

# **Digitization in the Real World**

**Lessons Learned from  
Small and Medium-Sized  
Digitization Projects**

Edited by

Kwong Bor Ng & Jason Kucsma



Metropolitan New York Library Council

Published in the United States of America by  
Metropolitan New York Library Council  
57 East 11th Street, 4th floor  
New York, NY 10003-4605  
p: (212) 228-2320 f: (212) 228-2598  
Web site: <http://www.metro.org>

ISBN: 978-0-615-379998-2

Cover Design: Jason Kucsma (*illustration by Smartone Design,  
licensed via iStockphoto.com*)

Reviewers Committee: Mark F. Anderson, Jill Annitto, Anna Craft, Jody DeRidder, Renate Evers, Wei Fang, Maureen M. Knapp, Sue Kunda, Mandy Mastrovita, Ken Middleton, Emily Pfothenauer, Mark Phillipson, Alice Platt, Mary Z. Rose, Stacy Schiff, Jennifer Weintraub, Andrew Weiss.

Copyright © 2010 by Metropolitan New York Library Council. No part of this book may be reproduced in any form or by any means, electronic or mechanical, including photocopying, without written permission from the publisher.

**The views expressed in this book are those of the authors, but not necessarily those of the publisher.**

# Digitizing Colorado State University's Historic Photograph Collection: A Case Study

Beth Oehlerts (Colorado State University Libraries)

## **Abstract**

This chapter discusses the experiences of creating a medium-sized digital collection of the earliest photographs from Colorado State University's Historic Photograph Collection. The entire collection of 500,000 photographs chronicles the history of Colorado State University, the city of Fort Collins, and Rocky Mountain National Park. Digitizing the first phase of this vast collection included capturing 5500 images from glass plate negatives, 750 images from magic lantern slides, and 7500 images from gelatin nitrate prints. These formats were chosen because they are the oldest and most fragile. The project was not without challenges, including utilizing untrained staff and students, coping with equipment problems, and creating avenues of communication to more than 40 people involved in the creation of the collection. This chapter will discuss these challenges and how we worked to resolve them.

**Keywords:** Archives, Digital collaboration, Digital collections, Digital projects, Digital scanners, Gelatin nitrate prints, Glass plate negatives, Magic lantern slides, Photograph collections.

## Introduction

Beginning in the late 1990s, Colorado State University Libraries (CSUL) started large-scale digitization activities in partnership with the Colorado Digitization Project (CDP). The first of these was the *Warren and Genevieve Garst Photographic Collection* of wild animal photographs donated to CSUL and digitized with CDP funding (<http://lib.colostate.edu/wildlife/>). *The Sidney Heitman Germans from Russia Collection* was another early digitization project (<http://lib.colostate.edu/gfr/index.html>) funded by the CDP. These projects, and a series of smaller ad hoc projects over the next three years, were created following the CDP's best practices guidelines and gave a few staff the opportunity to learn digitization.

In 2006 CSUL received the University Historic Photograph Collection (UHPC) a collection of 500,000 photographs documenting the history of Colorado State University, Fort Collins, and Rocky Mountain National Park. The collection came from the University's Office of Instructional Services and was given to CSUL's Archives and Special Collections Department with the understanding that it would be preserved and access given to a wider audience through digitization. Planning for the digitization of the first phase began in early 2007, as did creation of a finding aid for the earliest images in the collection, which included glass plate negatives, magic lantern slides, and gelatin nitrate prints from cellulose nitrate negatives. These photographs are also the most fragile materials in the collection. The goal was to digitize approximately 5,500 glass plate negatives, 750 lantern slides, and 7500 gelatin prints.

Although we had no budget line dedicated solely for digitization, our work did have the support of both the Colorado State University (CSU) and CSUL's administrative teams. Funds were provided by the Libraries' administrative team to purchase digital scanners, including an archival-quality overhead scanner used for later archival digitization projects. Administrative support provided extra funding in the project when we needed to hire students and acquire additional server space to house our digital master files.

