
Session 14: Case Studies in Digitizing: Image Databases. Course reserves, and Finding Aids

Monday, April 4, 2005, 4:00-5:30 PM

**Moderator:** Maryly Snow, UC Berkeley  
**Speakers:** Leslie Abrams, UC San Diego; Heather Cleary, Otis College of Art and Design; Ed Loera, Portland State University.

**Recorder:** Corey Schultz, Stanford University

**Sponsor:** Panel created and organized by the Visual Resources Division

**Summary:**

*Maryly Snow, UC Berkeley*

Part of this session was delivered at the 2005 VRA Conference, where the speakers included: Kristin Solias (UMASS Boston), Vickie O’Riordan (UC San Diego), and Ed Loera (Portland State University).

Creating digital projects rely on collaborations of different extents. This panel is not necessarily a session on “how to” do a digital project or strategic planning or costs or funding. UC Berkeley’s SPIRO (Slide and Photograph Image Retrieval Online) has 72,000 digital images available on the Web without password protection. This project was started in 1985 at UC Berkeley by the Architecture Visual Resources Library (AVRL) and Museum Informatics. A database was developed in 1987 and AVRL began to use digital images by 1988. In 1993, they got a faster scanner, and in 1995 began making course websites. In the past fifteen years, they have created 72,000 images; five thousand per year, scanned and archived and presented on the Web in two formats, of which 25,000 are rights-free images.

Lessons learned from this project:

1. The VR community is developing standards, but one must remember that each collection, and its institution, is unique; therefore, there may be diversity in solutions.
2. Use the “request, request, request” mode, especially if you do not have powerful allies.
3. Streamline the process and report and log every step needed to track images (and therefore track errors). Modify the systems as needed annually.
4. Develop methods of quality control: while filing slides and filing shelf list cards check cataloging and arrangement errors, and do a quality control check for the digital content by reviewing images and records together on the Web.
5. Remember that technology keeps changing; be firm in your goals but flexible in your methodology.
6. Participating in image aggregation projects requires a Digital Asset Management System (DAMS).
7. Need campus-wide planning initiative for digital records.
Loera managed the project Walker County Treasures (www.walkercountytreasures.com), which included a range of community resources from Walker County, including the Huntsville Public Library, Huntsville Art Commission, Peabody Memorial Library and Archives, Walker County Genealogical Society, Samuel Walker House Museum, and the Walker County Courthouse. They received a “Texas Treasures Grant” of $20,000 to create a repository for these resources, with a goal of digitizing and cataloging two thousand objects. The project team involved John Stokes as Image Consultant, Ed Loera as Cataloging Consultant, two student image workers, and two student catalogers.

Loera created a XML template using modified Dublin Core. Equipment included a repurposed Dell PC, an Epson Expression 1680 Scanner, PhotoShop (with additional plug-ins), and a repurposed Linux server (with a 80 GB hard drive). They debated between choosing commercial software over that of open source software. Though both have a learning curve, they felt that the open source would be better for this project as it was free, had an active user support group, and offered formal workshops. Loera chose the open source software “Greenstone” due to its recommendation on www.oss4lib.org. It is robust, dynamic, creates thumbnails, and is internationally known and supported. It also has a users’ support group, which allows them to get assistance, scripts, etc. from other users on the listserv.

They began with their historic photo collection and entered as much info as they had available. They chose the “finding aid” method and separated the collections because it allowed for faster response in searching (a more robust server was needed to effectively combine collections for integrated searching). When data is added to Greenstone, it reclassifies all items; therefore, updates had to be done at night when no one needed to access the system.

They exceeded their goals and put over five thousand images in their Web-accessible database. Future plans include adding additional images as well as oral histories, upgrading to a more robust server, and offering integrated searching between collections.

Leslie Abrams, Head, Art Library, UC San Diego
“The Good ($$$) the Bad (Surprises!) and the Beautiful (Digital Images & Metadata): UCSD’s Slide Digitization Grant Experience.”

