Newsletter of the Society of American Archivists
Science, Technology, and Health Care Roundtable

Summer 2012

Contents

Message From the Co-Chairs

John Rees
National Library of Medicine

Melanie Mueller
American Institute of Physics

Come one and all to the Science, Technology, and Healthcare (STHC) Roundtable this year at the Society of American Archivists Annual Conference at the San Diego Hilton Bayfront! The roundtable will meet from 3:15 to 5:15 p.m. on Wednesday August 8th, 2012 in the Sapphire O/P room. Of course, be sure to check the final on-site program for any last minute location changes. We had a relatively active year and look forward to describing those activities as well as hearing about your new ones for 2013. Personally, I'm interested in re-engaging our natural sciences colleagues and any archivists participating in big data curation projects, particularly as informed by the National Science Foundation's new data management plan grant policies. And as always, consider serving the Society by volunteering for a leadership position within the roundtable.

STHC is a forum for archivists working at institutions in the natural and social sciences, technology, and the health sciences. The roundtable provides a means for its members to share problems, projects, and products that they have in common. Each year, the roundtable's meeting provides opportunities for members to network, share experiences and successes, and discuss ways for archivists working within scientific, technology, or health care organizations to solve common challenges. However, all interested individuals are most welcome to attend!

Our annual business meeting, including the election of a new co-chair, will be conducted during our roundtable session. Our program this year will feature three presentations relevant to archivists from all three of our roundtable's specialty areas; please refer to the program below for further details.

(continued)

We are also pleased that the only STHC-endorsed session made it into the program this year: Session 107, "From Hidden Collection to International Incident: The John Cutler Papers and the Guatemala Syphilis Experiments," 10 a.m. August 9th.

Give your sun tanning a break and sail on over for what is sure to be an entertaining and informative Roundtable and Annual meeting!

STHC Roundtable 2012 Meeting

Wednesday, August 8, 2012, 3:15-5:15 p.m.
San Diego Hilton Bayfront, Sapphire O/P

Business Meeting, 3:15-3:45

Welcome and Introductions

Council Representative Announcements: Tom Frusciano

Approval of 2011 Minutes / Old Business

Reports:

- Annual Meeting Taskforce: Jodi Koste
- HIPAA/HITECH and/or Common Rule regulatory update: Phoebe Evans-Letocha
- *Archival Elements:* Liz Phillips
- STHC Website: Polina Ilieva
- STHC listserv
- Leadership Activities during 2011-12
- Steering Committee Membership
- STHC Co-Chair Candidate Report: John Rees
- SAA 2012 Annual Meeting Sessions Proposal Report: John Rees

Program Presentations, 3:45-4:45:

AIP update on total collection digitization
Chip Calhoun, Technical Services Archivist

Smithsonian Archives Field Book Project
Tammy Peters, Supervisory Archivist

This session will present an overview of the Field Book Project, a collaborative initiative of the National Museum of Natural History and the Smithsonian Institution Archives (SIA) to improve access to these important records of biodiversity research. The core goal of the Field Book Project is to provide one online location for locating field book content. Combining traditional archival description with library and museum approaches to description and content delivery, the Field Book Project draws from existing practices and standards. This session will present an overview of the project, its description and content delivery approach, and next steps.

Tammy currently supervises SIA's Archives and Information Management (AIM) and Reference Teams. AIM is responsible for identifying, acquiring and describing records and papers relating to Smithsonian history that form the Archives' collections. The Reference Team provides researchers information about and access to SIA holdings. Tammy has worked on numerous collections management and cataloging and description projects since joining SIA in 1995.

Tammy holds a B.A. in History from Bethel College (KS) and an M.A. in American Studies from Purdue University.

Computer History Museum Digital Repository Development
Paula Jabloner

As Director of Collections at the Computer History Museum where she has worked since 2004, Paula oversees the work of the collections staff engaged in managing, preserving, and providing online and in-house access to the Museum's collections. Prior to joining the Computer History Museum, Paula managed a number of archival collections and projects, including serving as Project Director for Silicon Valley History Online

She received a Master of Information and Library Studies from the University of Michigan and a Bachelor of Arts in History from the University of Massachusetts, Amherst. Paula is a member of a number of professional organizations including the Society of American Archivists and the Society of California Archivists.