## Literature Review

Of the recent literature addressing digitization, the focus is often on the technical and preservation-related issues that arise following the creation of a digital collection; less of the recent literature focuses on the staffing issues encountered when creating digital collections. Boock and Vondracek (2006) found that when institutions begin digitization activities they often add these tasks to staff responsibilities rather than hire new staff dedicated to digitization. In a 2005 survey of ARL libraries, they found a majority of the responding institutions capitalized on the existing knowledge and skills of staff, encouraging current positions to evolve rather than hiring new digitization staff, as was the case at CSUL. Boock continues the discussion, focusing on how the Oregon State University Libraries (OSUL) reassigned several staff positions to a new digitization production unit, initially using staff for scanning, quality control reviews, and metadata creation using a metadata schema (Boock, 2008). OSUL relies on student employees for almost all of its digital imaging and metadata assignment, performing no quality control reviews on their work. In examining the organizational implications of digitizing, Sutton (2004) recalls how early digitization efforts were seen as temporary endeavors, requiring the temporary re-assignment of staff.

D'Andrea and Martin (2001), reporting on digitization workflows at Temple University, discuss utilizing part-time student staff to supplement the work of digitization staff. D'Andrea and Martin note, and our experiences confirm, there are positives and negatives in hiring students to work on digitization projects. The authors recommend hiring students whose interests match the project, not just those who apply for the job, and fully informing students about the nature of the work.

## Establishing Best Practices

In 2005, in response to the growing need for local standards and procedures to guide the creation of CSUL's digital collections, CSUL created the Digitization of Local Collections Task Force and charged it

to review existing national and international standards and best practices, and to recommend the steps necessary in building high-quality local digital collections. Among its findings, as reported in the final report, *Digitization of Local Collections* (Digitization of Local Collections Task Force, 2005), the Task Force identified a lack of formal project management, a need to prioritize projects, and the need to determine the impact in-house digitization would have on staffing. Based on the Task Force's final report and the Western States Digital Imaging Best Practices, version 1.0 (Western States Digital Standards Group, 2003), CSUL then created a series of local best practices covering all aspects of the digitization process, starting with guidelines on how to handle and care for materials to be scanned. Our policy is to capture an image that is a true representation of the original. To do this all the objects in the UHPC digitization project were digitized as archival objects, with no alteration made to the master image.

With some of the glass plate negatives, lantern slides, and gelatin nitrate prints more than 100 years old, including glass plates that are cracked or chipped, we scan an object once to reduce the chance of damaging fragile materials by repeated handling. Master images are created as TIFF (Tagged Image File Format) files. The glass plates were scanned at 1200 dpi, the magic lantern slides at 2400 dpi, and the gelatin nitrate prints at 1200 dpi. The master images were deposited in a project folder on a central server prior to the application of metadata and ingest into our digital repository.

Our best practices documentation, including national and international digitization standards documentation resides on the CSUL staff wiki where it can be accessed by all employees. The wiki includes a wide range of information related to creating CSUL digital collections, including how to operate the digital scanners by format and instructions for creating project-tracking Excel spreadsheets that record the work of each staff member. The information gathered on these spreadsheets enables the supervisor to monitor and evaluate the work of the scanning staff, identify inefficiencies, and keep the project on schedule. Workflow analysis helped support our proposals to hire additional student scanning staff, purchase additional scanning

equipment, and identify changes in workflow that will improve future digitization projects.

## **Project Planning**

UHPC digitization project planning began in early 2007, with discussions and project meetings between the project archivist, the digitization librarian, and the metadata librarian. The project's participants knew what would be digitized, what metadata was needed, and where the collection would reside, but we had to develop a road map to help us achieve these goals.

That road map began to take shape with the UHPC finding aid, created by the UHPC project archivist. The digital projects librarian created the UHPC digitization project-tracking documentation and identified staff for scanning, created staff scanning schedules, and arranged for staff access to the Archives Annex where the materials were housed. The metadata librarian, working with the project archivist, developed a plan to use the information in the finding aid for the metadata. Then the digital services librarian developed the procedures for loading them into our content management system.

In September 2007, the UHPC digitization project scanning started with a small group of test scans to determine the best procedures for creating the digital images. Using CSUL's standards-based digitization guidelines, staff created approximately one dozen images. These scans were reviewed for accuracy and quality by both the project archivist and the scanning supervisor; when the pilot images met the requirements for the project, scanning began.

