



Infinity

The newsletter of the SAA Preservation Section
Fall 2007 / Spring 2008, Volume 23, number 1

NEWSLETTER HIGHLIGHTS:

From the Section Chair

From the Editor

Building a Digital Preservation Program

Digitizing Oversized Visual Materials

When Cost Matters: Deploying a Digital Preservation System on a Small Budget

The "Why" Factor: The Importance of Goals for Digitization

Five Issues in Planning Long-term Preservation in an Institutional Repository

SAA Preservation Section: Ballot for 2008 Elections

SAA Preservation Workshops: A Sampling

SAA Annual Meeting 2008: Dates and Hotel Information

Section Committee Members

FROM THE CHAIR

Dear Colleagues:

The steering committee and I have worked this year to advance the strategic planning initiatives that Steve Dalton presented in Chicago during the 2007 business meeting. As a reminder, they are as follows:

1. update our operating practices to reflect the responsibilities and activities of our committees;
2. conduct monthly one-hour meetings via teleconference;
3. develop intern opportunities for each of the Section's committees;
4. broaden involvement in the committees by bringing in more members from the Section;
5. create a committee structure that features co-chairs with staggered terms;
6. expand our business meeting to 3 hours wherein the last hour serves as an all-committee meeting;
7. continue to pursue electronic balloting.

I'm happy to report that the section will continue with electronic balloting again this year. Our committees will have an opportunity to meet after the business meeting in San Francisco. And I began the process of broadening involvement with a call for volunteers last fall, to which I received an encouraging response. Part of the broadening initiative is to engage graduate students in committee operations, and to encourage their participation at an early, and active level. This initiative will be addressed to a greater extent during the upcoming year.

Through our monthly calls, the steering committee has stayed on top of our top priority, which is to convert our "Operating Practices" into formal bylaws. This process gives us the opportunity to examine our leadership structure, evaluate our practices, and determine whether leaders and committees are doing what we say we should be doing. Folded into this is to create a committee structure that facilitates a natural progression of leaders from one year to the next. The long-term success of any section or committee rests on the ability to attract and motivate new volunteers in the organization and to have those "newbies" learn from the experience and longevity of those that have served and nurtured the organization well. I look forward to presenting the fruits of this year's work to you at the business meeting in San Francisco.

Mark on your calendars that our business meeting will be on Friday, August 29, from noon until 2:00 p.m. Additionally, committee meetings will be held immediately after the business meeting and will adjourn at 3:00 p.m. This year's business meeting will feature a panel exploring *The Drive to Digitize* and the impact that digitization has had on overall preservation programs and funding for them. Our presenters, still to be determined, will present the issues of managing resources (staff, space, and money) for digitization and preservation programs and the conflicts and advantages that occur. You will not want to miss it!

With this issue of *Infinity*, we continue our focus on digital activities, and I want to thank our contributors for their insights: Elise Dodeles, David Joyall, Anthony D. Smith, Jennifer Greenwood, and Kyle Conner.

Brenda Gunn
Chair
Preservation Section

FROM THE EDITOR

Dear Preservation Colleagues:

This issue of *Infinity* has been a long time coming! With a continued focus on digitization, we reached out to our friends and colleagues in the College/University, Manuscripts, Oral History and Visual Materials Sections. Their contributions have added depth and scope to our discussion of digital preservation, and to all who contributed I am truly grateful.

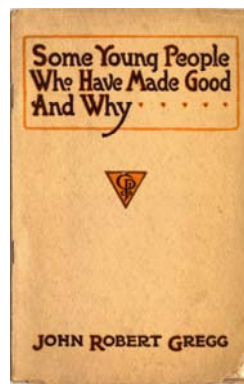
As we move into the summer months and those humidity sensors begin to rise, I am reminded that the Joint Annual Meeting of NAGARA, COSA, and SAA in San Francisco is just around the corner. The annual meeting in San Francisco will offer another opportunity for us to further our

conversation on digital preservation initiatives, and I look forward to seeing you all there. Before we convene in beautiful northern California, however, I'd like to take a moment to encourage you all to contribute to our summer edition of *Infinity*; submissions are due by July 1, 2008, and we're accepting pieces relating to all areas of preservation. I'm looking forward to reading about your latest endeavors!

