

Northwestern Pritzker School of Law

Pritzker Legal Research Center

Digital Preservation Plan: Dean Hansell Collection and Beyond

Developed by Miriam Heard

Created on 02-01-2017

Last Updated 03-03-2017

Version 02.01.01

Contents

Rationale for Digital Preservation	2
Statement of Organizational Commitment	4
Statement of Financial Commitment	5
Preservation Actions and Quality Control	7
Metadata Creation	9
Roles and Responsibilities	10
Training and Education	12
Monitoring and Review	13
References	14

Rationale for Digital Preservation

The Northwestern Pritzker School of Law community has identified a need within the Pritzker Legal Research Center (here forward PLRC) to extrapolate and preserve the donated collection of Dean Hansell. In the early 2000s, the Hon. Dean Hansell, JD alumnus of 1977, generously gifted the law school with his personal collection of English and European legal instruments. There are approximately 200 items ranging from indentures, deeds, contracts, trusts, wills, and other contractual legal documents dating back to as early as the 15th century. Most of the collection is from the 18th and 19th century, with some documents as current as the contract agreement for the Who to perform at Woodstock.

The designated community of PLRC and the university as a whole is diverse. The institution serves primarily the research faculty who actively publish and contribute to the larger law community. A subset is the students who are taught by the law professors, but at this point they do not actively utilize PLRC's archives. Secondly served is the network of other disciplines in the university that are dispersed between the Chicago, Illinois and Evanston, Illinois campuses. PLRC also strives to make the Hansell Collection available for public use through its eventual web presence.

The eclectic documents not only need to be physically identified and processed, but in order to allow for public and safe access, digitization is the next phase in the process of illuminating this one-of-a-kind collection.

The first step is to successfully digitize the Dean Hansell Collection. At the forefront of creating a successful preservation plan, the materials need to be properly identified and scanned into the server space.

Because the documents are handwritten, metadata standards will need to be as descriptive and complete as possible to ensure documents are discoverable. PLRC may eventually want to upload the collection again once more advanced text recognition software is developed. Actively staying abreast of emerging technologies will assist PLRC in discovering better solutions to its digital preservation efforts over time. It is also just as important in order to avoid any chances of software becoming obsolete before PLRC can safely migrate its collection to a more stable environment.

The structure of the institutions of PLRC and the university archives add another level of complexity to the digitization process. The librarian responsible for the Hansell Collection is new, and while this has been known in other institutions to provide opportunities to explore new methods of processing materials (Brillon & Craigle, 2013, p. 26), there is a desire to stay connected to the larger Northwestern University Libraries (NUL). The expectation is that this will be achieved through eliminating the siloed effect of the libraries. The establishment of temporary teams can be used to connect the libraries without exhausting resources (Thorp & Goldstein, 2010).

The realization that preservation is costly could result in a shortage of budget funding provided by the university to carry out many digitization goals. The librarian anticipates working closely with the director for PLRC to secure funds through outside groups.

The Dean Hansell Collection is just the beginning of digital preservation initiatives for PLRC. This project is being developed to establish a consistent workflow plan to apply to future projects.

Statement of Organizational Commitment

PLRC has created and filled the position of Special Collections and Archives Services Librarian to undertake the development and implementation of digitizing and maintaining the Dean Hansell Collection. Under the approval of the library director, and the guidance of the university archivist, the librarian will enlist the help of interns, and support staff to process the documents.

The librarian is responsible for consulting with the director and other departments to consider planning a future migration of the digital content to the university's server. The librarian needs to advocate for the collection locally to populate interest in the designated community. Collaborating with the university archives is also a large portion of the librarian's responsibility to establish and maintain an intercampus relationship. The librarian must stay active in researching and applying to grant funding to support future conservation efforts of the materials that are still in need of digitization.

The Special Collections and Archives Services Librarian position is a permanent fixture in PLRC. The law school expects whomever is in the position to continue with the preservation efforts of the previous employee as new projects arise. While the Dean Hansell Collection has a tight deadline to be web accessible by September 2017, the position will be utilized to process materials from other collections that have yet to be inventoried.