## **Utilizing Untrained Staff**

Initially the number of trained digital scanning staff at CSUL was relatively small, with a Library Technician I and a Library Technician II, who worked on earlier digitization projects and had experience with scanning photographs and text documents, and a Library Technician II, with limited scanning experience. A Library Technician III supervised the work. These four staff also had other job assignments: the Library Technician I assisted in preparing materials

for the bindery and assisted with a large inventory project; and the Library Technician IIs performed database maintenance tasks and also assisted with the inventory project. The Library Technician III, who reports directly to the digital projects librarian, supervises CSUL's in-house digital scanning, the database maintenance unit and the inventory project. She is responsible for creating and adjusting scanning staff schedules, monitoring scanning progress, contacting Library Technology Services (LTS) staff to maintain and repair scanning equipment, and performing quality control reviews of the scanned images.

The project began with 56 hours of scanning per week. Unable to hire additional staff, we examined staff duties to identify tasks that could be streamlined or eliminated altogether. A decrease in the number of print journal subscriptions and a reduction in the number of titles sent to a commercial bindery meant those saved hours could be applied to digitization. The same was true with tasks associated with database maintenance. For example, when approached by the CSUL Government Documents Unit to transfer the government documents database maintenance tasks to the Government Documents unit, we agreed and made the change. There were also some basic database maintenance tasks that could be performed by students, saving more time for digital activities.

We identified one of our Library Technician IIs whose knowledge of and experience with scanning, combined with her skills as an instructor, made her an excellent teacher. With these skills she trained our staff to create the high quality digital images we needed and now trains all our scanning staff and students.

Experience has shown us that the level of training needed for creating digital images varies by individual and must be tailored to the individual. Each person brings various levels of experience and understanding to digital scanning. Aside from the level of experience, the unique nature of the materials, coupled with unfamiliar models of equipment, requires individualized training. Training begins with scanner instruction and viewing Cornell University Library's *Moving Theory Into Practice Digital Imaging Tutorial* (Cornell University

Library, 2003.) Our digital imaging trainer spends 2-3 weeks teaching the trainee the local, national, and international standards and best practices, how to operate all of the digital scanning equipment, how to handle the materials being digitized, and any necessary project-specific scanning requirements. It usually takes staff, with multiple job responsibilities, one year to acquire the skills to create high quality archival digital images.

In December 2007, a couple of months into the project, the scanning was not progressing as quickly as expected. We hesitated to consider using student employees due to the fragile condition of the glass plate negatives and magic lantern slides, but there was no other option; the project needed to keep on schedule. We trained our most experienced and mature student employee first and eventually hired more students to assist with scanning. The first three hourly student scanners spent a combined average of 36 hours per week scanning. As summer approached we asked and received permission to offer summer employment to two of the students. The ability to keep them working during the summer, at a total of approximately 60 hours per week, helped us keep the project on schedule.

Midway through the project, in the spring of 2008, four members of the copy cataloging staff joined the UHPC digitization project. They were trained to assist the metadata librarian with metadata creation, a natural progression of their duties. The copy catalogers spent approximately 20 hours per week throughout the project assisting with metadata.

The UHPC project archivist, our authoritative resource for the project, was also our daily contact for the project. Creating images at the Annex proved to be a benefit for the scanning team because the project archivist's office is located at the Annex. She was available to answer questions about the extremely fragile materials and instruct staff on how they should be handled. At the start of digitization, the project archivist was responsible for performing quality control reviews of the images, but we soon discovered this arrangement was not feasible. The project archivist was still in the midst of processing the collection's 500,000 images, and the logical solution was to have

the scanning supervisor review the scans. To reduce the number of possible errors, scanning staff perform a review of their own scans prior to the end of their work day. Staff are expected to correct any errors found before a review by the supervisor. If the supervisor discovers images that need to be rescanned those are given back to the person who created the image. Correcting scanning errors became a learning opportunity and the person responsible for the scanning error was responsible for rescanning the image. If the problems were a result of malfunctioning scanning equipment, quality control reviews were performed on 100% of the images until the problems were corrected.

