

LDI Round 5 Grant

Fine Arts Library, Harvard College Library

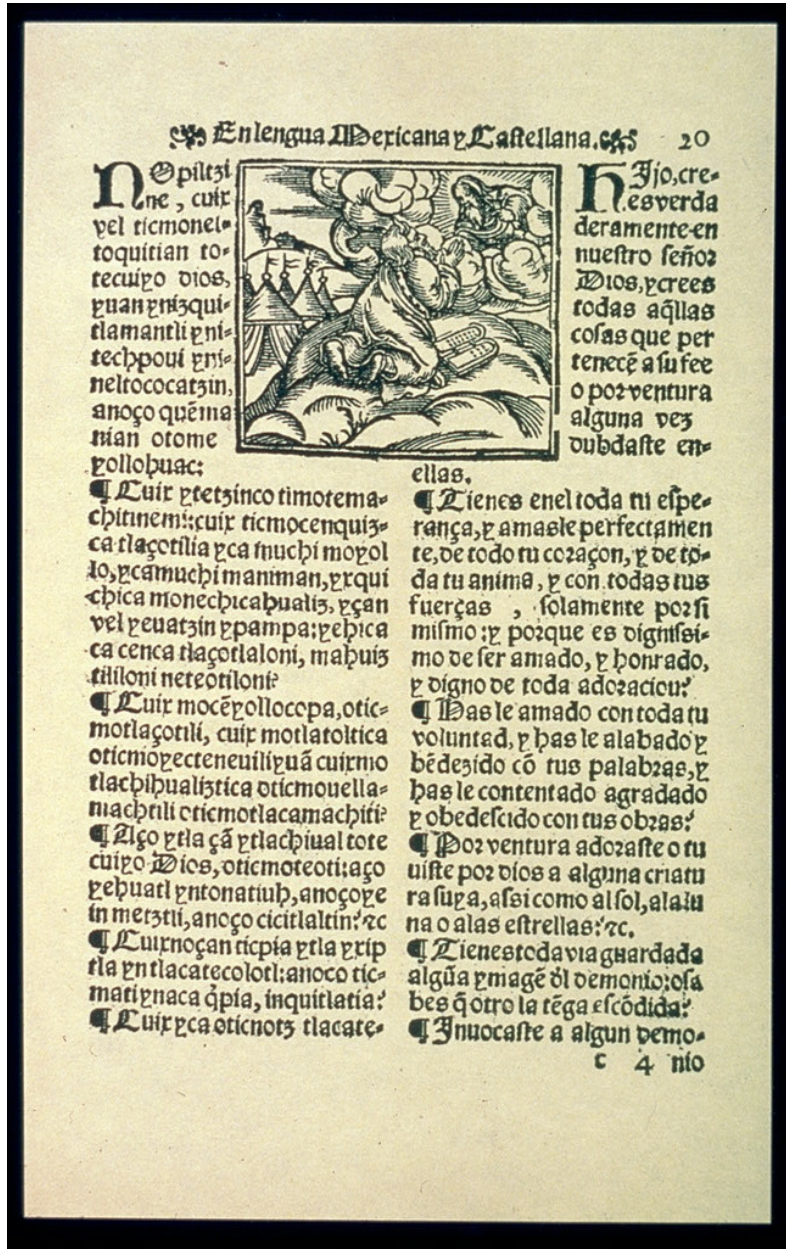
Digitization of the Slide Library

Final Project Report



Table of Contents:

1. Introduction
2. Methodology and Standards
3. Budget
4. Conclusions



Introduction

In May 2002, the Fine Arts Library submitted a proposal to the LDI grant review committee for a project described as follows:

The Fine Arts Library of the Harvard College Library is interested in participating in the LDI Grant Program, round 5, to support the digitization of approximately 100,000 35mm core teaching slides. The selection of which slides are included will be based on the definition of core course needs of History of Art and Architecture faculty, as well as some core course needs of FAS Classics faculty, and the faculties of Visual and Environmental Studies and the Graduate School of Design.

