIt-compliant packages (also known as "bags").[2] Additionally, some government archives, including the National Archives and Records Administration, the Washington State Archives, and the Hawaii State Archives, have developed homegrown software tools to package and transfer electronic records remotely via SFTP and HTTPS (Secure Shell File Transfer Protocol and Hypertext Transfer Protocol Secure).[3]

Several other state archives have taken steps to simplify and secure the transfer of electronic records to the archives. Institutions such as the Texas State Library and Archives Commission and the State Historical Society of North Dakota have not implemented a specific software tool, but clearly describe several methods for electronic records transfer to their repositories, such as FTP/SFTP and removable media.[4] Others, including the State Archives of North Carolina and the Vermont State Archives, utilize Bagger to package and transfer electronic records on removable media.[5] The WiSPER Project team leveraged portions of this preexisting work and documentation to select and implement a transfer tool and workflow.

Transfer Workflow Development

From the outset of the WiSPER Project, the team knew that the transfer workflow would have to meet several requirements, including

1. Transferring data in a manner compliant with relevant industry standards;
2. Allowing the remote transfer of electronic records without the use of external storage media such as flash drives or CDs;
3. Performing automated validation of transfers and notifying both parties; and
4. Requiring minimal setup and technical knowledge from the end user.

Initially, the project team planned to develop a homegrown transfer tool that satisfied these requirements. However, the WHS IT team was restructured and reduced toward the beginning of the WiSPER Project, and these changes meant that the project team no longer had a dedicated software developer as originally planned. The remaining members of the project team did not have the technical skills required to develop a software application and instead had to select a preexisting "off-the-shelf" tool.

This change coincided with the release of a new tool developed by AVP (previously AVPreserve), a consulting and software development firm that supports information management, particularly in libraries and archives. AVP developed a tool called Exactly in partnership with the Louis B. Nunn Center for Oral History at University of Kentucky Libraries. AVP describes Exactly as a "free and open source application that is easy-to-use for remotely and safely transferring any digital data from a sender to a recipient."[6] The project team learned about Exactly at the 2016 NDSA Digital Preservation conference and began reviewing the tool against their requirements, also comparing it to Bagger, the open source Library of Congress tool that has an established user base in libraries and archives.

Requirement 1: Transferring Data in a Manner Compliant with Relevant Industry Standards

Exactly packages and sends data by creating BagIt-compliant bags, discussed above. Bags include both descriptive metadata and a file manifest, which allows users to manually or programmatically decipher the contents and verify their integrity.[7] WHS staff determined that bags could be configured to comply with both PAIS (Producer-Archive Interface Specification; ISO 20104:2015) and PAIMAS (Producer-Archive Interface Methodology Abstract Standard; ISO 20652:2006) international standards.[8]

Requirement 2: Allowing the Remote Transfer of Electronic Records without the Use of External Storage Media such as Flash Drives or CDs

The ability to transfer electronic records remotely was one benefit of using Exactly instead of Bagger. Because Exactly is built explicitly to facilitate transfer, rather than just package data, the tool easily integrates with several data transfer mechanisms. While not impossible to transfer data using Bagger, it would require additional setup and

integrations for both WHS and partner agencies. Exactly facilitates remote data transfer through a variety of methods including FTP, SFTP, Google Drive, and Dropbox. Because WHS IT had a license for Dropbox for Business, the project team initially decided to leverage Exactly's Dropbox integration for electronic records transfers. The workflow would later move from Dropbox to SFTP for transfers, as discussed later in this article.

Requirement 3: Performing Automated Validation of Transfers and Notifying Both Parties

Upon completion of a successful transfer, Exactly has the ability to send automatic e-mail notifications to the sender and recipient, which include a summary of the transfer and a manifest of the files included. The recipient can then use Exactly to validate the transfer to ensure that all files have transferred successfully without data loss or corruption. Bagger can also validate bags, but does not have the option to configure automatic e-mail notifications.

Requirement 4: Requiring Minimal Setup and Technical Knowledge from the End User

Exactly is available for free download from AVP's website for use in Mac or Windows operating systems. AVP provides clear documentation on its website for downloading and installing the software, although the project team needed to adapt the documentation to the specific systems requirements of Wisconsin state agencies. Exactly provides a graphical user interface for packaging and sending bags, which agency records management staff needed only minimal training to use. When compared to Bagger, which also provides a graphical user interface for creating and transferring bags, the project team found the interface for Exactly more intuitive and friendly for users outside of the library and archives professions. Additionally, testing by project staff found that Exactly is easier to download and install than Bagger.