Additional changes in job responsibilities occurred in the spring of 2008 when we replaced most staff working on the large inventory project with student employees. The change gave us approximately 32 additional hours per week to devote to the UHPC digitization project. To accommodate the increase in hours we met with the archivists and the Archives and Special Collections Coordinator to request an extension of the hours at the Annex. The Archives staff was understandably reluctant to have others in the Annex when an archivist is not present, but they agreed to adjust the schedules of staff in the Annex and to extend the hours. The increase in the rate of production made meeting the project deadline a more realistic goal.

By July 2008 all of the glass plate negatives and magic lantern slides had been scanned and once the quality control reviews and all rescans were complete we notified the metadata librarian. She began the process of preparing the metadata for the copy catalogers. Upon completion, the metadata librarian reviewed their work and requested corrections where needed. The metadata librarian then notified the digital services librarian, who began the process of ingesting the images and metadata into the digital repository.

During the life of the UHPC digitization project we lost and gained staff and student employees. The impact on the project was challenging at times. In the late spring and again in the fall of 2008, two Library Technician I staff undertook several weeks of intensive training to acquire the skills to fill void of created by the loss of two

employees. The new staff were unfamiliar with digital scanners, they had little or no experience handling fragile photographic materials. Fortunately they quickly became proficient and by the end of 2008 we added 24 hours a week back to our scanning schedule.

The final format to be digitized was the gelatin nitrate prints. Work on digitizing this format began in August 2008 when we hired eight work study students to perform the gelatin nitrate print scanning. We began the fall semester with 132 hours per week: 80 hours of student scanning time, 32 hours of staff scanning, training, and quality control review, and 20 hours of staff metadata creation. We hoped the students would be able to work largely on their own after training. The scanning supervisor worked closely with them for the first couple of weeks; she then let them work on their own. A short while later, while performing quality control reviews of their work she noticed the students were making a large number of errors, many of them basic scanning errors. Our experiment to reduce the close supervision of the students was not a success. Distractions, or perhaps the lure of the holidays, resulted in a lack of concentration.

Scanning the UHPC collection took place at a site across campus. Saving the images to the Libraries' server from this remote site, and searching the web at the same time, caused the computers and scanners to slow and occasionally created serious technical problems. We now block internet access at all our scanning stations and more closely supervise the students' work.

There are advantages and disadvantages with using students to work on digitization projects. The advantages include lower labor costs, a more flexible workforce, and the opportunity to give students valuable skills and good work habits. One disadvantage is the substantial investment in teaching students the skills and knowledge needed to create high-quality digital images only to have them graduate after a year or two of employment. Due to the repetitive nature of the work, it is difficult for some students to concentrate on scanning images. Many times students are not able or willing to work during school breaks. In spite of the disadvantages, we found that the positives of employing students outweigh the negatives.

It is difficult to firmly estimate the number of hours of scanning time you will need when planning a digitization project. To compensate for staff leave and the uncertainty of student employment during breaks in the academic year or during final examinations, it is a good idea to budget extra time for digitization. Project downtime may occur because of equipment problems. In our case, while the scanners were being repaired we had to bring the scanning to a halt and reassign staff and students to other projects at CSU's Morgan Library. In addition to the scheduling problems on the UHPC digitization project, staff scanning schedules had to include 15 minutes to walk across campus to the Archives Annex, where the scanning was done. By using our project-tracking documentation we were able to assess the progress of the project at various points and make any adjustments needed to keep us on schedule. The willingness of everyone involved to adapt and remain flexible was a major factor in the success of the project.

## **Overcoming Equipment Problems**

The project began with just two Microtek Artixscan 1800f flatbed scanners. The Microtek scanners were purchased prior to the project startup and were chosen for their design, which includes a drawer where the glass plate rests emulsion-side up without touching glass or the light source. This separation is desirable when scanning fragile glass plate negatives. The scanners were installed and calibrated by the LTS staff, the equipment experts. During the first four months of scanning we encountered numerous equipment problems, including having to recalibrate the scanners every day, sometimes several times a day to eliminate colored lines on the digitized images.

