

# Digital Archives: Cal Poly's Experiences with Creation and Dissemination

Ilene Rockman

*California Polytechnic State University, San Luis Obispo*

---

Ilene Rockman, "Digital Archives: Cal Poly's Experiences with Creation and Dissemination." *Proceedings of the IATUL Conferences*.  
Paper 41.

<http://docs.lib.purdue.edu/iatul/1996/papers/41>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact [epubs@purdue.edu](mailto:epubs@purdue.edu) for additional information.

# **Digital Archives: Cal Poly's Experiences With the Creation and Dissemination [✱]**

*Rockman, Ilene F.  
California Polytechnic State University*

---

Academic libraries recognize that their mission is to provide quality services and collections in support of the instructional and research needs of the parent institutions they serve. Over the past decade, these libraries have experienced significant technological changes, as they have attempted to provide convenient and timely access to locally held information, primarily through online public access catalogs. In recent years, technological advancements have accelerated the ability to support a robust physical infrastructure essential to maintain a wide variety of functions, such as the delivery of text, image-based, electronic, and digitized resources. Librarians, and others, have actively designed easy-to-use interfaces. The worldwide web has greatly facilitated these efforts.

This paper will discuss some of the salient issues related to transforming selected collections to digital formats, not only for preservation purposes, but also for increasing access to unique and sometimes fragile sources. The paper will also share the experiences of one library relative to the selection of materials, mechanisms for document conversion and imaging, access issues (with sensitivity to copyright consideration), and finding aids to support the access.

## **Why Digitize?**

An important function of any archival agency is to communicate with researchers about the holdings of the agency. The archivist strives to describe collections, and the record groups within them, in such a way that researchers can determine whether or not there is pertinent information related to the nature of their inquiries. Traditionally, archival and manuscript materials cannot leave the agency premises, and are often exposed to repeated handling.

The concept of a digital archives suggests the possibility that selected materials can be perused independent of physical location and time constraints, minimizing paging and reshelving activities by library or archival staff members.

In 1995, the Kennedy Library of California Polytechnic State University decided to investigate the feasibility of digitizing a small portion of its most prominent special collection, the Julia Morgan collection. Researchers from all over the world have travelled to the Kennedy Library to pursue research related to Ms. Morgan, the premier female architect of California, best known for her design work of the San Simeon State Historical Monument (also known popularly as the Hearst Castle). The library was interested in determining if a small portion of the collection could be delivered electronically to researchers. The permanent collection, housed in 30 acid-free record boxes and one map case, includes personal and professional papers, office records, correspondence, photographic materials, architectural drawings and blueprints, and nonpaper items. Delivering selected parts of the collection through the

worldwide web could make the materials more accessible to national and international scholars.

### **What to Digitize?**

Based upon previous interest shown by those searching the text description of the collection through the library's gopher site, a decision was made to extract a small portion from the collection representative of correspondence, drawings, and photographic images. Not only would this sample represent core areas from the collection, but it would present a diversity of materials from which to gain valuable digital experience.

Selection of the unique materials was completed by the department head, in conjunction with a graphic designer and a web specialist. All parties were concerned with the condition of the original materials, the accurate representation of these primary source materials, and the aesthetic principles of how these materials would scan and display on the computer screen. Also considered was the need to select a content rich sample which could form the basis of a descriptive interactive finding aid to be created in standard generalized markup language (SGML), an internationally recognized standard for representing text in electronic form.

The SGML-based Encoded Archival Description (EAD) format in use at several universities, such as the University of California at Berkeley [1], served as a model for the Morgan project. SGML was preferred to Hypertext Markup Language (HTML) because it permits encoding of the intricacies that might be lost after transforming items into electronic text. [2] Although HTML is based on SGML, it has limited tags and is used to describe only the general structure and content of documents for presentation. SGML is useful for standardizing archival inventories and registers, serving to identify logical components of a record, and providing intellectual control over a collection. Because SGML can support description, navigation, presentation, and indexing of materials, it has broader applications for long-term use. Therefore, items were selected with this application in mind.

