

OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

Building the Perfect Repository for Archival Collections: Lessons Learned from the Henry A. Kissinger Papers Project

Kevin L. Glick, Manuscripts & Archives, Yale University Library, kevin.glick@yale.edu

Rebecca Hirsch, Manuscripts & Archives, Yale University Library, rebecca.hirsch@yale.edu

Steelsen Smith, Library Info. Technology, Yale University Library, steelsen.smith@yale.edu

Session Type (select one)

- ✓ Panel (length of panel, but actually a presentation on a single topic)

Abstract

A vision for the perfect repository necessarily incorporates rights management and system integration, but, more importantly, is based upon the needs of researchers. Archival collections require access to rich descriptive content and easy browsing of a hierarchical arrangements and archival bonds of collections that cannot be adequately represented with systems designed for monographs, serials, or stand-alone still images. This presentation demonstrates how the Kissinger Papers project at Yale was used as an opportunity to conceive of, and develop a repository tailored to these needs that would be flexible enough for all types of archival collections. The presentation also addresses a unique way to handle rights management that allows for the digitization of an entire collection while still maintaining granular control over researcher access and requesting workflows.

Conference Themes

- ✓ Integrating with External Systems
- ✓ Managing Rights
- ✓ Building the Perfect Repository

Keywords

archival materials, rights management, repository systems integration

Audience

The audience for this session comprises all staff involved in the design and development of digital repositories built to include the description of, and content from, archival collections - both digitized and born-digital. This includes archivists, librarians, repository designers, and developers; disparate groups that have tended to not work closely in the past.

Background

This session will address three of the major themes highlighted in the Call for Proposals for Open Repositories 2015: “Managing Rights,” “Integrating with External Systems,” and “Building the Perfect Repository.” The presenters will discuss these themes specifically in the context of archival collections,

OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

but the insights and the some of the solutions developed, especially regarding rights management and integration with user management systems, are applicable to many different digital repositories regardless of the types of digital content.

Additionally, this session fits into the conference's theme of Looking Back/Moving Forward, as our system development process started with existing digital library repositories and archival descriptive standards like EAD. We then built off of this foundation, leveraging the work done by others in addition to our own. This system has been envisioned as the a first step toward connecting together a number of related systems used to describe, display, track, manage, monitor, and control access to archival materials online. It represents the next step in providing access to large, complex collections of primary source material.

Presentation content

Archival collections are inherently different from the materials that are typically managed by libraries, such as books and serials and they are described by a different set of cataloging standards. Most existing digital repositories were developed to deal with single-item works, whether they are monographs or single images of art objects or historical documents. Archival records differ from other types of objects because much of their inherent value lies in the interrelationships between each record and other records resulting from the same activity, referred to as the archival bond. The archival bond places a record in context and gives additional meaning to the record. It includes the relationships between different records that relate to a specific transaction, as well as the relationship between the records of preceding and subsequent transactions. The archival bond is so important that archivists consider most individual records to be of little or no value with removed from this context. Thus, the archival context must be presented to the end user so that they may fully understand the meaning and value of the archives.

In contrast to the traditional book-centric digital library model, this presentation will address the needs of institutions that have digitized archival collections that they want to display in whole or in part in the context of their archival finding aid, the standard form of metadata used by archivists to describe the context and archival bond of archival collections. Most current digital library access systems in the United State divorce digital materials, including archives, from their contextual description. This method de-contextualizes digitized archival materials, making it difficult for a researcher to reconstruct the relationships between the item they are viewing and the components of the collection of which it is a part, especially when archival collections are returned as part of large search results.

Another issue is that much of the literature about the digitization of library materials advises organizations to first focus on digitizing (and therefore providing digital access to) those collections that are in the public domain, or for which the institution owns the copyright, and have no donor restrictions or privacy issues. This has the result that most of the archival collections that have been digitized in their entirety are older 19th century collections. However, as the world becomes increasingly digital, both donors and users have come to expect (or at least hope for) the digitization of 20th century (or even 21st century) archival collections. Indeed, the digitization of these collections is essential to enabling the

OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

growing field of digital humanities to do its work. Nonetheless, collections may not have uniform access policies - privacy, copyright, and donor restrictions may apply differentially to different parts of a collection. Applying granular control to researcher access allows the entire collection to be made available digitally while ensuring that only its public components are available without auditable staff approval.

