

Issues in Electronic Publication of Image Databases: Report from a Study of the Museum Educational Site Licensing Project

Howard Besser

UCLA School of Education & Information
howard@sims.berkeley.edu

Robert Yamashita

CSU San Marcos Dept of Science, Technology & Culture
yamashta@coyote.csusm.edu

This paper summarizes the findings of a two-year study of digital image distribution and publication, focusing on the Museum Educational Site Licensing Project (MESL). This study (Besser & Yamashita 1998) evaluated the costs, infrastructure, and efforts involved in implementing the MESL project, as well as user reaction to functionality. The study also examined costs of running analog slide libraries and compared these to costs and functionality associated with digital image distribution.

The MESL Project was the first US attempt to take a collection of images and accompanying metadata from a variety of museums and publish these in digital form on campus networks. It was a two-year experimental collaboration among seven museums and seven universities that distributed over 9,000 digital images and associated text for classroom use.

The study discusses cost-center models for looking at the distribution of digital and analog images, including creating digital images and metadata, mounting and distributing digital images, maintaining a distribution house, running a slide library, and an analysis of hybrid image libraries. It presents a comparison of user interfaces and search engines from the MESL universities. It also reports on the results of focus groups discussing faculty adoption of digital images for classroom use.

Background

The MESL Project was an ambitious attempt to distribute an identical set of 10,000 images and metadata from 7 cultural heritage institutions on 7 different university Intranets (Stephenson & McClung 1998). This ambitious multi-year project was launched in 1994, soon after the World Wide Web had begun. The project built on ten years of progress in distribution of images and metadata that represented works of art and other cultural heritage objects (Besser 1997b). The technical issues that this project tackled have been discussed elsewhere (Besser & Stephenson 1996).

Search and Interface Comparison

All seven MESL universities mounted the identical set of approximately 10,000 images and accompanying text records—each in their own way. These implementations varied widely, each university making different choices as to the search options, the indexed fields, display choices, and the overall look-and-feel of the access systems. Methods for access control, authentication, and the choice of text fields that displayed with an image differed not just from one university to another, but even within some universities these changed over time.

Because an identical set of 10,000 images and metadata was distributed to all seven universities, and because each university implemented its own delivery system, the project provided a unique opportunity to compare different implementations of the same data set.

Faculty Attitudes

One must understand the various obstacles to teaching with digital images before such image use is widely adopted within the instructional process. We conducted a study of faculty who teach classes in the visual arts using digital images. We wanted early adopters of this technology to tell us what problems they had faced and to speculate about which impediments needed to be overcome in order for their colleagues to begin using digital images. We conducted two focus group sessions with faculty from many different universities, and supplemented this data with observations from Beth Sandore's MESL-supported pre-and-post classroom survey of faculty and students using MESL images (Sandore 1998). We examined the results of a recent survey (Corbetta-Noyes 1998) of faculty and slide curators on using digital images in the classroom. We compared our findings to transcripts from a discussion on MESL impact among MESL project coordinators at the final participants' meeting. We also compared our findings to Eileen Fry's study of art history slide librarians and faculty (Fry 1998).

We discovered that, from the perspective of the faculty interviewed, current university infrastructures were woefully inadequate for using digital images in the classroom. The issues that dominated both focus group sessions were technical support and training; acquisition and development of digital tools (software and hardware); time commitment to learn and develop the new technology; and some form of academic (if unrecompensed) recognition for their efforts from the universities. One participant characterized the dilemma of starting to use digital images as a "*vortex of need*". The complexity of the aggregate components (equipment, training, support, and more) was the overwhelming barrier for most faculty. Participants perceived digital image technology as a significant means of reaching a broader audience and there was a general consensus that new technologies stimulate and encourage new ways of thinking. Likely barriers to widespread acceptance were:

Image Quality: Participants universally asserted that image quality was critically important, and that the digital images they were currently using were inadequate. However, when asked to elaborate (*What kinds of image quality do you need for what types of purposes? Is there a baseline of image quality that would get significant buy-in from your colleagues?*), most participants agreed that for many purposes, the quality of digital images was no worse than slides.