Nominations and election of Co-Chair

Call for New Steering Committee members

2013 Program Committee Representative: TBD

2013 Session Ideas

New Business from the Floor

Adjourn, 5:15

Our chief concern is to ensure that the STHC Roundtable reflects the interests of its participants. We welcome all suggestions relating to the above topics or concerning any other issues members

might like to see addressed at our meetings. Please do not hesitate to get in touch with either of us:

John Rees
National Library of Medicine

301-496-8953
Fax: 301-402-7034
reesj@mail.nlm.nih.gov

Melanie Mueller
American Institute of Physics

301-209-3177
Fax: 301-209-0882
mmueller@aip.org

Around and About Archives

Physician Price Fixing in 19th Century Virginia: An Online Exhibit

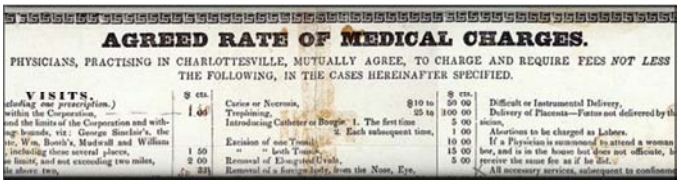
Joan Echtenkamp Klein
University of Virginia

http://blog.hsl.virginia.edu/feebill/

What would you pay for a house visit from a doctor whose office was within a mile of where you lived? How about a dollar with one prescription thrown in for good measure? Or maybe you need your tonsils out. Fifteen dollars will do it. Have a broken arm? Ten dollars will take care of setting it, unless it is a compound fracture and then it would be twice as much. A dollar will cover the extraction of a tooth.

Where are these prices being offered and who are the practitioners? The place is Charlottesville, Virginia, and the twelve doctors making such offers are some of the most respected men in town and include faculty members of the School of Medicine at the University of Virginia. Actually, all the teachers in the School of Medicine, a grand total of four, have agreed to these medical charges because the year is 1848.

What can we learn about medical practice in the mid-nineteenth century by examining the document, generally called a fee bill, which is the inspiration for this exhibit? What was like to live in central Virginia in this time period and who were the men who signed the bill? Visit the newest web exhibit (http://blog.hsl.virginia.edu/feebill/) from the Claude Moore Health Sciences Library Historical Collections and find out. The exhibit features an essay on physician fee bills by Todd L. Savitt, Ph.D.



AGREED RATE OF MEDICAL CHARGES.

PHYSICIANS, PRACTISING IN CHARLOTTESVILLE, MUTUALLY AGREE, TO CHARGE AND REQUIRE FEES *NOT LESS* THE FOLLOWING, IN THE CASES HEREINAFTER SPECIFIED.

VISITS.

Agreed Rate of Medical Charges, 1848.

Children's Hospital Boston Archives Unveils History Wall Exhibit

Sheila Spalding
Children's Hospital Boston

The Children's Hospital Boston Archives is pleased to announce the completion of the Children's Hospital History Wall. Unveiled at a ceremony last September, this museum-quality exhibit is a tribute to the hospital's dramatic and colorful 142 years. The interactive exhibit features archival film footage and photos, a timeline of significant events in the hospital's history and in-depth panels on major discoveries at Children's including the culturing of the polio virus and the development of chemotherapy, surgery, and the largest pediatric research enterprise in the world.

Speakers at the dedication ceremony included Sandra Fenwick, President of Children's Hospital; Dr. Hardy Hendren III, Chief-of-Surgery Emeritus; Dr. Mark Rockoff, Associate Anesthesiologist-in-Chief and Chairman of the Archives Committee; and special guest Lorraine Sweeney Nicoli. Mrs. Nicoli ushered in the field of pediatric cardiac surgery in 1938 when she became the first patient to undergo

repair for a patent ductus arteriosis (PDA), a hole in the heart, at age seven.

