Leading the Way Toward Low-Cost Digital Collections:
The S.D. Cabaniss Digitization Project

Amanda Axley Presnell
University of Alabama Libraries

ABSTRACT

Beginning in January 2010, the University of Alabama Digital Services department embarked upon a 14-month project to digitize the papers of Septimus D. Cabaniss, a Huntsville-area attorney during the Civil War era. The project is funded through a grant awarded by the National Historical Publications and Records Commission. In addition to providing digital access to a collection of national historical significance, the University of Alabama is digitizing this collection in a revolutionary way that will provide more immediate access to materials at a much lower cost than other methods. This model will allow other institutions to mass-digitize large collections by taking advantage of detailed series descriptions already available in a collection finding aid rather than creating item-level metadata for each object. Digitized material is presented within the context of the finding aid, allowing users to locate items in a way that is similar to the physical encounter with the collection if the individual was to visit the University of Alabama's W.S. Hoole Special Collections Library.

CONTACT

Amanda Presnell
University of Alabama Libraries
Email: aapresnell@ua.edu
Phone: (205) 348-0515
Website:
http://www.lib.ua.edu/libraries/
hoole/cabaniss

PROJECT PURPOSE AND SIGNIFICANCE

The purposes of this project are to scan and make available online the complete papers of S.D. Cabaniss (approx. 40,000 images) in a cost effective manner, demonstrating and testing a simple model whose efficiency and usability we intend to prove.

This model will enable other institutions to digitize massive quantities of materials at a minimal cost, while providing effective search and retrieval for scholars within the context of the collection.

The Cabaniss collection was chosen for this project due to its detailed account of a southern attorney's work to award African-American slave children the estate of their white father. Cabaniss' largest client was Samuel Townsend, a man who openly admitted to fathering slave children and sought to take responsibility for their care by leaving them his estate. Included in the collection are Townsend's will, letters from the slaves themselves dated before and after the Civil War, Cabaniss' detailing of the sophisticated trust used to free the slaves, and Cabaniss' correspondence with important political leaders, including William Seward and Stephen Douglas.

METHODS

The previously-created online finding aid serves as the primary point of entry to the online material (see Figure 1). Links to digitized content are added to the finding aid weekly, providing access to new content within weeks of scanning. This methodology provides users with the same context, provenance, and order as encountered by a researcher perusing the physical materials in the reading room. Additionally, this method's efficient online access to digitized materials has a remarkably low cost.

Students have been employed for flat-bed scanning under the supervision of a Digitization Specialist, who also performs overhead scanning, material preparation, access development and usability testing.

Usability testing will be performed at the conclusion of the project to provide feedback and demonstrate the widespread applicability of this model for other institutions.
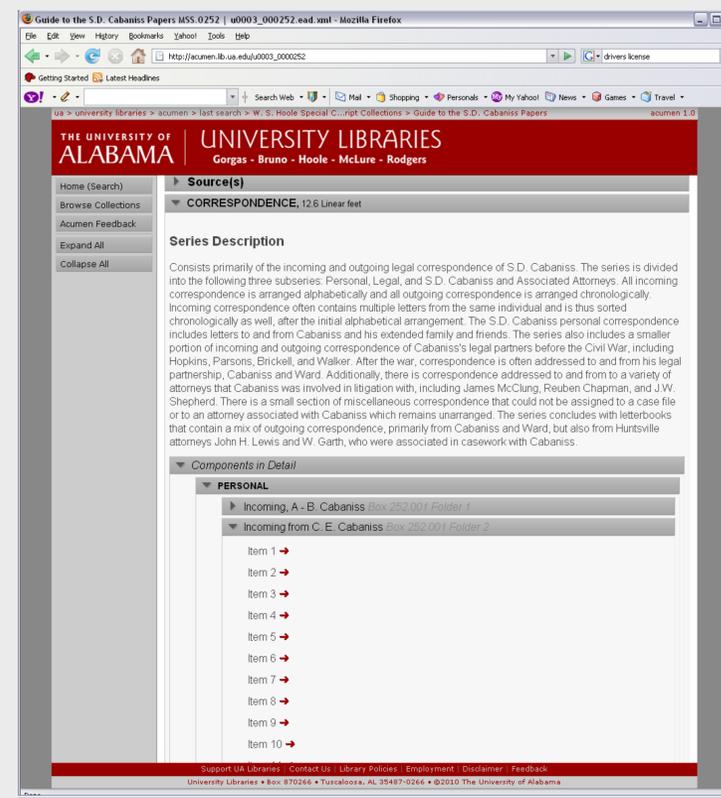