Technical Support and Training: The relationship between technical support staff and faculty was an issue in both focus groups. Another obstacle to implementation was the lack of adequate resources for technical support and training, which were often project-driven (and therefore temporary) rather than actively integrated within a given program.

Tools (software and hardware): Participants voiced a need for a variety of tools with which to manage, manipulate, and display images and descriptive data. They offered examples of both commercially developed "generic" tools such as PhotoShop and locally developed "custom" tools. The communications gulf between those who created the tools (programmers and technical staff) and those who used them (faculty, students, and visual resources staff) was listed as an impediment. Participants discussed the need for computer technicians with subject expertise.

Time and Recognition: Both groups were concerned with the value placed on their work

by their universities and with the enormous time commitment required to develop digital projects. Institutions rarely offered faculty release time or credit toward tenure for work with digital image technology. This was perceived as a formidable barrier to non-tenured teachers.

Metadata: In discussing the need for metadata to accompany images, participants stated the importance of being able to customize for their students both the interpretive data (commentary) and descriptive data (attribution, dates, provenance) that accompanies an image. Because art historians were eager to contribute their scholarship to an image database as long as they received credit for their comments, this issue was closely tied to that of recognition by the university. Participants speculated that the descriptive data about images should flow in both directions between the providers of the data (museum curators) and the universities (art historians).

Resources and Collaboration: Several participants had good experiences in forging alliances with colleagues from other departments to share resources. This was an interesting finding because it proved contrary to faculty's tendencies to favor autonomy and solo endeavors over collaborative projects.

Costs

The study compared the costs of running a university slide library with the costs incurred during the MESL project (Besser 1999). Though these two entities are not quite equivalent, they do replicate many of the same functions. Because no one is likely to repeat the MESL design in much the same way, the study tried to isolate and compare costs that are likely to persist in other models. And because technological development is rapidly altering many types of costs, the study sought to focus on the relationship of costs between different models (rather than the precise costs in any given model).

Slide library costs appear to be mainly in the areas of cataloging, acquisition, data entry, and filing. Although most of these costs disappear in the MESL model, new costs arise, mainly in the conversion and error-checking processes, as well as the very substantial (and expensive) ongoing process of maintaining collection management (text) records and clearing rights.

Key cost differences between analog and digital center around acquiring the image and putting it into service. Analog images are acquired individually or in small batches, and each must be individually cataloged. The MESL images were acquired in two very large groups, and came complete with cataloging information. The MESL images cost nothing to acquire, while the analog slides cost an average of \$3.25 each (and the analog acquisition process took almost five minutes per slide). In a production-level digital distribution system, the cost for universities to license digital images will certainly cost something more than the \$0 that they cost during MESL.

Although an initial reading of MESL per-image costs is more than double that of an analog slide library, when one removes the museum's cost of clearing rights and creating collection management records (a process museums may need to do anyway), the per-image time devoted to MESL drops to about half the time of acquiring an analog slide and putting it into distribution. This would suggest that, if all the other problems were overcome, a digital distribution model might support a modest cashflow back to the museum to help defray the cost of rights clearance and developing collection management records. But if this cashflow comes in the form of a licensing fee, one must be cautious in figuring an annual fee that is essentially replacing a one-time cost.

But while per-image digital distribution¹ costs may be lower, fixed overhead costs for MESL were substantially higher than for analog slide libraries. (And the figures gathered do not take into consideration essential infrastructure elements such as network installation and classroom projectors.) Though some of the costs for security and creating functionality are start-up costs, changing technology and evolving user needs will likely result in substantial ongoing costs in these areas. Running a digital project requires a much more complex installation than running a slide library. Digital delivery systems need ongoing technical support and much more sophisticated user training than do analog slide libraries.²

In examining university use of digital images, it is important not to ignore the great value of museum information. We believe that perhaps the most important thing that the emerging digital image consortia can provide to universities is the rich set of information about their objects.

Digital image distribution does not currently appear to be cost effective for universities when compared with existing analog distribution schemes. While the per-image cost to universities appears viable, the overhead and infrastructure needed to support digital image distribution is significantly higher than that needed to support an analog system. But as costs diminish and as other university activities begin to share the infrastructure costs, these types of schemes are likely to become more viable.