Having selected a tool for electronic records transfer, the project team needed to develop a workflow that would integrate the tool, establish the appropriate procedures for use, and comply with PAIMAS, a standard for information transfer between an information producer and an archives. It "defines the methodology for the structure of actions that are required from the initial time of contact between the Producer and the Archive until the objects of information are received and validated by the Archive."[9] Elements of the WHS workflow, such as the transfer agreement and the information collected in Exactly metadata templates, ensure that the workflow meets PAIMAS requirements. This workflow is outlined in Figure 1.

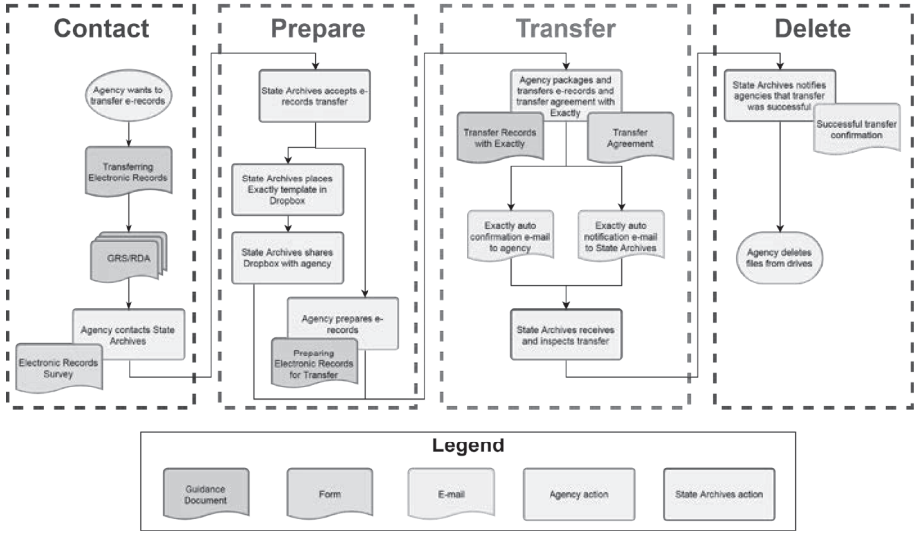

Figure 1: Preliminary workflow for electronic records transfer

The workflow is separated into four basic steps, which cover the electronic records transfer process from initial contact to deletion by the creating agency.

Step 1: Contact

Prior to transferring electronic records, state agencies must contact WHS staff via e-mail regarding the proposed transfer. The project team created a survey to collect information about content (i.e., applicable records schedules, records creators, date range), file formats, and access permissions (i.e., personally identifiable information, confidential records).[10]

The survey creates a clear line of communication between the state agency and the WHS about these records, including which records schedules cover them, who created them, and their size and complexity. It also establishes expectations about access provisions by collecting information about steps taken by the agency to identify confidential records and noting that, although WHS archivists will make efforts to screen for these records, it is primarily the responsibility of the transferring agency to complete this work. The survey provides WHS archivists a chance to offer feedback and guidance prior to transfer, and to refuse transfers of records that should not be sent to the WHS. Once WHS archivists have reviewed and approved the proposed transfer, the agency is notified that it has permission to move to the next step.

Step 2: Prepare

Upon transfer approval, agency staff consult the documentation created by the project team to prepare electronic records for transfer. This document includes guidance on creating records inventories, purging nonrecord items such as drafts, organizing records

and adding metadata, and managing records that have statutory restrictions or personally identifiable information.

Step 3: Transfer

After preparing electronic records for transfer, agency staff uses Exactly to package and transfer the records along with an electronic records transfer agreement. This agreement includes stipulations on transfer and access specific to electronic records.[11] Agency staff must also provide information about the transfer by completing the metadata template loaded into Exactly. The metadata template collects information about the records schedules, the arrangement of the records, the record types, and restrictions on access.[12] It is important to collect metadata at this step because the details of the transfer may have changed since the agency submitted an electronic records survey.

The completed transfer triggers e-mail notifications that are sent to both the sender and recipient. WHS archivists retrieve the transferred bag, move it to secured network storage, and validate the bag using Exactly. The transferred records are now ready to be "unpacked" from the bag and accessioned.

Step 4: Delete

After validation, WHS staff send a final confirmation e-mail to agency staff, informing them that the electronic records held by the state agency can now be deleted. The transferred records in WHS custody are the official records, per the transfer agreement, and the records held by the agency are now duplicates that should be deleted to eliminate any confusion about the official record holder. Once these records have been deleted, the transfer process is complete.

Although WHS staff notifies the agency when records can and should be deleted, the WHS does not have the ability to confirm or enforce the deletion of records by state agencies. This may be an area for future development, though it is unlikely that the workflow would ever include automatic deletion of records without the intervention or approval of the state agency.