Equipment problems plagued us through much of this project and adding staff to scanning team did not help the situation. With more people scanning the scanners were in operation for 50 or more hours per week. This high rate of production took a toll on all the scanning equipment and the scanners were showing the stress of overuse. We needed more scanning equipment for both technical reasons and to increase production. The digital projects librarian recommended staff

operate two scanners at a time, as one way to increase production. With 20 minutes to scan one glass plate, staff had time to prepare one image for scanning while the other was being scanned. We decided to follow this recommendation and requested that CSUL purchase two Epson Expression 10000 XL scanners, a model that best suits our needs. The scanners arrived in January 2008 but then we had to wait. It took LTS a couple of weeks to install the scanners. At the same time one of our Microtek scanners malfunctioned. Attempts to repair the scanner were unsuccessful and the only option was to return it to the manufacturer for refurbishment. The equipment installation delays and failures caused us to suspend the UHPC scanning for two weeks. We had to reassign staff and students to other projects in the interim.

By February 2008 the four scanners were in place and working, but then we were experiencing other problems. Our new Epson scanners were producing Newton's Rings, a series of concentric, alternating light and dark rings centered at the point of contact between the glass surface of the scanner and the glass plate negative, on the scanned images (Illueca, Vazquez, Hernández, & Viqueira, 2002). The older scanners were beginning to produce banding artifacts (straight lines) on the images. As this only happened with the glass plate negatives, we switched from scanning glass plate negatives to scanning the magic lantern slides until we could determine a solution to the problems. A resolution to the Newton's rings problem came from the digital projects librarian, who thought an acid-free paperboard (barrier board) frame, similar to mat boards used in picture framing, might resolve the problem. The frame raises the glass plate from the scanner glass just enough to allow air to circulate between the plates of glass and eliminates the Newton's rings. During the digitization of the University of North Carolina at Chapel Hill's William Blake Archive, scanners encountered Newton's Rings when creating images from transparencies. Their solution was to scan the images directly and not through glass (Viscomi, 2002). This solution was not practical for us. Using a frame to separate the glass plate negatives, which are extremely fragile, from the scanner glass provided a safe method to protect the emulsion side of the plates. The

banding artifacts problem was resolved when we began a daily cleaning and re-calibrating of the scanners.

There was also the challenge of creating quality images from over-exposed and under-exposed glass plate negatives. We adjusted the scanner to capture either a darker or lighter image to correct the poor exposure. One point of pride for us was that we did not break or damage any of the fragile glass plate negatives in the collection.

## **Improving Interdepartmental Communication**

Prior to the UHPC digitization project most digitization work was small ad hoc projects that did not require planning meetings, project-tracking documentation, detailed condition assessment guidelines, or extensive quality control reviews. As a result, not everyone participating in creating digital projects was aware of a project's status. Our digitization procedures changed with digitizing the UHPC collection. During the digitization project not all of us worked in the same location. The project archivist worked in one building, the metadata librarian, the digital projects librarian, and the digital services librarian worked in another. We needed to develop a project management structure that could foster communication and promote collaboration. Project communication, which was taking place via email, often resulted in further 'follow-up' emails. We realized that everyone involved in the project needed to know what was happening and why.

In early 2008 the Repositories Matrix Team, CSUL's administrative and policy-making team for digital activities, created the Digital Projects Management Plan Working Group (Working Group). The Working Group's charge is to monitor the progress of all digitization projects. It includes representatives from Archives and Special Collections, Metadata and Preservation Services, and Digital Repositories Services; every department participating directly in digitization. The Working Group began meeting weekly in February 2008. The initial focus was on the issues and problems surrounding the UHPC project. Now members report on digitization project news, digital equipment problems and equipment purchases, staffing issues,

and conclude with a 'round robin' of project status reports. All policy questions are forwarded to the CSUL Repository Matrix Team. Building on the Working Group model, we now schedule pre-project planning meetings for all new digitization projects. These meetings focus on the scope of the project, staffing needs, equipment needs, and the creation of a digitization timeline. Attendees include faculty and staff who will be involved in the creation of the digital collection.