Statement of Financial Commitment

PLRC has several software programs in place to digitize the Dean Hansell Collection, as well as a limited supply of human resources. The scanning segment of digitization is to be executed in Opus, the software provided by the large book scanner purchased at the end of the 2014 fiscal year. After scanning, the person processing the material will navigate in Omeka, a free web subscription software package that serves as the digital repository for the collection. Eventually when the collection has been completely digitized, PLRC will migrate the metadata to ArchivesSpace, the web catalog for the entire university's archives. This service is maintained by the university archives, at no cost to the law school.

PLRC will need to find funding to support the conservation projects necessary to capture the images from tightly folded fragile documents. It plans to explore preservation avenues that offer free and low-cost evaluations before grant applications are in place to provide additional funding.

PLRC has set a goal to move the digitized documents from Omeka to a server provided by the university or law school. The information technology department of the law school as it currently stands is not equipped to handle such a task, so funding for consultation or temporary skilled labor is expected to be acquired through a few options. One such resource is potential donors. By creating a public space for what is available to be digitized, the library aims to attract alumni and other public facing stakeholders to realize the value of the collection and offer to support continued maintenance. Another option is to review fiscal spending and find a category where funds can be migrated to digitization projects. Lastly, grants that cater to digital preservation, archives, and small collections will be explored to continue to provide financial stability to PLRC.

Another future financial burden is the technology required to process and store the Dean Hansell Collection. The library is prepared to replace or repair the scanner and has allocated funds for equipment failure into its annual budget. Larger documents that do not fit on the scanner will need to be shipped offsite to a processing facility. The librarian is responsible for researching these costs and reporting to the director to determine the best course of action. As with the migration of the university's database for catalog records, there has been anticipation of another major migration in the future as better technology surpasses the functionality of ArchivesSpace. Since this software is jointly shared across all university libraries, it is expected that the funds will be pooled.

The librarian is currently limited to one intern and a support staff and together they contribute an estimated five hours a business week to processing the Hansell Collection. The librarian expects that unless there is a significant budget increase, they will need to rely on limited free labor to continue with digital preservation projects. The librarian will work with the director to access funding, and communicate with the financial manager for the operations

budget of the library to ensure they stay within budget.

Preservation Actions and Quality Control

PLRC has approximately seven months to digitize the Dean Hansell Collection before Alumni Week at Northwestern Pritzker School of Law. The librarian has ambitiously designed a series of checkpoints to ensure the collection is available for public accessibility on a website or blog.

The steps are as follows:

Digitization

- Accurately describe all documents
- Create proper housing for documents
- Develop testing environment to determine optimal scan settings
 - Design scanning standards document based on tests
 - Scans should be mindful of server storage usage; rasterize and flatten images if possible
 - Images for web should be .jpeg with a resolution between 75dpi and 150dpi
- Test load documents to Omeka
 - Design title/description standards for uploaded documents to ensure accessibility
- Scan documents in Opus
 - Look at preview to ensure entire document is captured and clear. A clear document is crisp, with distinct text and images that offer the best representation of the original document whether color grayscale or black and white as determined by the librarian.
 - Make adjustments as needed; rescan
 - Save a duplicate 300dpi TIFF for marketing and reproduction use
 - Use descriptions to generate metadata
 - Using a spreadsheet, document items scanned and date for reference and place holder
 - Double check that the document scanned matches metadata and series folder, box, etc.
 - Save files to secure server and create a duplicate folder on another drive in case there is a server crash
- Upload document images to Omeka digital repository
 - Double check that images are accessible to view online
 - Create appropriate title description using established standards
- Emulate print finding aid of physical items into ArchivesSpace to reflect Omeka repository item links
 - Generate accession ID

- Use EAD Best Practices at NUL Guide document to create cohesive shared archive metadata