Implementation

Following the development of the initial workflow, the project team started working with the Wisconsin Department of Workforce Development (DWD) on implementation. A major state cabinet agency, DWD collaborated with the project team by piloting electronic records management processes and identifying electronic records eligible for transfer to the WHS. Records staff at DWD tested Exactly and worked with DWD IT staff to gain approval and implement the software. About four months after the project team developed the preliminary workflow, DWD was able to successfully start transferring records to the WHS. However, during these four months, the WiSPER Project encountered several obstacles to implementation that delayed progress but also influenced and improved subsequent iterations of the workflow.

Many of the implementation challenges were a direct result of the project team's tool selection. While Exactly met all of WHS's criteria for a transfer application, state agency IT staff consistently resisted the use of open source tools, which they viewed as posing security risks due to lack of support. This resistance stood in stark contrast to the organizational culture of the WHS, where staff members frequently use open source tools for common library and archival tasks and have the ability to download applications to their own computers. The project team had several meetings with DWD IT staff to "make the case" for Exactly.

The primary concern voiced by DWD IT staff was the potential security risk to the data transferred using Exactly and Dropbox, particularly for transfers of confidential records. The project team began conducting research into alternatives to Dropbox which, like many cloud storage providers, has been subject to wider scrutiny for its data security practices.[13] To mitigate security risks related to Dropbox, the project team worked with IT staff at the WHS and the UW–Madison Department of Information Technology (DoIT) to set up an SFTP server to replace Dropbox in the transfer workflow. SFTP leverages an SSH connection to transfer files, making it a secure version of FTP. In addition to using a secure protocol, setting up the SFTP server meant that WHS IT and DoIT would directly administer the transfer storage, giving the WHS tighter control over access to data temporarily stored on the server. The SFTP server integrated with Exactly in a way similar to Dropbox, meaning that the workflow could still utilize Exactly's data validation capabilities. The updated transfer workflow, using SFTP instead of Dropbox, is illustrated in Figure 2.

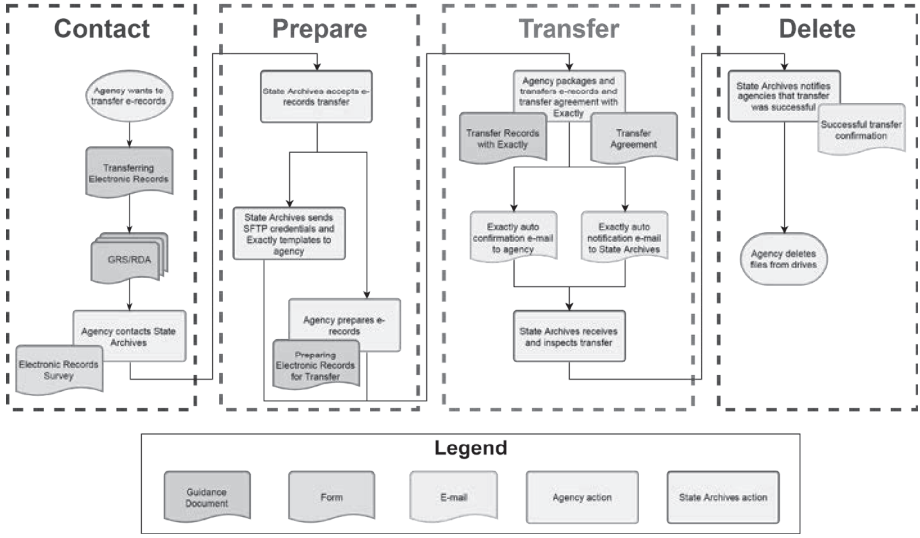

Figure 2: Updated workflow for electronic records transfers

Although they had not specifically requested the change, the switch to SFTP transfer meant that DWD IT staff were more willing to assess and approve Exactly for confidential data transfer. However, this technical improvement delayed the workflow implementation by about two months.

The second major challenge to implementation arose following tool approval during the installation and configuration of Exactly at DWD. Due to the WHS's dual role as state agency and North American history library for UW–Madison, the university's DoIT oversees IT services at the WHS, while almost all other state agencies in Wisconsin receive IT services through the Division of Enterprise Technology (DET). This means that, although the WHS is a state agency, WHS servers and applications are administered separately from those at other state agencies. This was an underlying obstacle throughout the entire WiSPER Project, which relied heavily on IT collaboration. During the implementation of Exactly at DWD, the project team had many complicated conversations and meetings with IT staff at the WHS, DoIT, DWD, and DET to troubleshoot errors and create firewall exceptions as needed.