The collaboration that formed during the UHPC project contributed much toward improving interdepartmental communication. The project required a close working relationship along with frequent meetings and emails between the project archivist, the digital projects librarian, the metadata librarian, and the digital imaging staff. Group members were willing to share their knowledge and skill with others and remain flexible while workflows shifted and changed course during the life of the project. Those involved with building digital collections at CSUL have learned that building quality digital collections requires planning, good communication, and a commitment to a collaborative endeavor.

### **Summary/Lessons Learned**

Digitization of the UHPC collection provided several lessons. Within months of commencing the project, with enough data to create an accurate picture of our progress, we realized that we needed more staff, more equipment, and more time to complete the project. While hiring additional digitization staff was not possible, our option was to train staff in other units to assist with portions of the work. As mentioned above, the willingness of everyone involved to remain flexible was a major factor in the success of the project.

When budgeting time for a digitization project, experience taught us that 20% more time should be added to the timeline, to allow for problems of all types to be resolved. Developing a realistic project budget and project timeframe will also help administrators and others understand the true costs of building a valuable digital collection. A short pilot project, conducted prior to the start of the project, can reveal quite a lot. It will show you how much time it really takes to

scan the materials; which scanner settings should be used for a particular format; and whether it is possible to make textual documents searchable. Although we did conduct a short pilot, it was designed to learn how to operate the scanning equipment. We should have gathered other information, such as how long it takes to scan a glass plate negative, because soon after the scanning began we realized that we had wildly underestimated the time we thought was needed to scan these negatives. The 5 minutes scheduled per scan in fact turned into 20 minutes. A more comprehensive pilot would have revealed, among other things, what was technically feasible.

The UHPC digitization project underscored for us the importance of regularly sharing information with the project's participants. Good communication, as mentioned above, was vital to understanding the UHPC project's goals and unique digitization requirements. With more than 40 people involved in its creation, sharing information helped us understand how to handle the fragile archival materials, avoid many image rescans, understand the reasons for many of our equipment problems, create solutions, and keep on track to meet the project's deadlines.

## **Conclusion**

By June 2009 the photographs in the first phase of the project were digitized, and the images and metadata began to be loaded into the digital repository. Each digitization project is unique, and each has the potential to present new challenges. Though it might not be possible to anticipate every outcome, there are some things that can be done to prepare the project for success. Obtain the support of your administration or funding agency for the project; develop good project planning skills and implement them; create tools for communicating with all project participants and document all the processes and workflows of the project. Subsequent digitization projects from this vast collection will benefit from what we learned during the first phase of the UHPC digitization.

## References

- Boock, M. (2008). Organizing for digitization at Oregon State University: A case study and comparison with ARL libraries. *Journal of Academic Librarianship*, 34, 445-451.
- Boock, M., & Vondracek, R. (2006). Organizing for digitization: A survey. *Libraries and the Academy*, 6, 197-217.
- Cornell University Library. (2003) *Moving theory into practice digital imaging tutorial*. Retrieved on March 31, 2010 from <http://www.library.cornell.edu/preservation/tutorial>
- D'Andrea, P. & Martin, K. (2001). Careful considerations: Planning and managing digitization projects. *Collection Management*, 26(3), 15-28.
- Digitization of Local Collections Task Force. (2005). *Digitization of local collections*. Retrieved on March 31, 2010 from <http://digitool.library.colostate.edu/>
- Illueca, C., Vasquez, C., Hernández, C., & Viqueira, V. (2002). The use of Newton's rings for characterizing ophthalmic lenses. *Ophthalmic and Physiological Optics*, 18, 361-362.
- Sutton, S. (2004). Navigating the point of no return: Organizational implications of digitization in special collections. *Libraries and the Academy*, 4, 233-243.
- Viscomi, J. (2002). Digital facsimile: Reading the William Blake Archive. *Computers and the Humanities*, 36, 27-48.
- Western States Digital Standards Group. (2003). *Western States digital imaging best practices. Version 1.0*. Retrieved on March 31, 2010 from [http://www.bcr.org/dps/cdp/best/wsdibp\\_v1.pdf](http://www.bcr.org/dps/cdp/best/wsdibp_v1.pdf)