The Children's Hospital History Wall is free to the public and can be viewed near the Patient Entertainment Center in the main building at 300 Longwood Avenue during regular business hours. For more information, visit the Children's Hospital Boston Archives online at www.childrenshospital.org/archives .

New Online RCPE Archive Catalogue Reveals Historical Medical Treasures

Alison Scott
Royal College of Physicians of Edinburgh

The Royal College of Physicians of Edinburgh (RCPE) Sibbald Library and Archive has tens of thousands of manuscripts and letters within its collections which it has previously not had the resources to catalogue. As a result of a recent ongoing project funded by the Wellcome Trust Research Resources Scheme, over 8,000 items have been catalogued in order to open up access to this material. The majority of the collections consist of lecture notes and research by some of the most eminent doctors of their day. These demonstrate developments in medical theory and practice from 1700 onwards in one of the most important centers for medical education in the world. Some fascinating items discovered during this work include a draft post mortem report on Napoleon Bonaparte (written on St. Helena on the day he died), correspondence with the Marquis de Lafayette, and notes regarding the conduct of Dr. Robert Knox (of Burke and Hare notoriety).

The Archives catalogue can be searched at: http://www.rcpe.ac.uk/library/find/index.php

For the full press release, please see: http://www.rcpe.ac.uk/press-releases/2012/new-online-archive-catalogue.php

An early X-Ray taken by J.W. Gifford

MGH Museum Opens in Boston

Jeffrey Mifflin
Massachusetts General Hospital

The Massachusetts General Hospital (MGH) announces the opening of the Paul S. Russell Museum of Medical History and Innovation at the corner of Cambridge and North Grove Streets in Boston. The museum is intended to function as a portal between the hospital and the Boston community, showcasing historic as well as modern milestones in medicine, surgery, and clinical practice. Interactive video displays share space with historical objects, tracing the progress of health care over the course of two centuries. The museum's rooftop garden is an open-air retreat, featuring carefully selected plantings and sweeping views of Boston's Beacon Hill neighborhood. The museum is free and open to the public, Monday - Friday, 9:00 AM - 5:00 PM. The Russell Museum inherits art and artifacts previously cared for by the MGH Archives and Special Collections; the exact relationship between the museum and the archives has not yet been determined.

Processing Grants Available at AIP

Joe Anderson
American Institute of Physics

The History Programs of the American Institute of Physics offer annual grants to process and describe archival collections in our fields of interest. This year the **deadline for applications is August 15, 2012.** The grants are intended to make accessible records and papers that document the history of modern physics and allied fields (astronomy, geophysics, optics, etc.). Grants may be up to $10,000 each and can be used to cover direct expenses connected with preserving, inventorying, arranging, describing, or cataloging appropriate collections. All archival repositories in the U.S. and abroad may apply, including archives at universities, corporations, historical societies, and government agencies. Individuals are not eligible to apply.

The AIP History Programs' mission is to help preserve and make known the history of modern physics/allied sciences, and the grant program is intended to help support significant work to make original sources in these fields accessible to researchers. Preference will be given to medium-size or larger projects for which the grant will be matched by the parent organization or other funding sources. For grant guidelines check the Center's website at http://www.aip.org/history/grntgde.htm

The 2011 recipients are the Fermi National Accelerator Laboratory Archives, $10,000, to process the John D. Linsley Papers; Huntington Library, $9,450, to process the Alan Sandage Papers; and Lowell Observatory Archives, $9,600, to process the William A. Baum Papers. See http://aip.org/history/grants_previous.html for a list of earlier recipients.

Recently Opened Radiology Collections at the Center for the History of Medicine

Meghan Bannon
Harvard University

The Center for the History of Medicine at the Countway Library at Harvard Medical School is pleased to announce the opening of five manuscript collections. The collections were part of a series of processing projects focusing on radiology and were funded by the Countway Library's Lloyd E. Hawes Fund for Radiology.

Lauriston Sale Taylor Papers: Taylor, a radiation physicist, was the founder and President of the United States Advisory Committee on X-Ray and Radium Protection (later the National Council on Radiation Protection and Measurements), and Associate Director of the National Bureau of Standards from 1962 to 1965. His research focused on ionizing radiation and radiation protection standards.