Tonia Sutherland
Editor, *Infinity*

Building a Digital Preservation Program

By Elise Dodeles, Preservation Technician
Rider University's Moore Library



Cover of John Robert Gregg's
Some Young People Have Made Good and Why

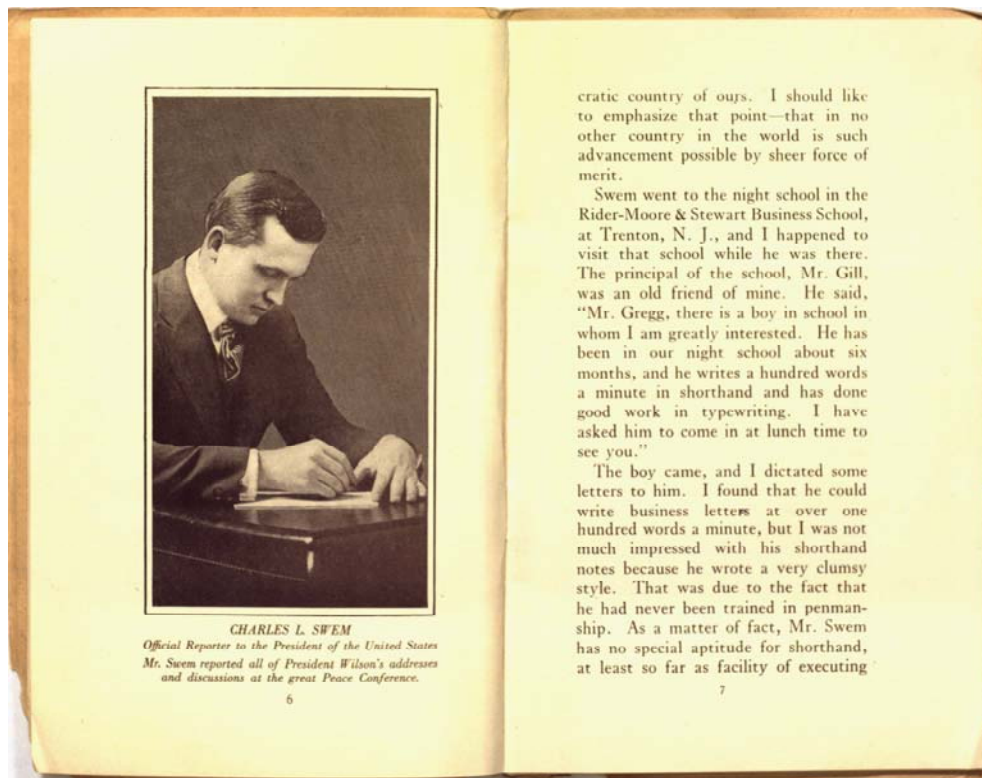
As Preservation Technician at Rider University in New Jersey, I oversee our digital preservation program. Thus far, we have concentrated our attention on our Louis A. Leslie Collection of Shorthand Materials. I choose items for digital preservation and by extension our digital library, based on the following criteria:

- Condition of item. (i.e., item has brittle pages, or cannot be handled without damaging further). Sometimes this is my first priority, but I usually balance this with
- Copyright concerns. Some items are no longer under copyright restrictions, so these will usually get precedence. Those items with copyright restrictions need permissions from the copyright holder.
- Importance to the history of shorthand and its development.
- Importance to the growth of Rider University. Our University is in a unique position in relation to the Leslie Collection. We began as a Business School that stressed the teaching of shorthand, typewriting and penmanship. John E. Gill, an early Dean of the Rider-Moore & Stewart Business School, was friends with John Robert Gregg, creator of Gregg shorthand. Gregg, in a 1917 work entitled *Some Young People Who Have Made Good*

and Why speaks of a Rider student, Charles L. Swem, who went on to become Official Reporter to President Woodrow Wilson.

The above-mentioned pamphlet was chosen for digital preservation because it met all 4 criteria. The front cover of the 40 page booklet has torn and damaged edges. Its date of 1917 is before the cut off date of 1923 that has been recommended in our University's legal guide. Gregg shorthand was the most popular form of shorthand taught in this country from the 1930s-1970s. Gregg's thoughts on automation, women and business, and perseverance through hard work speak to the history of shorthand, business education and the culture of the times. As mentioned previously, Gregg had a significant relationship with the Rider Business School.