### **How to Digitize?**

Using the existing equipment on site in the Kennedy Library, the library began to develop the project. Available equipment included:

**Powermac 8100/100 server** with 24MB RAM. The server has a speed of 100MHz, with built-in ethernet capability. A 256K cache is standard.

**La Cie SilverScanner II** 24-bit color flatbed scanner, with the capability to scan up to 8.5" X 11" paper. The document feeder permits scanning multiple pages at a time. Later, a transparency adaptor was added.

**Adobe Photoshop 3.0**, a photomanipulation program which is more powerful than optical character recognition (OCR) software.

**Apple QuickTake digital camera**, offering the capability to take high resolution pictures which could be later sized and manipulated by the Photoshop software.

Recognized standards regarding document imaging were employed and tests were successful, but it was apparent that another imaging system needed to be configured for the project. Strong consideration was given to having equipment on site in the Special Collections and University Archives Department, rather than in another location in the library (as had been the case). Fortunately, ethernet connectivity was not a problem, so that interoperability for effective communication across networks was not an issue. User access was restricted at first to protect the integrity of the project. Later, hypertext links to the library's homepage were made.

Additional consideration was given to the following questions:

- Does the digital copy capture acceptable details of the original without enticing users to capture and reproduce it without permission?
- Will the digital conversion damage or harm the original materials?
- Is the process cost effective for the intended use?
- Will the institution assume responsibility for preserving both the original material and the digitized image? Changes in computer hardware and software mean that the digital image may have to be periodically transferred to a new storage medium. Are costs manageable if electronic environments change? [3]

### **Technical Issues**

Scanning was set at 72 dots per inch (DPI) so that the scan could be viewed on the computer screen. It was set in the Joint Photographic Experts Group (JPEG) compression format, rather than in the Graphic Interchange Format (GIF). Conversion ratios were 8:1 to permit access to fragile materials while still maintaining a level of screen legibility. JPEG, a widely used industry standard for compression, supports full color image, unlike GIF. JPEG also uses a sophisticated technique called a discrete cosine transformation to produce a sliding scale of graphics compression [4]. Large compression ratios are possible, for faster download speeds, and full-color images (24 or 32-bit images) are supported. JPEG is especially well suited to photographs and complex images.

### **The Site**

A prototype digital archives web site was developed and tested for the Julia Morgan Collection. It is now available for viewing at <http://www.lib.calpoly.edu/library-info/morgan/morgan.html>

### **Lessons Learned**

After almost a year of planning and testing, the staff has reflected on the progress to date. They offer the following advice:

- Don't be afraid to consider establishing an archive of digital materials as a way to "open up" collection access. Start small, and strongly consider adopting national standards.
- Choose to provide access without compromising the integrity of the original materials. Images can be reduced to provide access, while discouraging illegal use and copyright violations.

- When possible, engage the talents of a graphic designer and experienced webmaster to maximize the time and energy of all involved with the project. Consider how the project may increase in scope and scale, and be mindful of the need to migrate to additional hardware platforms as it grows in size.
- Have fun - this is an exciting development in information access, and a terrific learning experience for all involved.

## References

1. Berkeley Digital Library SunSITE <http://sunsite.berkeley.edu/FindingAids/>
  2. PRICE-WILKIN, JOHN. A gateway between the World-Wide Web and PAT: exploiting SGML through the Web. The public-access computer systems review 5, no. 7, 1994: 5-27. To retrieve this file, use the following URL: <gopher://info.lib.uh.edu:70/00/articles/e-journals/uhlibrary/pacsreview/v5/n7/pricewil.5n7> or <http://info.lib.uh.edu/pacsrev.html>
  3. SCHMIDLE, RICHARD E. The smile and promise of digital imaging: preserving photographs in a digital world. Library Hi Tech News 130 March 1996: pp. 14-16.
  4. LYNCH, PATRICK J. Creating graphics for the World Wide Web: file formats for Web pages. Syllabus 9 (4) January 1996: pp. 46-48.
- \*) Appreciation is extended to Nancy E. Loe, Head of the Special Collections and University Archives, Kennedy Library, Cal Poly, San Luis Obispo, California for her help with the preparation of this paper.