Felix Fleischner Papers: Fleischner was Clinical Professor of Radiology at Harvard Medical School and the first full-time radiologist and Head of the Department at Beth Israel Hospital.

Morris Simon Papers: Simon, Professor of Radiology at Harvard Medical School and Radiologist-in-Chief at Beth Israel Hospital (now Beth Israel Deaconess Medical Center) from 1963 to 1970, was also the inventor of several medical devices, including the Simon Nitinol Filter, which is used to trap and dislodge blood clots.

Merrill Clary Sosman Papers: Sosman was Clinical Professor of Radiology at Harvard Medical School, Boston, Massachusetts (1948-1956), and Roentgenologist-in-Chief at Peter Bent Brigham Hospital, Boston (1922-1956). Sosman was a leader in diagnosis by x-ray and significantly contributed to the establishment of a Department of Radiology at Harvard Medical School.

Fleischner Society Records: The Fleischner Society was a thoracic radiology society founded in 1969 and named for Felix Fleischner.

To celebrate the opening of these important Center collections in the history of radiology, the Center hosted an event entitled, "Beneath the Surface: The Development and Cultural Impact of Radiology," held on March 1, 2012. Lectures focused on the history of radiology, including the development of the X-ray, the pioneering "radiology martyrs," and radiology's pervasive influence on visual culture. The video of the event can now be viewed online at: https://cms.www.countway.harvard.edu/wp/?p=5394.

For information regarding access to these collections, please contact the Public Services Staff at chm@hms.harvard.edu . Electronic finding aids are available on Harvard's OASIS web site at http://oasis.lib.harvard.edu .

Conferences, Meetings, and Workshops

The Science, Technology, and Health Care Roundtable will be meeting on Wednesday, August 8, 2012 from 3:15 - 5:15 p.m. in the San Diego Hilton Bayfront, Sapphire O/P. STHC will host [three presentations: the first by Chip Calhoun of the American Institute of Physics; the second by Tammy Peters of the Smithsonian Institution Archives; and the third by Paula Jabloner of the Computer History Museum. For the full agenda see "Message from the Co-Chairs".

Pre-Conference Tours/Open Houses:

For information on pre-conference tours see: http://www2.archivists.org/conference/2012/san-diego/repository-tours

For the full SAA program, please see the following: http://www2.archivists.org/conference/2012/san-diego

STHC-Themed Programs:

Please be sure to read the abstracts for other sessions, because we might have missed some.

STHC Roundtable Meeting
3:15 – 5:15 p.m., Wednesday, August 8th
Sapphire O/P

101. Bitstreams Beyond Borders: The Value of Digital Forensics to Archivists
10:00 a.m. - 11:30 a.m., Thursday, August 9
Sapphire EI

107. From Hidden Collection to International Incident: The John Cutler Papers and the Guatemala Syphilis Experiments
10 a.m. – 11:30 a.m., Thursday, August 9
Sapphire 410

403. Beyond Documents: The Archivist's Role in Research Data Curation
10:00 a.m. - 11:30 a.m., Friday, August 10
Sapphire GH

609. Creating an International Consortium: The Atomic Bomb Casualty Commission
12:30 PM - 1:30 PM, Saturday, August 11
Indigo D

Articles

Balancing Volume and Value: Appraising the Records of Big Science

Laura O'Hara
SLAC Archives and History Office

Originally presented at Society of California Archivists Annual General Meeting, April 27-28, 2012, Ventura, CA

SLAC National Accelerator Laboratory (SLAC) is a U.S. Department of Energy (DOE) physics laboratory operated by Stanford University. Established fifty years ago as a particle physics research center, SLAC is now a multipurpose laboratory for astrophysics, photon science, accelerator and particle physics research. SLAC has a history of reinventing itself, building on former success and repurposing its technology. The Archives ensures that the laboratory's history is identified, collected, preserved, and made accessible to the SLAC and Stanford communities, to researchers, and to the public.