Quality Control/Accessibility/Maintenance

- Run quarterly check sum of server data in Opus scan folder for the Dean Hansell Collection
- Once a quarter, physically browse Omeka web space to ensure users see scanned images
- Once a quarter run link resolver application to validate hyperlinks
- Stay up to date on software updates to ensure all browsers still have access to Omeka and ArchivesSpace
 - Run test upload from Opus to Omeka to ArchivesSpace for significant releases to test functionality changes
 - Update any and all standardization/best practices documents to reflect major updates if they affect usability
- Research new technology as it passes beta testing to consider preparing for migration
 - Keep in mind developments in text recognition software for scanned images

Metadata Creation

PLRC will adopt the procedures of the EAD Best Practices at NUL Guide to populate the metadata used in the Hansell Collection as it generates records in ArchivesSpace. ArchivesSpace will house the information about the collection that is provided by the physical finding aid. The bulk of the metadata that is significant to each item in the collection will be developed in the scanning software, Opus.

Opus allows for some customizable metadata, as one would find on the back of a photograph. The librarian will delegate the metadata design of Opus to a support staff. The end result will be a modified version of Dublin Core to capture the document title, prominent date, country of origin, language, and any authoritative figures. The Opus files that are uploaded to Omeka will be stored on the managed Omeka server as well as the server delegated to PLRC.

In concert with the Opus metadata, an Excel document will host the manual processing metadata associated with the digitization of the resources. This includes scan dates, record creator name, content box, and folder number of the physical items that are processed. The Excel file serves as a reference to when the material was digitized and where it came from to reduce the chance of redundancy and track progress. PLRC will store the Excel file in the same folder as the Opus scans on the physical law school server. An updated PDF and text file will be made to prevent inaccessibility if the Excel file gets destroyed or corrupted.

Roles and Responsibilities

PLRC will use the Hansell Collection as a processing guide to determine how to assess the length of time needed for future articles and collections to be digitized. At its current standing, the Special Collections and Archives Services Librarian will enlist the help of interns and support staff from the Access Services and Technical Services departments to execute the project.

The acting librarian will assume the majority of the responsibilities to assess and digitize a collection. They will come into the position with a working knowledge of archival processing best practices and will use NUL's guides to familiarize themselves with NUL's application of those guidelines in the PLRC setting. Figure 1 illustrates the time spent in a given month on one collection compared to the librarian's other responsibilities in PLRC. The librarian is responsible for reviewing and naming the collection and creating the accession ID. They are also responsible for consulting with NUL to determine if any materials are of more value in the main archives of the university. The librarian is responsible for recruiting interns and soliciting the assistance from the support staff by working with their department supervisors to schedule manageable time on the project.