This presentation will involve a theoretical and practical discussion of what the ideal access system for archival collections looks like and how it functions, with a focus on two of the unique elements of the mass digitization of archival collections: the need to provide access while maintaining and presenting invaluable context of the archival bond and archival description and also the need to exercise granular control over access rights. The ideal access system for archival collections should also be fully integrated with other critical systems, including archival collections management (Archivists' Toolkit, ArchivesSpace, etc.), archival search, discovery, and description systems (EAD finding aids), logistical systems, user management systems, and AV or born digital records access systems. We will not only discuss what we view as the ideal system, but also demonstrate an access system in beta-testing at Yale University Library that addresses many of the raised issues, including providing context and rights management.

The proposed presentation will comprise three presenters from Yale University Library: Kevin Glick, Head of Systems and University Archives, Manuscripts & Archives; Rebecca Hirsch, Kissinger Project Archivist, Manuscripts & Archives; and Steelsen Smith, Systems Specialist, Library IT. Note that this presentation is envisioned as longer than a single paper presentation, as long as a panel, although this presentation is not a true panel of disparate peers.

Mr. Glick will talk about the integration of an ideal primary front end system for archives with other systems, including: ArchivesSpace/EAD finding aids; request management system; user tracking system; access system for AV and born digital records.

Some of this integration work has already occurred at Yale University Library for the access system to the Kissinger papers, while more is under active development. Mr. Glick's background as the Head of Systems for Manuscripts & Archives makes him well-suited to discussing the integration of these various services into one interconnected technology environment that provides both the front-end accessibility to the researcher as well as also providing staff with the ability to manage both users and digital objects, and physical original materials in an efficient manner.

Ms. Hirsch will discuss how the method in which archives are described in finding aids affects the manner in which they can be presented online and how they must be described in search results. In order for researchers to correctly interpret the meaning and relevancy of searches, the archival context of those results, both collection and component, must be clearly described from the same page. Traditional digital repositories designed for monographs and other single-item results generally do not provide enough context for archival collections, ignoring the large aggregations of material described in multiple parent-

OR2015 | 10th International Conference on Open Repositories

June 8-11, 2015, Indianapolis, Indiana, USA

child relationships. Instead, the ideal access system will make it easy to navigate those parent-child relationships without having to perform additional repetitive searches. Ms. Hirsch will demonstrate the ways in which the access system that has been designed for the Kissinger Papers project presents the context and archival bond of the collection. She will also compare the Kissinger papers access system as built with the wireframes used for development and screenshots of access systems built by other institutions that were considered while coming up with the requirements for the Kissinger system.

Mr. Glick will introduce the topic of restrictions with a conceptual discussion the role of restrictions in the management of archival collection management. Then, Mr. Smith will discuss the rights management system that has been developed for the Kissinger papers project. The project contains materials with three levels of access: open, open with permission and restricted. In order to provide researchers with the greatest possible access while still complying with restrictions imposed by the donor, a rights management system was built that allows users to ask for and be granted access to collection materials at different levels - whole collection, series, sub-series, heading or folders. This system is integrated with both the front-end access applications and the library's user and request management platform. Mr. Smith will talk about the development of this rights management system in technical detail, and demonstrate how it extends existing staff tools without altering the underlying operation of that software. The technical description will not delve into application code or a level of detail inaccessible to an archivist or librarian. This portion of the presentation will include an active demonstration of how the rights management process works for the Kissinger Papers project.

Conclusion

A vision for the perfect repository necessarily incorporates rights management and system integration, but, more importantly, is based upon the needs of researchers. Archival collections require access to rich descriptive content and easy browsing of a hierarchical arrangements and archival bonds of collections that cannot be adequately represented with systems designed for monographs. This presentation demonstrates how the Kissinger Papers project at Yale was used as an opportunity to conceive of, and develop a repository tailored to these needs that would be flexible enough for all types of archival collections. The presentation also addresses a unique way to handle rights management that allows for the digitization of an entire collection while still maintaining granular control over researcher access and requesting workflows. Finally, the presentation explains how this repository has been, and will continue to be integrated with other systems, allowing it to enter the library ecosystem without introducing extra manual workflows or altering the behavior of critical software already in use.