Additionally, to keep this project afloat, we have received several John Robert Gregg Fund grants from the New York Community Trust. The Trust has made it a point to fund projects such as ours that are interested in the dissemination and education of shorthand skills and knowledge. We received our first grant 18 years ago when the library was given the Louis A. Leslie Collection. The money was left untouched for 2-4 years. With no official archives, the librarian in charge of the collection categorized the items broadly and housed them in acid free boxes. This and the next 2 grants in the amount of \$250,000 were used for further cataloging, preservation and housing. The last two grants have been used to extend the digital preservation of the shorthand items. After the request for our last grant was declined, our Chairperson, who had built a solid relationship with the Program Officer at the NY Community Trust, was given the invaluable advice to consider the grant request as an extension of previous grants, rather than a new grant. This change proved crucial as it allowed our project to continue to be funded.



Charles L. Swem, Official Reporter to President Woodrow Wilson

Digitizing Oversized Visual Materials

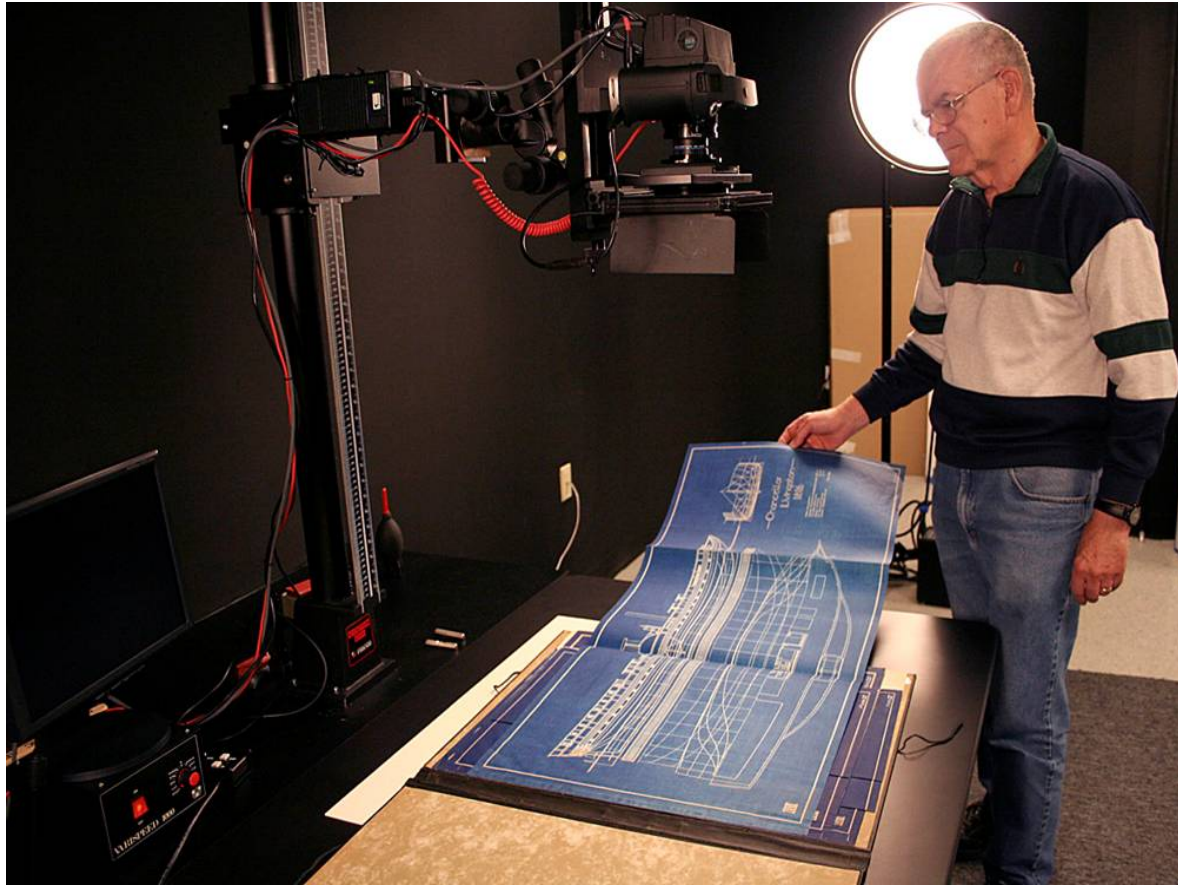
by David Joyall, Technical Photographer
Northeast Document Conservation Center (NEDCC)

Most historical collections and archives include objects that don't fit conveniently on a table top scanner. Most scanners are designed to fit flat paper up to 11"X17". Dealing with oversized materials always complicates a project; whether it's finding storage space, conservation treatments or digitizing. Physical defects that commonly plague historic paper documents such as mold damage, brittleness and tears are compounded when dealing with large items. During the digitization process, the first priority is minimizing the stress on the object.