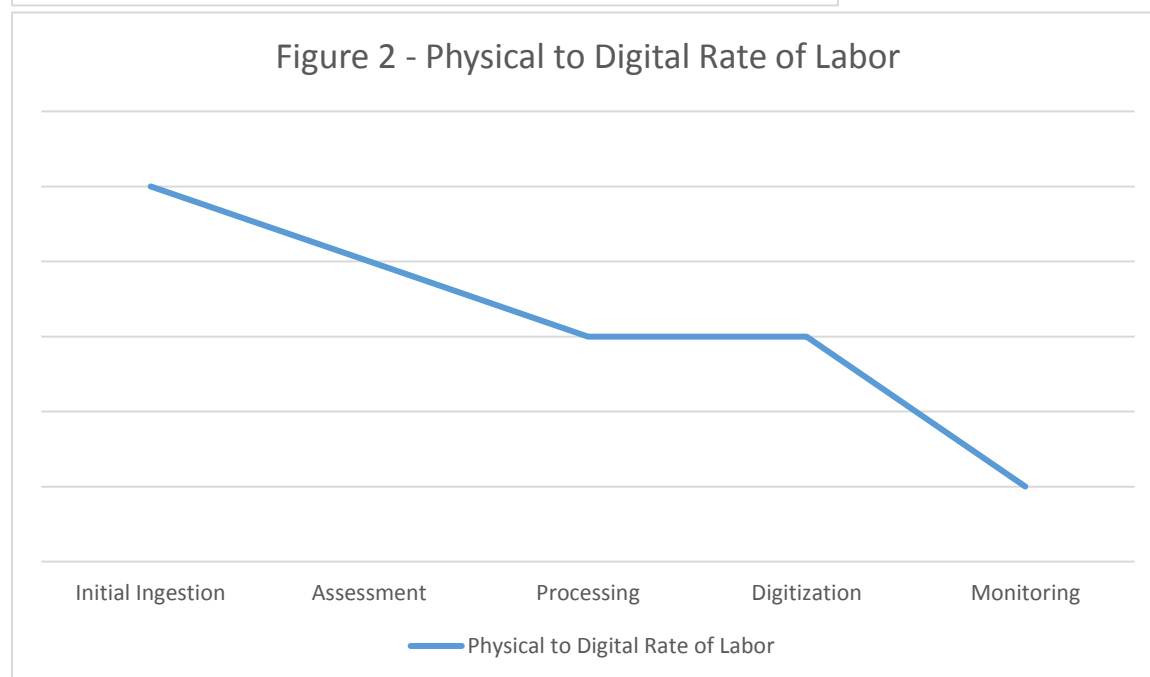
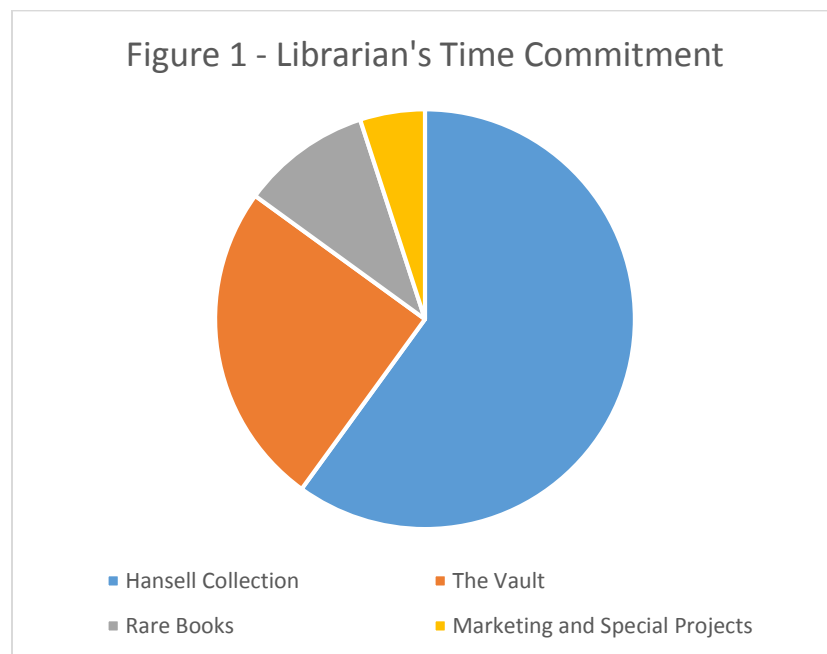
The librarian will arrange and provide training to the assistants in order to ensure a successful onboarding with the project. They will review the work of the assistants as quality control. The librarian will also work closely with the technology department to regulate server space for the digitized collection, and will collaborate with the web applications department to develop and implement new discovery interfaces or special interactive projects. The librarian is also responsible for forging a bond with the main archives as emerging technologies present collaborative instances for training and development.

The role of the intern is determined by the librarian based on their existing skills and working knowledge of collection evaluation and processing. The intern will most likely be used to inventory, document, and describe the physical items meant for digitization. They need to be detailed oriented and able to multitask since there may be more than one collection being processed at a time.

The intern will be a library student or recent library school graduate so that they have been exposed to courses in relation to the tasks they are given. The intern's weekly participation is decided based on their credit requirements and/or availability, but will not be a full-time work week appointment. The duration of the internship is also arranged, thus more than one intern is likely to process a single collection.

The support staff who will help the librarian are responsible for most of the processing work. They will supplement the intern's duties by housing the physical material, digitizing it, and following metadata standards to ensure the material is discoverable online. While their time is devoted to their department responsibilities, one can expect 20% of their week to include handling a project once it is prepped for digital preservation.