The implementation process highlighted a discovery made during several parts of the WiSPER Project: partnering with DET would greatly smooth the process of implementing a software application across several state agencies. During the final year of the project, members of the project team began participating in an Electronic Records Working Group organized by the Department of Administration (the parent agency of DET), which focused on developing guidance for agency compliance with state electronic records management regulations. During these meetings, the project team was able to meet and make contacts with DET staff.

The project team began working with DET staff by proposing that DET add Exactly to its list of approved and supported software applications. An application does not need to be on this list to be utilized by a state agency, but DET will only offer technical support for listed applications. After several conversations, it emerged that DET would essentially never agree to offer support for an open source tool. The project team then suggested that DET officially approve the use of Exactly without supporting it, so that each agency's IT staff would not need to evaluate it. However, no model existed for obtaining this type of approval from DET, and no official approval materialized. Finally, the project team worked with DET to attempt to find an alternative enterprise-wide solution that would meet the requirements of the workflow. These efforts stalled during the testing phase, and it was clear that any solution offered by DET would not meet the project team's more robust requirements for metadata and transfer validation because it would not be built for the purpose of official transfers of electronic records. After these unsuccessful efforts, the project team pulled away from this attempted collaboration and finished the WiSPER Project using Exactly. Even without official DET approval, state agencies were still able to use Exactly—each agency would just need to evaluate and approve the software on a case-by-case basis.

When the three-year WiSPER Project concluded, DWD had successfully transferred 50.3 gigabytes of scheduled electronic records to the WHS using Exactly and all steps of the electronic records transfer workflow. In addition, the WHS had begun implementing components of the workflow at other state agencies. The Department of Transportation successfully utilized parts of the workflow, including several forms, to transfer records to the WHS, and the project team had begun working with their IT staff to approve Exactly. The project team rolled out Exactly at two other state agencies: the Public Service Commission and the Department of Revenue. Feedback indicated that the implementation of SFTP transfer was a vast improvement over data transfer on physical media, which presented its own security risks. However, until the WHS receives DET approval for Exactly or finds an alternative transfer application supported by DET, state agencies will continue to resist downloading and using this free and open source application.

Discussion

As previously discussed, many of the challenges faced during the workflow implementation phase stemmed from the project team's choice of an open source transfer application. While Exactly appeared to be near perfect for the workflow's requirements and the WHS was comfortable with using open source software, Exactly met with resistance when the project team tried to implement it with state agencies. A lack of clear communication channels between multiple agencies with very different organizational cultures compounded this obstacle. It is clear that, in the future, any project using a software application and technical infrastructure to transfer, transform, store, or manage data must involve DET from the outset. At the beginning of the WiSPER Project, WHS archives staff did not have established contacts with DET, which was not a partner on the project. During the course of the WiSPER Project, however, WHS archivists cultivated a more trusted relationship with DET, and they are hoping to leverage this in future projects.

WHS archivists will continue to introduce the workflow and Exactly at other state agencies. The endeavor is an iterative process, and the WHS will make improvements where possible. Although Exactly currently involves a significant setup process for state agencies and a slight learning curve, it has worked well for the purposes of electronic records transfer, and it has certainly been an improvement over past *ad hoc* transfers of physical media. Although much effort has focused on the technical aspects of the workflow, as a result of these developments, WHS archivists now have many non-technical processes in place that have greatly simplified electronic records processing. Furthermore, the WHS now has multiple guidance documents, forms, and dedicated space on network storage that will all aid in future transfers.[14]

As a result of this workflow development, the project team has also begun exploring the use of Exactly for other purposes in the WHS archives. In addition to the WiSPER Project, WHS is making a parallel effort to improve standards and procedures around preservation of all digital content. This work has included exploring options

for straightforward digital accessioning tools that archivists who do not specialize in digital archives could use. Because Exactly's interface is easy to use and it creates BagIt-compliant bags, the archives may be able to use it for accessioning digital records received on physical transfer media and for creating bags that can store digital accessions for the medium to long term. Another unanticipated benefit of the project is that the SFTP server set up by WHS IT and DoIT creates a secure "dropbox" for future transfers, regardless of the application or client being used to package and send the transfer.

Conclusion

The development and implementation of an electronic records transfer workflow helped the WiSPER Project team better understand the collaboration required to implement a software tool meant for multiple state agencies to use. The project team learned that, although a tool may work perfectly for the WHS and satisfy all of their business requirements, other state agencies can still create significant resistance. The project team also learned to make all workflows as technology-agnostic as possible, in case tools change. Almost every component of the WHS transfer workflow could also apply to transfers that do not use Exactly. The project team hopes that other similar institutions will find this discussion helpful in developing their own transfer workflows and that they can adapt pieces of this workflow to their own institutional contexts.