Large objects such as posters and maps need special digital camera and lighting equipment to photograph the object and create file sizes large enough for reproduction. By mounting the digital camera on a copy stand, we can photograph objects safely on a horizontal surface. Output needs will dictate the choice of camera. For web use, digital SLRs will be more than adequate. Digital Single Lens Reflect cameras can capture file sizes up to 21 megapixels. For more accurate results, medium format cameras fitted with a digital back are necessary. Digital backs can record an enormous amount of information. The sensor array of a digital back can capture up to 39 megapixels in one shot. Because the subject matter we are photographing are inanimate objects, we can take advantage of the multi shot mode of these backs and create files sizes well over 500 mb .

Color management is important for any digital project. For oversize objects, color management starts with the room where the capture takes place. Lighting for these oversized jobs has to be even. Fortunately, special software is available to detect uneven lighting and compensate for unwanted dark or bright areas. Walls need to be painted a dark neutral gray to absorb excess light bouncing off the objects. Light that bounces off light colored walls and ceilings can affect tone contrast and create unwanted color cast. A color viewing booth fitted with 5000K lamps offers a controlled environment for comparing originals to images on a monitor and prints. Fluorescent bulbs in the overhead fixtures can be replaced with bulbs similar to the ones used in the viewing booth to create a large white light viewing area.

There are many other two dimensional objects that could be handled with a digital camera/ copy stand system. Books and atlases with weak bindings can sit on a book cradle and be photographed. Architectural drawings and blueprints that need to be copied with no distortion benefit from the latest designs of digital camera lenses. With the right lens, edge to edge sharpness is achieved. Prints that are framed and behind glass can be copied with minor adjustments to the lighting. When compared with conventional scanners, digital cameras offer a safe and effective alternative for these hard to handle materials.



Digitizing blueprints using a Sinar digital camera

When Cost Matters: Deploying a Digital Preservation System on a Small Budget

by Anthony D. Smith
Director, Digital Initiatives and Services
University of Miami

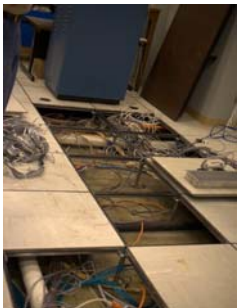


The University of Tennessee Libraries is taking a new approach to digital preservation with its adoption of *Dark Archive in the Sunshine State* (DAITSS) system, which will serve as its new digital repository. DAITSS was developed by the Florida Center for Library Automation (FCLA) to help support the preservation functions of format normalization, mass format

migration, and migration on request. Digital objects are deposited into the DAITSS system using a standard submission information package (SIP), which is constructed using the Metadata Encoding and Transmission Standard (METS). The University of Tennessee implementation of DAITSS will initially be used to archive digitized masters from various libraries, museums, and archives throughout the state. Archiving of these materials is being carried out as part of an Institute of Museum and Library Services (IMLS), grant-funded project titled *The Growth of Democracy in Tennessee*.

UT Libraries purchased the dedicated hardware to run the DAITSS system based in part recommendations from the FCLA development team. The system was configured to provide disk storage to accommodate a three-year growth period and to allow for the addition of new storage arrays with minimal difficulty. A rack-mounted system was purchased to be housed in a dedicated full-height rack (\$1600). A Dell PowerEdge 2850 was purchased (\$3600) to perform the system processing. The PowerEdge was configured with two Intel Xeon, 3.0 GHz processors, four gigabytes of memory and it has its own internal 73 gigabyte disk drive for Red Hat Linux and the DAITSS installation files. We also purchased a VTrak M500i SATA RAID Storage System (\$4900) to handle our storage requirements. The M500i is described as a “budget-maximizing, high-performance RAID storage solution optimized for organizations deploying small to medium application clusters, disk-to-disk backup and mid-range SANs.” It uses an iSCSI interface and can house up to fifteen inexpensive serial ATA disk drives. The UT Libraries system was deployed using fifteen 500 GB drives (\$375/each), allowing for approximately 6.5 terabytes of storage after RAID allocation. Additional M500i storage systems can be purchased as needed to meet digital archiving requirements. The full-height rack can accommodate additional storage units by daisy-chaining the new unit to the existing M500i. One terabyte serial ATA drives are now available meaning that a new M500i can be added with a 13 terabyte capacity for the same price as the original 6.5 terabyte array. UT Libraries also purchased a dedicated tape library system (\$32,000) to accommodate the massive backup requirements of the digital archive. The new tape library is a Dell PowerVault 136T. The system utilizes LTO-3 technology, meaning each tape can store up to 800 gigabytes. The library can house 36 tapes (\$40/each) at once and has two drives for greater backup efficiency. Incremental backups of the digital archive are scheduled to occur daily with full backups once a week. UT Libraries conducts an off-site tape rotation cycle that utilizes three sets of tapes.