Key per-image savings to universities from digital image distribution consortia are likely to come from the receipt of museum collection management information. Though there are still some serious questions whether this information can substitute effectively for what faculty currently use in analog slide collections, this particular cost center holds the potential of offering a great deal of cost savings. We offer an analogy of effect of computerization on library cataloging. With the advent of automation (and copy cataloging), most of the cost of cataloging was not entirely eliminated, but shifted from the local site to an external centralized spot. Because an individual book was not repeatedly cataloged at different sites as before, this greatly reduced the overall total costs among all members of the system. But the actual cost of cataloging a book was not reduced; the cost savings came because the same book was not repeatedly cataloged by each site. It's also interesting to note that most sites continued to have a cataloging cost for each book—the cost of finding the centralized copy cataloging and, if necessary, adopting it to local needs. So with the advent of copy cataloging, the cost of original cataloging shifted to a central site, leaving a greatly reduced cost of copy cataloging at each individual site. Of course this process only worked for items that were held by multiple sites (and the total savings to the system was greatest for items held by a wide number of sites).

We do not believe that the emerging digital image distribution consortia can financially sustain themselves from the university market alone. Likely scenarios for subsidizing this activity include: museum contributions because they view this activity as part of their mission; leveraging the cost of a university distribution system by using the same system to deliver individual images to the museum's traditional high-paying analog image customers (advertising agencies and pre-press operations); and encouraging the production of added-value products which incorporate the images and metadata, and managing the sales of these to K-12 schools and individual consumers.

Even though there may be similarities between digital images and previous forms of image surrogates, it is important to note that there is a fundamental difference: digital images are available remotely. Previous transitions to new media have not altered the need for physical space to house a collection and a set of personnel to staff that space. Yet, if there were a

complete conversion to a fully digital world, there would be no need for physical staffing of a circulation facility since there would be no items to “check out”, nor would there be a need for a large physical space to house this type of surrogate. However, in a digital environment, physical space does not completely disappear; its use simply gets “reassigned” (in the form of online storage costs, server space, and workstation labs for accessing the material). A digital environment would still require staff for user training and technical support (critically important in a digital environment, but likely to be handled centrally for an entire university campus). Staff might also be needed for image selection and collection management, or these functions may be ceded to a consortium that provides digital images to most universities in North America.³

The allure of high technology and the desire to appear “technologically savvy” can drive administrators to pressure departments to move more quickly into the digital realm. Such mandates are justifiable given that space and labor are expensive commodities on college campuses. However, rapid moves to fully automate thus eliminating older technologies will probably *not* result in any cost savings—only a relatively expensive short-term “conversion” charge. Our observations suggest that the base cost of maintaining the functionality of the system will remain, and probably increase (although the increase will be masked by larger “campus-wide” technology initiatives). What is needed is a rational transition to the new technologies. Such a transition will require the support of the central campus administration. The acquisition of new digital image collections will severely impact local department (and college) budgets. The cost to convert legacy data and custom and campus specific images into new formats will be enormous. The local development will require technological sophistication so these data can be used in conjunction with purchased digital formats. Most importantly, these moves will need to be done over time and in conjunction with existing facilities.

University administrators considering digital image distribution schemes need to be aware of a number of critical issues. One overwhelming important issue is the perishability of information in digital form (Lyman and Besser 1998). Three key questions that administrators need to address are: What might be lost if digital distribution schemes replace analog slide libraries? Will target groups use digitally-distributed images? And what entity within the university will contract for digital distribution rights?

Conclusion

The study discovered that, as a whole, the digital distribution environment appears to be good for individual usage, and provides access from multiple locations. But digital image distribution in its existing form is problematic for group viewing situations, such as in the classroom. Electronic classrooms, computing and network infrastructure, technical and instructional support, and image quality issues need to be addressed before digital distribution to the classroom becomes viable.