The FAL slide library forms a remarkable resource for teaching and research, documenting the entire range of works of art and architecture from prehistoric times to the present. Indeed, the use of visual images for teaching is essential in the fields of the fine arts. The images in the slide collection are important as a whole, especially given this collection's breadth and depth, even though the slides do not hold any intrinsic value in and of themselves. The faculty and students of the Department of the History of Art and Architecture use the collection most heavily, for study and classroom display. Other users include Visual and Environmental Studies, Graduate School of Design, and FAS Classics faculty and students. This study collection is open to the larger Harvard community as well.

As part of ongoing operations of the FAL Slide and Digital Images Department we are currently scanning new acquisitions as well as scanning slides for which there are records in VIA but no images.

This project was conceived of as a partner to the LDI/ACCESS-funded retrospective conversion ("recon") project. The digitization project endeavored to encompass 100,000 slides for which the data was converted and accessible in OLIVIA. Our original timeline anticipated a start time for the scanning project 11 months after the start of the recon project, and an overlap period between the projects of close to two years. In the end, however, we had to extend the timeline for the recon project.

The scanning project kicked off officially in December 2003, and ran through January 31, 2006. This required us to support the Project Coordinator for more time than had originally been budgeted for. As the scanning grant did not include sufficient funding to support the Project Coordinator, we were able to use some of the additional data recon funding to do so. The Fine Arts Library requested, and received, \$320,409 for the LDI 5 project. An additional \$64,236 was transferred from the Access project to support the Project Coordinator.

Methodology and Standards

We determined that outsourcing the scanning would be the more efficient method, rather than doing it in-house. The costs to the Fine Arts Library for in-house scanning would have been high. In order to scan 100,000 slides in two years the Library would have had to support additional staff, as well as invest in expensive equipment, such as scanners and servers. This level of investment would be far over and above what the Library would need to support even an expanded on-going scanning operation. Moreover it is unlikely that an in-house project staff could have maintained the ambitious schedule we followed.

At the same time, we believed that some aspects of pre- and post-scanning processing should be done in house. This included creating batch data in OLIVIA to accompany the scans in order to facilitate linking the images with the appropriate data. By using previously created and tested software for project tracking and depositing images into the DRS, we would have better control over the processes and the quality. We had a proven workflow to accommodate the entire process thus we believed that keeping this portion of the project under local control would ensure the best results.

All of the images created during this project were deposited into the DRS, and linked in OLIVIA to the appropriate cataloging record. For each image, five files were created. These included an archival tiff file, two projectable-sized files (one medium and one large, the latter of which was linked in OLIVIA and viewable in VIA), and finally a small file which meets the requirements for use with the ICG Course Website tool. When an image is requested by ICG the Fine Arts Library will ensure that it is moved from the DRS by whatever means work best with the website tool.

We used a workflow largely informed by ongoing slide scanning operations at the FAL Digital Imaging Lab. We are processed around 1,500 slides per week. As part of the pre-scanning process, the in-house Scanning Assistants batched the images in OLIVIA and the pertinent batch data accompanied the slides to the scanning vendor and back. We sent the physical slides and received back the slides and image files. Due to the heavy use of the collection we felt it was important to minimize the length of time slides were out of the library. Our contract negotiations with the chosen vendor also addressed this issue. The vendor sent the images to us on DVDs.

The image files we received included the original "raw" scan which was machine color-corrected at the time of scanning, as well as a sharpened version of this image which served as the master image from which derivatives were created. From the vendor-created master image we required three derivative images, for use with various image delivery systems (e.g., VIA, the ICG website tool).

Image specifications adhered to the standards which were established when the FAL Digital Imaging Lab was created and are being used for ongoing operations. See Appendix 1, for specifications.