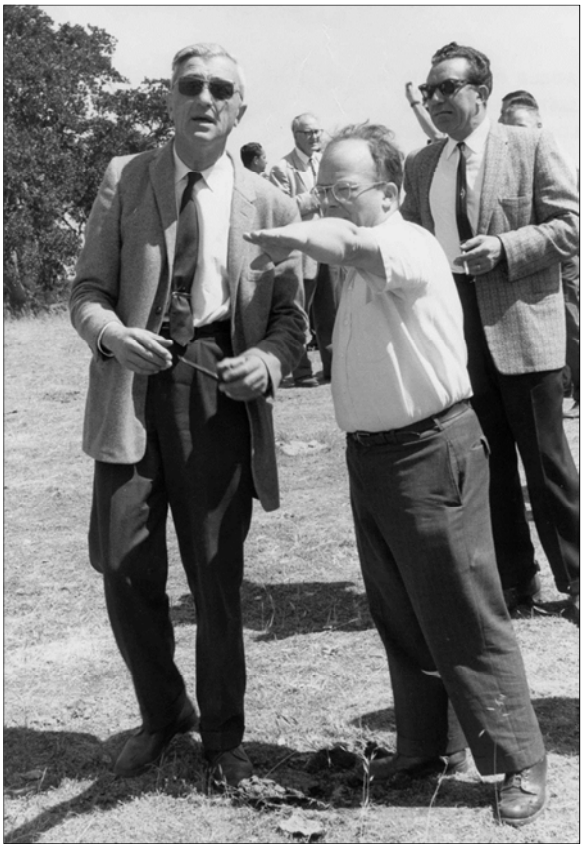
Big Science really took hold post-World War II, followed shortly by the growth of new fields of historical study including the history of science and technology. At roughly the same time, archives and archival appraisal underwent a significant evolution with T.R. Schellenberg, who argued that archivists

should be active agents who select what will be preserved rather than passive accepters of what someone else chooses to send to the archives. The movement to document science and the development of archival appraisal grew up together. While I would love to go into all sorts of historical details—the state of science archives (or lack thereof); watershed conferences of historians of science and scientists concerned about their legacy; a study by the American Institute of Physics (AIP) of recordkeeping at DOE laboratories; the evolution of the DOE Research & Development Records Schedule; and the creation of the SLAC Archives—I will instead focus on the practical application of the Research and Development (R&D) Schedule in real-life situations and other approaches to the types of records found in voluminous modern collections.

The SLAC Archives is a Stanford University coordinate archives holding faculty and pre-contract papers. We also hold records subject to U.S. federal regulations, including DOE and National Archives and Records Administration (NARA) regulations. So we use the federal records schedules and can send scheduled material to NARA, but we also hold Stanford records which NARA does not take. The DOE Records Management Order, which conveys its records requirements, is not a part of the Stanford-DOE contract at this time, so we only use the DOE R&D Records Schedule and other federal records schedules as best practice. Stanford has no formal records requirements.

The R&D schedule is our most important tool for appraisal. Largely based on a study and recommendations from the AIP, the schedule recognizes that not everything should be retained, that there should be periods of reevaluation, and that the stakeholders should be involved. The schedule goes into much more detail, but it boils down to this: experiments or projects are evaluated for level of importance. Level One projects achieve national or international distinction such as a Nobel Prize; those records are permanent. Level Two projects are usually first of a kind or hold implications for the future; their records are retained for twenty-five years. Generally, if a project is going to attain national or international distinction it will do so in those twenty-five years and the records can be reappraised as Level One. Level Three projects are defined by not being Level One or Level Two and their records are retained for ten years with the belief that ten years is enough time for any graduate student to complete his or her thesis.

In addition to the records schedules, we have other tools for getting the goods—some are physical, some are more of an attitude. One is our network of formal and informal contacts of people who hold the records or who can influence the people who do. Another is our knowledge of the organization—it is necessary to know about the person, department, or experiment whose records you are appraising and where they fit in the organization and mission. With this knowledge, we check our existing holdings to see what we already have and what our gaps are in a big-picture way. Armed with this, we can make an appraisal call which starts with a visit to learn what the records creators have and provide them with information about what we do or don't want. Our web page What Should You Keep/What Can You Throw Away? is always a great relief to the records creators as they generally believe the archivists will want to save everything. Depending on what they have, we provide supplies or return to pack the material ourselves.