The total hardware needed to deploy the dedicated digital archive system cost about \$60,000 with the ability to add another 13 terabytes for roughly \$10,000. This system was developed to provide long-term digital stewardship of scanned images of primary source materials from state-wide cultural heritage institutions. It can very easily accommodate a variety of born digital materials from within the university community as well.



“Here's the area we selected for the new tape library. We first had to locate a power source and network connections. That meant we had to go under the dreaded floor.”



The “Why” Factor: The Importance of Goals for Digitization

by Jennifer Greenwood, Freelance Archivist and Associate Curator
Paul Gray PC Museum, Claremont, CA

“**W**hat should I digitize?” This is a loaded question and the answer lies in understanding the goals and mission of the archival digitization project. It is not enough to understand what to digitize, but also why the desire to digitize, and ultimately store it, is so strong.

When I was new to the archival profession, I would sit down with my clients and casually discuss their digitization “wish lists.” While this was mildly effective, I still found that there was some uncertainty during the actual decision making process. Many of my clients were not certain about their goals or desires for their own collections. They simply knew that they wanted their collections in some electronic format.

So, after years of meeting with clients, I began to understand the benefit of structure and developed a questionnaire. During my initial meeting with my client I present him or her with a list of guided questions, including:

1. What is the main focus of your collection? (grandparents’ lives, children’s birthdays, etc.)
2. Why do you want to digitize this particular collection? What is the purpose of having this collection in electronic format?
3. Will this be an ongoing digitization project? As new photographs and/or documents are received, do you plan to add them to this digitized collection?

I found that these 3 questions, more than any others, allowed my clients to focus on the ultimate goals of digitizing their collections, consequently allowing me to better assist them as their treasured heirloom pieces made the leap into the digital realm.

The “why” factor is an integral piece of the digitization puzzle because it gives a collection substance and depth that is impossible to achieve by simply digitizing a mass of information. By creating a cohesive foundation of goals, and implementing an organizational structure based

on those goals, digitization can create a multidimensional product that provides information regarding a particular subject or time frame.

Digitization is an exciting opportunity to distribute information more widely than ever. Through the utilization of responsible, organized, traditional archival techniques, the future looks brighter than ever for electronic collections.

Five Issues in Planning Long-term Preservation in an Institutional Repository

by Kyle Conner, International Monetary Fund, Washington, DC

Design of a long-term preservation plan for a new institutional repository is no easy undertaking. Preserving information is the goal, but it is impossible to save everything, in every format. Here I raise five vital issues that a successful preservation plan should address.

First, with hundreds of file types known to exist, selection of trusted “standard formats” must be done to control the variety of material and to more effectively design supporting systems. Proprietary formats such as .pdf, .xls, and .rm are not archival as format support is only guaranteed as long as the company chooses to do so. When a company collapses, reliable format support can’t be guaranteed and formats become obsolete. Subsequently, migration becomes an option for retaining the data. Emulation, decreased interactability, and data loss are issues that should be addressed when developing an effective policy regarding format migration.

Secondly, intellectual property and privacy laws regarding the deposited items must be addressed. Prior published materials must be deposited with an authorized statement from the copyright holder guaranteeing certain rights to the works. Authors must also grant the repository the right to make some of the material available to encourage access and searching capabilities. Authorization to make copies should be also granted, as their use in backup systems is crucial in case of a malfunction. Restrictions on access should be addressed in such a way that restricted information can not be revealed to users or even hackers. Placing sensitive documents on a separate server or removing them completely from the system are good ways of monitoring and restricting access to the documents.

Thirdly, maintaining the storage itself is a major concern. Having the physical space and security to house the necessary servers and network connections is crucial to keep the service stable. Modularity and scalability are key features to keep in mind to accommodate system growth while reducing future costs. To provide secure retention of data, routine backups must be made with the data distributed to protect against loss and damage. Partnering with another institution to exchange copies and backups to prevent loss from disasters, both natural and electronic, is an effective way to accomplish this.