Digitization project with ARTstor included the following: University Librarian, Brian Schottlaender, was the Principal Investigator; Leslie Abrams, Head of the Art Library, oversaw the budget and administration of the grant; and Vickie O’Riordan, VR Curator, was the Project Manager. The digitization project was funded by a grant by the Mellon Foundation with the goal of digitizing 200,000 slides and their data (80 percent of their collection) for inclusion into ARTstor over the course of eighteen months. Scanning was done by an outside vendor, 125 miles away; slides therefore had to leave campus, but the VR collection had to stay open for business during the project. UCSD prepared and delivered 12,500 slides per month to the vendor, which had two to three months to scan the slides before returning them, also at a rate of 12,500 per month. UCSD would be fined for late or missed shipments. UCSD had to
maintain a continuous cycle between shipments and inform and track patrons when subject materials were out being digitized.

In order to get additional resources for this project, UCSD had to estimate costs and workflow. In order to calculate what labor, space, and equipment that was needed for this project, they did time and motion studies to determine how long this project would take. These studies allowed them to calculate the time and space needed. Temporary staff was hired for this project, including several full-time hires as well as student workers.

Firstly, UCSD identified vendor slides and ensured that they were not included in the project. They pulled and cataloged uncataloged slides and created a project management website that included an online delivery schedule which listed what was out and when it would be back. They studied previous course outlines and future class schedules and tried to send subject areas that were not being currently taught, and then created a visual map of the physical collection, including what to send and when. The VR Curator pulled low-use slides to act as a backup in case they had difficulty sending their quota of 12,500 slides per month.

Slides were pulled, cleaned, placed in plastic sleeves, bar coded and scanned, and the library’s binding van took the slides to the scanning vendor. When the slides and digital images were returned, the digital images were given to IT staff and the slides were refiled.

When the first batch back was received from the vendor, they noticed that their vendors were overly enthusiastic with the post-process work of color correcting, contrast, etc. Thankfully, this was recognized early on and they were able to quickly resolve it. They also learned that quality control had to be more intense.

Overall, the project was extremely successful; not only does it make up the bulk of ARTstor, it also came in under time and under budget.

Heather Cleary, Otis College of Art and Design
“Creating an Online Image Repository: Small Steps and Large Leaps.”

Otis, an art and design school in southern California with one thousand students, received a Getty Electronic Cataloging Initiative grant to construct an artist's books database. Otis has made a strategic decision to commit to digitization in ten years, and have decided not to add new slides to the collection. Rather, they have digitized selected slides from the existing collection and have created digital images from other sources on a class-by-class basis with instructor lists. They have projected that digital and analog collections will coexist for seven years. MDID (called Otis-DID at Otis) was chosen as the delivery mechanism for this digital collection. Numerous demonstrations have been given to Otis lecturers and administration in order to get “buy-in” from the instructional community.

In January 2002, they began the pilot project with one VR librarian, one scanner, one contract programmer, and one enthusiastic professor. They surveyed the lecturer and the students on using digital images in the classroom; the response was that they all liked the digital images and the new technology, found them easy to use, offered 24/7 access, could be saved, edited, revised ,etc.
They then designed the following workflow; image lists were given by instructors, and VR staff would use these lists to gather slides and books. The materials were scanned, cataloged in their database, and trained students retouched the scanned images. Both master and retouched images were saved on CD then uploaded to the MDID server. The lecturer then created study guides and lectures via MDID. The library also subscribed to the digital image collections of AMICO and Corbis, and bought digital images sets from vendors such as Saskia and Davis. A year was spent normalizing the data and creating pick lists of authorities to assist student catalogers.

Equipment included one Windows 2000 server and a scanning station with a slide scanner with a bulk slide feeder. Otis has several smart classrooms with built-in PCs, a laptop hookup (with directions), wireless mouse, wireless keyboard, a VCR player, stereo, and a digital projector.

For copyright issues, they depend on Fair Use. They track the sources for all images and only provide images for study and educational purposes.

Currently, the VR library is staffed by one VR librarian, one programmer, four to five student workers, and two grant-funded interns. The collection has grown from four hundred to sixteen thousand digital images as of March 2005 and twenty lecturers use MDID.

Lessons learned: plan in context with user needs; create pilot projects; form an enthusiastic user base; give numerous demonstrations; create user-friendly technology set-up; and collaborate with other groups.