Following a careful review of responses received to our Request for Proposal, we chose JJT, Inc. as the vendor for this project. This decision was based on both cost and vendor responsiveness. The contract negotiations took far longer than we had expected (this was something we also experienced during contract negotiations with the vendor for the LDI Access project), but were concluded in summer of 2004. However, it still took many frustrating weeks to settle issues with JJT and finally get into production. These issues included quality problems with their scans, schedule problems, and general misunderstandings on the part of JJT about exactly what it was we wished them to do. These problems delayed the start of production by 4-6 weeks. Once in production, the stream of 1,500 slides per week to the vendor, and 1,500 image sets returned was established as planned.

After a period of relative calm and productivity, we once again hit snags with our vendor, JJT, Inc. in the summer of 2005. Upon reviewing images returned from JJT, the project assistants noticed prominent "ghosting" within the images. This was particularly noticeable with high-contrast, abstract images. This problem was brought to the attention of JJT, who admitted that this was a known problem, and that the set-up they were using was known not to be optimal for slide work for that very reason. Setting aside the question of why this was never brought up when they were bidding on the job, we spent some amount of time developing an alternate plan.

The vendor decided to send the slides from Plymouth, MA to JJT's facility in Baltimore, MD, which apparently is better equipped. JJT took on the additional costs of transporting the materials to and from Baltimore. To the extent that it was possible for the staff at the Fine Arts Library, we attempted to send 3,000 slides every two weeks, as opposed to 1,500 every week. At the same time, JJT re-scanned about 10 batches (15,000 slides) to eliminate the ghosting on the images.

The Department of the History of Art and Architecture received three Provost's grants that summer (2005) to digitize and catalog slides used by faculty for teaching. These are slides that reside in faculty members' personal collections, but which they feel should be available through VIA. We scanned slides for two of those faculty members through this project, totaling about 10,000 slides. We felt that this was well within the parameters of the grant, since they are images

that will be part of the library's (digital) collection, and which will be widely accessible on VIA and will support teaching and research at Harvard.

Given the last episode with JJT, it seemed likely that we will need to extend the project by about two months (through February, 2006 instead of the end of December 2005) in order to get to our goal of 100,000 images. We felt confident, however, that we would be able to reach that goal within this time period, barring any further issues with the vendor. However, in November we reported the following:

There is both good news and bad news to report for this period. First the bad: we have had continuing issues with our slide scanning vendor, JJT, Inc. Problems involved quality of scans, turnaround time, and communication. After battling with them and making continuous compromises (e.g. a turnaround time of 8 days instead of one week, which eventually climbed to 10 days) without positive returns, the decision was made to send them a ten-day warning that unless changes were made immediately the contract would be terminated. This letter was sent to JJT on October 28. We received a reply on November 4 indicating that JJT would, indeed, attempt once again to comply with the terms of the contract, including a 7 day turnaround. While we remain somewhat skeptical, we hope that they will be able to live up to these promises. At this point any failure on their part to meet our conditions will result in immediate termination of the contract.

The good news was that we were three-quarters of the way through the scanning. At that point we still thought that if the vendor did, indeed, make good on their promises we had a good chance of making up the lost time and completing the project at the end of January.

Due to Fine Arts Library move preparations the space used by LDI5 project staff had to be re-purposed. The decision was made that the end date of LDI5 would be no later than January 31, 2006, regardless of the project status at that time.

Progress was made with the vendor, but unfortunately not enough to make up for the time lost, given that the project end date of January 31, 2006 was pre-determined. In the end, 83,245 images were scanned, deposited into the DRS along with their derivative images, and linked to their cataloging records in OLIVIA. These images are all available through VIA at this time. The failure to achieve the goal of 100,000 images scanned can be laid squarely at the feet of the vendor, who has, since the completion of the project, gone bankrupt. Had we been able to extend the project for another few months we most likely would have gotten all 100,000 images done, however that was not possible. On the other hand, this project, as did the Access project, concluded with unspent funds.

Budget

As noted above, the scanning project was awarded matching grants of \$160,204.50 from Harvard College Library and the Library Digital Initiative program. An additional \$64,236 was transferred from the Access project to support the Project Coordinator. Total funds for this project were \$384,645. Of this over \$23,000 remains unspent, which will be dispersed to the two funding agents in accordance with the original grant funding procedures. Had we been able to extend the project these funds would have sufficed to complete the scanning and post-production work.