As a physical collection is processed and digitized, PLRC can predict a trend in human resource output. Figure 2 describes the lifecycle of processing a collection.



Training and Education

In order to maintain Northwestern's grand reputation of being an innovative and leading institution, it is imperative that the departments are aware of emerging technologies and best practices. The university libraries of Northwestern are proactive in their efforts to achieve this goal. Several examples are evident, but one to be mentioned is the recent migration from Archon to ArchivesSpace. Through the migration, the development of the EAD Best Practices at NUL Guide live document came to be that the libraries update it as new trends in documentation occur.

The NUL community has a strong workplace development budget to support staff in their efforts to grow within their department. These funds are primarily disbursed to librarians and technical assistants who attend conferences and off-site training sessions. The librarian and staff will have access to an allowance to learn about new trends in technology at tech shows, business meetings, and conferences. The librarian will have access to technology journals through accounts purchased by the main library so he/she will be able to read about related subjects.

Webinars and trainings across NUL--such as those provided during the onboarding process of ArchivesSpace--will serve as another resource for advancement in best practices in digital preservation strategies. Other free resources such as blogs and forums can supplement learning and development when conferences are inaccessible.

In most situations because funds are limited to each fiscal year, the librarian will take on the train the trainer role and report back to the involved staff if all cannot be in attendance of special off-site opportunities. The librarian is expected to stay active in relevant communities such as Society of American Archivists and other special interest groups.

Monitoring and Review

PLRC should exert caution when ingesting new accessions. One such reason is the sheer amount of neglect of the archives and special collections from before the employment of the librarian. Items left unlabeled in unmarked, non archival boxes is a perfect example of what PLRC should eliminate from the Hansell Collection and beyond. As the collection is digitally documented and preserved, PLRC will strive to maintain a systematic approach to reviewing and monitoring its digital resources and this plan.

The first step of approach is to revise the digital preservation plan as new technology emerges. Digital documents are assessed as described in the fourth section of this document: Preservation Actions and Quality Control. PLRC will assess if new developments meet the needs of the community and will enhance the collections.

PLRC will also revise the plan if there is a shift in community need that may or may not be impacted by emerging best practices and technology. First and foremost, the collection is public facing, and if there is a voiced majority opinion, this document shall reflect that in the best interest of PLRC stakeholders. Analytical data such as website tracking stats, discussion panels, and surveys will be used to monitor and assess how the collection is being accessed and how to make improvements through the documentation.

The plan should be reviewed in every instance that a new collection is ingested. As employee positions transform over time, it is crucial to stay on track with established guides in order to maintain consistency and stability throughout the collections.

The acting librarian will review the plan as changes develop in the EAD Best Practices at NUL Guide. They will consult with the head archivist to determine if retraining should be offered to support staff working on digitizing material, and if so, the plan document will be updated to reflect those changes or restructuring.

Lastly, the librarian should review the preservation plan at the end of every fiscal year whether or not a new collection is ingested or new technologies are adopted. This will ensure a fresh look at the document before a new year to catch any new developments or make changes that may have been missed throughout earlier.

References

- Brillon, A. & Craigle, V. (2013, September). Automated library services: One law library's experience. *AALL Spectrum*, 18(1), 26-27. Retrieved from <http://eds.b.ebscohost.com.offcampus.lib.washington.edu/ehost/pdfviewer/pdfviewer?sid=daf22c32-a492-4c65-b6ab-2e2d97c251d5%40sessionmgr120&vid=13&hid=104>
- Thorp, H. & Goldstein, B. (2010, September 3). How to create a problem-solving institution (and avoid organizational silos). *Chronicle of Higher Education*, 57(2), A43-A44. Retrieved from <http://eds.b.ebscohost.com.offcampus.lib.washington.edu/ehost/command/detail?sid=daf22c32-a492-4c65-b6ab-2e2d97c251d5%40sessionmgr120&vid=22&hid=104&bdata=JnNpdGU9ZW9vc3QtbGl2ZQ%3d%3d#AN=53564326&db=a9h>