The fourth issue to address is providing consistent and reliable access to materials. Retaining a valid permanent URL for each item and standardized metadata is crucial for

citations and for long-term use. An effective metadata scheme should address the document's characteristics as well as indicate its provenance, terms of use and guarantee its authenticity and integrity. Being able to routinely find datasets is necessary to retain good relationships with users as well as donors.

Finally, guaranteeing the authenticity and integrity of the records is crucial. Digital integrity ensures data has not been distorted or altered since the time it was created and deposited. One way of accomplishing this is to secure data with a checksum or one-way hash. Following this "content locking" a testing schedule should be developed to monitor documents periodically to guarantee they retain these locks.

These five issues form a framework for any successful IR policy. Further investigation into the current state of institutional repositories is necessary to flesh out any successful policy. By keeping these issues in mind and by utilizing organizations such as the Registry of Open Access Repositories (ROAR) a successful IR policy is just around the corner.

Preservation Section – Election Alert!

Voting will open in mid-June for the 2008 Preservation Section Election, which will be conducted entirely by electronic ballot. Members will be notified via e-mail when the ballot is available. In the meantime, here is our wonderful slate of candidates:

Chair-Elect: (Choose One)

Pat Morris
Rebecca Hatcher

Steering Committee: (Choose One)

Anselm Huelsbergen
Laurie Gemmill

Nominating Committee: (Choose One)

Winter Shanck
Linda Peterson
Rachel Onuf

SAA Preservation Workshops: A Sampling

| Upcoming Conferences/Workshops | | |
|--|--|---|
| Available | Conference/Workshop | Location |
| On Demand | Thinking Digital ... Practical Session to Help You Get Started Web Seminar | Online/On-Demand |
| On Demand | Electronic Records: Preservation of PDF Web Seminar | Online/On-Demand |
| 2008 Dates | Conference/Workshop | Location |
| May 19, 2008 | Building Digital Collections #0838 | Paterno Library, Pennsylvania State University University Park, PA |
| Jun 12, 2008 - Jun 13, 2008 | IT Training for Practicing Archivists Series | Haverford College Haverford, PA |
| <i>Including the following programs:</i> | | |
| Jun 12, 2008 | Digitization of Archival Materials #0854 | |
| Jun 13, 2008 | Digital Libraries and Digital Archives #0855 | |
| Aug 24, 2008 - Aug 31, 2008 | ARCHIVES 2008: Archival R/Evolution & Identities | Hilton San Francisco San Francisco, CA |
| <i>Including the following programs:</i> | | |
| Aug 26, 2008 | The Essentials of Digital Repositories #0913 | |
| Aug 26, 2008 | Preserving Your Audio/Video Assets #0916 | |
| Oct 17, 2008 | Legal Aspects of Photography Rights, Archive Mgmt, & Permissions #0919 | New York State Archives, Cultural Education Center Albany, NY |

For questions, please contact: education@archivists.org.

SAA ANNUAL MEETING 2008 PREVIEW



Archives 2008
Joint Annual Meeting of NAGARA, COSA, and SAA
San Francisco, CA
August 26 - 30, 2008

Note that the Preservation Section Meeting will be on Friday, August 29, from noon until 2:00 p.m. Additionally, committee meetings will be held immediately after the business meeting and will adjourn at 3:00 p.m. This year's business meeting will feature a panel exploring *The Drive to Digitize* and the impact that digitization has had on overall preservation programs and funding for them.

Preservation Section Committee (<http://www.archivists.org/saagroups/preserv/index.html>)

Brenda Gunn (Chair)

bgunn@mail.utexas.edu

Steve Dalton (Past-Chair & Chair, Nominating Committee)

daltonst@bc.edu

Elizabeth A. Slomba (Vice Chair)

eslomba@unh.edu

Rebecca Hatcher (Member-at-Large & Co-Chair, Education Committee)

rebecca.hatcher@yale.edu

Susan Koutsky (Member-at-Large & Co-Chair, Education Committee)

skoutsky@umd.edu

Julie Graham (Web Liaison)

jgraham@library.ucla.edu

Patricia Morris (Chair, Program Committee)

Patricia.Morris@colorado.edu

Tonia Sutherland (*Infinity* Editor)

tsutherland@library.umass.edu