W.K.H. Panofsky, first director of SLAC, and Felix Bloch at SLAC site dedication, 8/10/1962. Courtesy SLAC National Accelerator Laboratory, Archives and History Office, Muffley Collection.

One of our final tools is the disposal dossier for material that has been appraised as Level Two or Three. A disposal dossier includes the appraisal memo and an inventory. The appraisal memo is

signed off by the experiment spokesperson, thereby providing approval to proceed with disposal. The first few times we went through this process, we weren't sure of the reception we'd get. Imagine going to someone and saying "we've reviewed your experiment and records and, though the experiment was a big part of your life, the records just aren't worthy of being saved," but everyone approached has been more than fine with the appraisal to dispose.

Appraisal is very much a continuum in our workflow. The initial appraisal occurs when we make our house call at which we learn what they have and redirect some of it to records management. Then during packing we may do some appraisal on the fly, at the very gross level. There's a trade-off on whether appraisal time is spent in the packing step or when the material reaches our space. Once the material is in our space we accession and describe it, which often includes inventorying and rehousing; this is the point where we can appraise at a large-grain level, removing handfuls of material, redirecting material to records management, or noting where material duplicates or overlaps existing holdings, and so on.

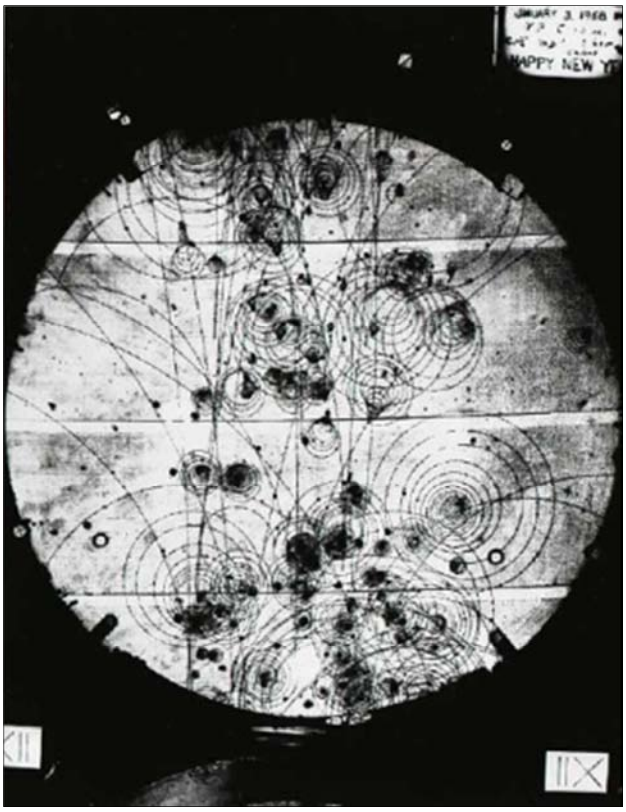
We also appraise during reference. This is because of our triage approach to processing. The triage approach acknowledges that not all records are of equal importance. Everything we receive gets a basic, minimal record in the collections database. Most of our accessions have now progressed to the intermediate level of processing, which includes a comprehensive folder list and an expanded database description. The third level, what we all know as full processing culminating in a formal finding aid, we have only done for a very select few collections. This triage approach allows us to perform some level of reference with all of our material almost immediately and then any further processing is reference-driven. It also means that as we use an accession for reference we continue to appraise, process, and describe. Only once something has reach the final level of processing is it no longer under scrutiny: full processing is the final appraisal.