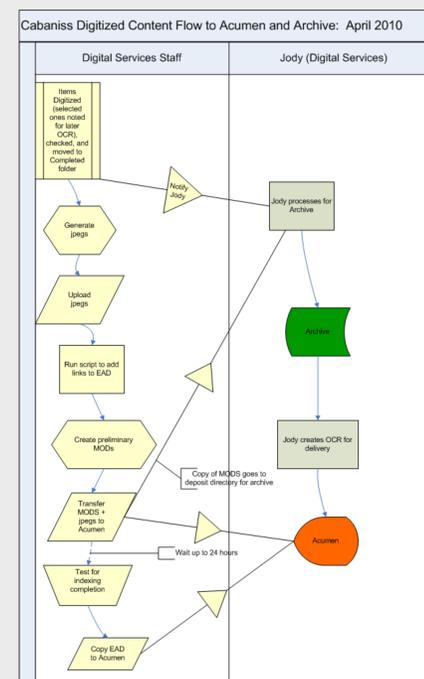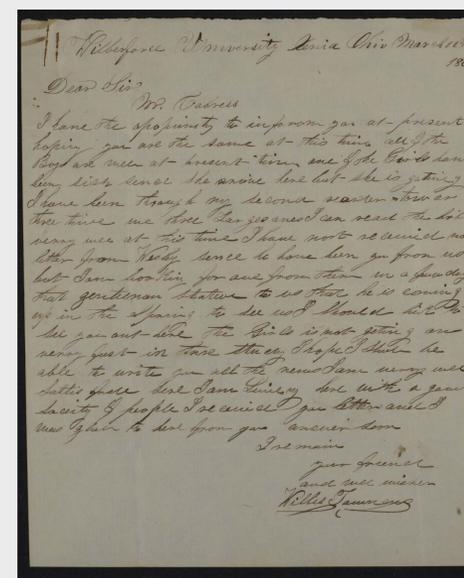*Figure 1 – Digitized items are presented online within the context of the finding aid.*

Figure 2 - Workflow of the digitization process.

An example of an item in the Cabaniss collection – a letter from Willis Townsend, son of Samuel Townsend, to S.D. Cabaniss. Willis writes to Cabaniss about his studies at Wilberforce University in Ohio.

METADATA

In the absence of descriptive metadata for each item, we are presenting scanned items within the context of the finding aid. We are using scripts to insert links to the digitized items which are organized in directory folders similarly to the way they would be encountered in the physical collection. We wished to link out of the finding aid rather than present the scans themselves within the finding aid, as that would be too cluttered and confusing.

Because the traditionally-used software Archivist Toolkit did not meet our needs for inserting links into the finding aid, one of the university's metadata librarians created a unique set of Encoded Archival Description tags, based on those used by the Library of Congress.

File names and directories use a taxonomical structure that allows the following information to be communicated within the filenames and directories themselves: collection number, box number, folder number, item number (sequence encountered in the physical collection), number of pages and item identifier.

DIGITIZATION & WORKFLOW

Physical items are prepared for scanning by the Digitization Specialist by pulling a box from the special collections stacks and perusing the items to determine if special attention is needed prior to scanning.

Scans are saved to a departmental shared drive and are organized within directories for each box and folder. Workers scan each of the items, place scans in the proper directory, and record their work on a spreadsheet created for this purpose. At this point, items are spot checked for image quality; correct naming and sequence is checked using Perl scripts.

Content is delivered via the Acumen website (the university's locally-created delivery system for digital collections), using open source software which was developed in-house for this purpose. This software uses Windows and Linux Perl scripts to organize digital items in openly accessible web directories.

Workflow is outlined in *Figure 2*.