The interface and the ability to query and manipulate the database is critical for future use. Additional tools for examining, organizing and saving retrieved sets are also necessary. The MESL model of localized control over distribution discouraged development of expensive retrieval systems. A more centralized model would be able to spread the development costs over a wide body of sites, and likely lead to better retrieval tools. But local customization of such a system may still be desirable, and this poses an interesting research issue in system design.

The different metadata vocabulary and general language used by different institutions made the creation of an integrated and consistent database problematic at best. It is glaringly evident

that a project like this needs guidelines and standards at many levels (from field delimiters to controlled vocabulary).

The lack of comprehensive content made the database extremely problematic for classroom purposes. For a digital image publication scheme to be successful, it must be able to provide a critical core of important images. But the contents of a "critical core" are constantly shifting due to new approaches to disciplinary understanding.

Because faculty content needs can be robust and shifting, a digital image distribution scheme will almost certainly also need to give faculty the option of integrating locally produced material. Future systems must be both extensible and easy to supplement.

Citations

- Besser, Howard. (1999). The MESL Experience versus Slide Libraries: Comparison and Analysis, *Visual Resources: An International Journal of Documentation* 14 (4), pages 477-499.
- Besser, Howard. (1998). MESL Implementation at the Universities, in Christie Stephenson and Patricia McClung (eds.), *Delivering Digital Images: Cultural Heritage Resources for Education*, volume 1, Los Angeles: Getty Information Institute, pages 70-84.
- Besser, Howard. (1997a). Comparing Five Implementations of the Museum Educational Site Licensing Project: "If the Museum Data's the Same, Why's It Look So Different?". *Proceedings of the Fourth International Conference on Hypermedia and Interactivity in Museums*, Pittsburgh, PA: *Archives and Museum Informatics*, pages 317-325.
- Besser, Howard. (1997b). Image Databases: The First Decade, the Present, and the Future. *Digital Image Access & Retrieval* (Papers presented at the 1996 Clinic on Library Applications of Data Processing, March 24-26, 1996), editors P. Bryan Heydorn, and Beth Sandore, Urbana, IL: University of Illinois, pages 11-28.
- Besser, Howard and Christie Stephenson. (1996). The Museum Educational Site Licensing Project: Technical Issues in the Distribution of Museum Images and Textual Data to Universities, *E.V.A. '96 London (Electronic Imaging and the Visual Arts)*, editor James Hemsley, Hampshire, UK: Vasari Ltd., pages 5:1-5:15.
- Besser, Howard and Robert Yamashita. (1998). *The Cost of Digital Image Distribution: The Social and Economic Implications of the Production, Distribution and Usage of Image Data*, a report to the Andrew W. Mellon Foundation, Berkeley: School of Information Management & Systems, 1998. (<http://sunsite.berkeley.edu/Imaging/Databases/1998mellon/>)
- Corbetta-Noyes, Lorna. (1998). *Using Digital Images in the Classroom* [Survey Results]: <http://www.rlg.org/home/disummary.html>
- Fry, Eileen. (1998). Slide Circulation Study, *VRA Bulletin*, Fall.

Lyman, Peter and Howard Besser. (1998). Defining the Problem of Our Vanishing Memory: Background, Current Status, Models for Resolution, in Margaret MacLean and Ben H. Davis (eds.). *Time and Bits: Managing Digital Continuity*, Los Angeles: Getty Information Institute and Getty Conservation Institute, 1998, pages 11-20.

Sandore, Beth. (1998). Findings of the Instructor/Student Survey, in Christie Stephenson and Patricia McClung (eds.), *Delivering Digital Images: Cultural Heritage Resources for Education*, volume 1, Los Angeles: Getty Information Institute, pages 100-121.

Stephenson, Christie and Patricia McClung (eds.). (1998). *Delivering Digital Images: Cultural Heritage Resources for Education*, volume 1, Los Angeles: Getty Information Institute.

Notes

1. MESL University deployment is an overhead cost, not a per-image distribution cost. This type of deployment involves creating a set of tools that are needed, and that total costs within that cost center do not change if more images are added to the system.
2. Even though no costs for slide library training and outreach were reported, costs in this area would be extremely small when compared to the cost of training and outreach for a digital distribution system.
3. Currently planned consortia are not really international due to differing intellectual property laws.