To illustrate appraisal on the ground at SLAC I present two recent, very different house calls: one to a department that was moving, the other to a scientist who was retiring. The department, Radiation Physics, was moving into a smaller space from a building that they had occupied from almost the beginning of SLAC. They had decades of accumulated material. The history of SLAC's Radiation Physics department is a slice of the history of the field of health physics. The department had everything from the beginning including personal collections of articles, departmental memos and reports, publications, and radiological surveys. They are meticulous records creators (and savers to a fault, though not so meticulous in this respect). We had to appraise material in several different rooms, offices, workspaces, and storage spaces. There was internal order in some places, but also cabinets of "old stuff" (and every time I thought we saw the light at the end of the tunnel, they would find another cabinet or three). When we first met with them, they thought we would say everything had to be kept. But by applying the records schedules we were able to divert certain records to records management because, while they have long retention periods, they are not permanent. A check of our holdings found that we had very little in the way of previous accessions from this group, but there was a written history of the department so we knew who the key players were. This packing trip is a great example of appraisal on the fly as we had limited time to clear out the space, but also limited space at our end to receive it, so we were motivated to whittle the volume down before transfer. As we packed we could immediately recognize whole shelves that were not of archival value, but never a whole cabinet. There were many personal collections of articles from non-SLAC sources and publications which we left behind.

Sometimes we had to consider materials a little more closely—for example, a shelf that at first glance appeared to be simply numbered Environment, Safety and Health (ES&H) publications. We do not need those: we already have them from more authentic sources. But a closer examination showed that they were the revision copies from a former leader in the department. So when a copy was clean, we could leave it, but when a copy was marked up with what were clearly his notes, we packed it. These may still be determined to be non-permanent in future handling, but on the fly, they indicate some value. In the end we took a quite a bit. There's probably chaff, but there are also some important gems which will be revealed as we continue to appraise and describe. After all, this is an old department and important to the history of SLAC and of the field of health physics.

The retiring scientist on the other hand was practically a dream client. His career also spanned decades of SLAC history. A check of our existing holdings showed that we already had some material from his early years and we had a brief

biography. When he contacted us, he already had a folder list of his files with his own disposition appraisal. His categories were: leave with his department; take home; send to the Archives; send to Radiation Physics (he had been chair of the Radiation Safety Committee); dispose; or shred. When he called Radiation Physics, he learned that they were already in the process of transferring their material to the Archives. We were now being offered everything he had earmarked for them as well as what he already planned to send us. As I packed what he was ready to give us, it was good stuff: committee minutes; high-level reports; and planning documents for big, important (probably Level One) experiments. He also shared his disposition list with the Archives. In reviewing it, we agreed with his appraisal of what should be shredded -- that was all personnel files -- but what he had earmarked for disposal was of great interest to us. When asked about those files, he said "I didn't think you'd want those, they're just full of fiddly details," but if we wanted them, we were free to take them. We did and, though this was a small accession, I have already used it several times for reference requests, even before it made it onto the shelf.



Bubble chamber event, 1/3/1968. Courtesy SLAC National Accelerator Laboratory, Archives and History Office, Zawojski Collection.

Appraising an individual's records, particularly an individual with a long career, also reveals what until recently was a weakness of the R&D schedule. Previously the schedule separated federal records—those created by DOE staff as government employees—from their overlapping professional and personal records created in their parallel careers as university faculty, leaders in their professions, consultants and advisors. They may start as post-docs and later become principal investigators or spokespersons, perhaps administrators, and then respected statesmen involved in national and international science policy. So, John Stoner, the archivist at Lawrence Berkeley Lab, with Jean Deken at SLAC and Lee Michael at the National Renewable Energy Laboratory, proposed a revision to the schedule. Rather than try to separate such an intertwined mix of federal records, academic papers, and professional papers, the R&D Records Schedule was revised in 2008 to include a Researchers Collection item. Through this scheduling innovation, DOE and the national labs are succeeding in preserving the integrity of the records of individual researchers whose efforts span multiple roles, projects, and experiments.