ABOUT THE AUTHORS

Hannah Wang is the electronic records and digital preservation archivist at the Wisconsin Historical Society, where she works with Wisconsin state agencies to manage, transfer, and preserve their electronic records. She also coordinates born-digital processing workflows and digital preservation planning for the archives at WHS. In addition to her work at WHS, she has taught classes in archives and records management at the University of Wisconsin–Madison iSchool. Wang has a master's degree in information science from the University of North Carolina–Chapel Hill.

Laura Poplett is a digital archivist at the Wisconsin Historical Society, where she has worked to implement a Preservica repository and public access website for Wisconsin state electronic records. She served as the electronic records project archivist for the WHS WiSPER Project and has a master's degree in information studies from the University of Wisconsin–Madison.

NOTES

The authors of this article would like to acknowledge and thank the NHPRC for its support of this three-year grant project. They would also like to thank all past and present members of the WiSPER Project team—Abbie Norderhaug, Jacob Riehl, Matt Blessing, Sarah Grimm, Carey Clifton-Myers, and Helmut Knies—for all their hard work and dedication through the many twists and turns of this project.

1. National Archives, "National Historical Publications & Records Commission: State Government Electronic Records Announcement," August 15, 2016, https://www.archives.gov/nhprc/announcement/electronic.html.
2. Library of Congress, *Bagger*, Java, 2019, https://github.com/LibraryOfCongress/bagger; John Kunze et al., "The BagIt File Packaging Format (V1.0)," https://tools.ietf.org/html/draft-kunze-bagit-17.
3. Library of Congress, " Digital Preservation: 2010 Digital Preservation Partners Meeting," July 21, 2010, http://digitalpreservation.gov/meetings/ndiipp10.html; Keone Yutaro Hiraide, "Kukini: The Challenges in the Design, Implementation, and Evaluation of a Digital Records Transfer Tool for the Hawaii State Digital Archives" (master's thesis, University of Hawaii at Manoa, 2014), http://hdl.handle.net/10125/100329; "Electronic Archives: Transferring Records," National Archives, August 15, 2016, https://www.archives.gov/era/training/transferring-records.html.
4. "Archives—Electronic Records," State Historical Society of North Dakota, http://history.nd.gov/archives/electronicrecords.html; Texas State Library and Archives Commission, "Transferring Electronic Records," Texas Digital Archive, https://tsl.access.preservica.com/tslac-digital-preservation-framework/p3985.
5. "Bagger GUI User Guide," State Archives of North Carolina, March 14, 2019, https://archives.ncdcr.gov/documents/bagger-gui-user-guide; Zachary Whitaker, personal communication with Hannah Wang, May 1, 2019.
6. "Exactly," AVP, https://www.weareavp.com/products/exactly.
7. Kunze et al., "The BagIt File Packaging Format (V1.0)."
8. Consultative Committee for Space Data Systems, "Producer-Archive Interface Methodology Abstract Standard," 2004, https://public.ccsds.org/Pubs/651x0m1.pdf; Consultative Committee for Space Data Systems, "Producer-Archive Interface Specification (PAIS)," 2014, https://public.ccsds.org/pubs/651x1b1.pdf.
9. Consultative Committee for Space Data Systems, "Producer-Archive Interface Methodology Abstract Standard."
10. "Electronic Records Survey," Wisconsin Historical Society, last modified October 13, 2017, https://www.wisconsinhistory.org/pdfs/la/ERecordsSurvey-Final.pdf.
11. "Electronic Records Transfer Agreement," Wisconsin Historical Society, last modified January 25, 2018, https://www.wisconsinhistory.org/pdfs/la/ERecordsTransferAgreement-Final.pdf.
12. "Electronic Records Transfer Information," Wisconsin Historical Society, last modified October 13, 2017, https://www.wisconsinhistory.org/pdfs/la/erecords-transfer-metadata-20171009.pdf.
13. C. Chu et al., "Security Concerns in Popular Cloud Storage Services," *IEEE Pervasive Computing* 12, no. 4 (2013): 50–57, https://doi.org/10.1109/MPRV.2013.72; Te-Shun Chou, "Security Threats on Cloud Computing Vulnerabilities," *International Journal of Computer Science & Information Technology* 5, no. 3 (2013): 79–88.
14. All workflows, forms, and guidance documents discussed here were produced as a part of an NHPRC Electronic Records Grant and can be found on the Wisconsin Historical Society website, https://www.wisconsinhistory.org/Records/Article/CS15454.