Conclusions

Two important lessons were learned from this project:

- 1) Contract negotiations take far longer than expected. It would not be unreasonable to plan for 4-5 months in a project schedule for the process. The contract must be reviewed by the university's Office of the General Counsel, and one must schedule time with them in advance. In fact, part of the initial project planning should involve conversations with the OGC if an outside vendor is to be involved;
- 2) Working with vendors can pose real challenges. The Scanning project and its companion Data Digitization project (Access grant-funded) exposed us to two sides of the spectrum. The data vendor was quite easy to work with, responsive, accurate and responsible. The

scanning vendor, on the other hand, proved quite problematic. Specifications and standards that we believed to have been clearly stated throughout the negotiations, and within the final contract, were continually questioned. The vendor kept trying to re-negotiate the contract, whenever they found they were unable to do that which they had contracted to do. Cancellation of the contract was threatened at one point, due to the vendor's lack of communication coupled with their apparent inability to do the work as required. While there was no way to have known at the beginning that the vendor would not work out as expected, the result was that far more time was spent just dealing with the vendor, and as noted above, the project goal was not realized.

On the other hand, the digitization of over 80,000 slides is a major accomplishment, one not paralleled by any of the Fine Arts Library's peers. The addition of these images to VIA increases the value of that resource as a teaching and research tool for the study of the history of art and architecture.



APPENDIX 1: Image Specifications

As mentioned above image specifications and workflows used in this project adhere to the standards which were established when the FAL Digital Imaging Lab was created. These are the same standards as those upon which current slide scanning in the FAL DIL is based.

Also noted earlier, the nature of this collection is such that any individual image or slide holds little intrinsic value. These images generally represent photographic representations of original art works. The majority of the images are photographs of photographs found in exhibition catalogs and monographs (so called “copy-stand” photography). The vast majority of images in this collection are easily replaced. For these reasons we feel that there is no need to generate or store an archival image of any greater size or quality than is indicated here. With the large number of slide scans the Library ultimately will be storing, we feel that Level 2 Preservation Service (as defined within the DRS Policy guide, accessed 04 March 2002 at <http://hul.harvard.edu/ois/systems/drs/policyguide.html#preservation>) is most appropriate.

The standards for image specifications currently in use include creating large and medium sized scans appropriate for projection within a classroom or lecture hall. While these images are not currently accessible from VIA, they are stored in the DRS. As technological advancements occur within Harvard’s image delivery systems this situation may change.

A total of five files will be created for each image, including the file for use with the ICG website tool. All five files will be deposited at the same time into the DRS. The table below describes these files, and their purposes:

| Description | Type | File Name | Pixels on longest side | Use |
|--------------------------------|-------|-----------|------------------------|---|
| Raw image ¹ | .tiff | Prdmaster | 2126 | Archival master |
| Sharpened | .jpg | Prdwork | 2126 | Projectable image |
| Large deliverable | .jpg | Lgdl | 1024 | Projectable image, Link in OLIVIA for VIA |
| Medium deliverable | .jpg | Mddl | 800 | |
| Small deliverable ² | .jpg | Smdl | 800 | ICG Course Website Tool |

All of the files will be scanned at 1,500 DPI, unless there is a difference in price for scanning at 1,200 DPI which is equally as accepted.

Currently we are using Linocolor (version 6.x) for sharpening and Photoshop (version 6.x) for color matching. Photoshop could be used for both processes, and if the sharpening is done in-house we will most likely use Photoshop. The software being used to create the derivative images (which are all created from the sharpened, corrected Prdwork image) is DeBabelizer (Equilibrium). We will use this product for this project as well, if we are creating the derivatives in-house.

¹ Black point/white point auto-assigned by the scanner.

² This is the same as the medium deliverable, but at medium compression.