Resources

AIP Study of Multi-Institutional Collaborations. Phase I: High-Energy Physics, American Institute of Physics, 1992

AIP Study of Multi-Institutional Collaborations. Final Report, American Institute of Physics, May 2001

Joe Anderson, American Institute of Physics, Niels Bohr Library and Archives with the Center for History of Physics, "Pragmatic Appraisal: Collecting the Records of Modern Science," Opening presentation, 5th Annual Scientific Archives Conference, Rio de Janeiro, Brazil, 9/27/2011

Department of Energy, Research and Development Records Schedule (N1-434-96-9, N1-434-07-01, and N1-434-08-02) Revision 2, June 2008, http://energy.gov/sites/prod/files/cioprod/documents/RD_revised.pdf

SLAC Archives & History Office, "What Should You Keep/What Can You Throw Away?" http://www.slac.stanford.edu/history/archnonarch.shtml

The iPhone and the Tule Shoe

Leigh Johnsen
San Joaquin County Historical Society and Museum

I like my iPhone. I think of it as a Swiss Army Knife on steroids. It can make telephone calls, send text messages, keep my personal calendar, and enable me to write notes to myself, play games, listen to music, find my way on city streets, and surf the Internet. Sometimes I gaze at this pocket-size marvel with amazement.

But my iPhone has a problem: its brilliance often blinds me to less dazzling technologies from earlier times, which can also be impressive. So it is with the tule shoe, a low-tech invention that has won my appreciation over the four years I've worked as archivist of city, county, and private records at the San Joaquin County Historical Museum, in Northern California.

The tule shoe helped change the course of California agriculture. Its story began in 1851, only a year after California entered the Union. Two years had passed since the Gold Rush, and savvy settlers had figured out by then that the value of California's rich soil--especially in the state's Central Valley--outstripped the worth of its gold.

The richest soil was in the Sacramento-San Joaquin Delta. The Delta covers a broad swath of land at the inner reaches of the San Francisco Bay. As with other deltas, its value arose in part from the presence of rich topsoil washed downriver for millennia. But even more important was a thick layer of decayed vegetation that built up in the shallow waters where Sacramento and San Joaquin Rivers meet the Bay.

This layer of vegetation came largely from tule reeds, and they dominated the landscape. Illustrations from the last half of the nineteenth century show a swampy marshland for miles on end. Channels snaked their way through the reeds, thus enabling ships to make their way between San Francisco and the burgeoning city of Stockton, which stood at the edge of the Delta. So abundant were tules at that time that in its early days Stockton bore the name "Tuleburgh," the German equivalent of "Tule Town."

In 1851, the state of California opened the marshland for sale and settlement. Records show that the land went quickly. But the new owners faced obstacles. Levees were required to keep the land from flooding, and settlers needed to cut down and burn living vegetation. Then they needed to plow and prepare for planting. The product was a rich yet soft and spongy form of highly organic soil that caught fire easily, swirled into suffocating clouds of dust on dry days, could develop cracks up to a quarter of a mile long and twenty feet deep, and tended to swallow wagons and horses.

Tule Shoe. Gene Celli Collection, San Joaquin County Historical Society and Museum.

Tule shoes guarded horses from sinking into the soil. In appearance and theory, they resembled snowshoes. Imagine for a moment an ordinary horseshoe. Now place an iron ring around it, attach spokes radiating outward, and connect them to the circle. The result distributed the horse's weight, prevented it from sinking, and enabled it to plow, pull wagons, and do other tasks needed to tame

the marshy land. According to some sources, workers were known to attach wooden boards to the bottom, which distributed the horse's weight even more widely.

Students of California's agricultural history credit Chinese laborers with inventing tule shoes. Whatever their origins, they were pivotal for the Delta's development. Thanks to them, land values rose, water was managed more closely, crop selection widened, government watchfulness rose, and new shipping routes developed. In addition, men and women made fortunes, laborers arrived, and ethnic diversity widened. The success of the tule shoe also helped turn the uniqueness of the Sacramento-San Joaquin Delta into a breeding ground for other inventions, the most notable being the track-laying tractor that eventually became the Caterpillar.

Appreciation of the tule shoe may seem out of place in the twenty-first century. But its story sends me a message that my iPhone can't. The story of the tule shoe informs me about incremental change over time and of the potential for humans to overcome obstacles. It reminds me of men and women through millennia who met challenges they faced with creative, yet often modest solutions and ended up triggering an avalanche of changes they never foresaw.

Personally, I think that's rather nifty.

The ABCC Collection in the Texas Medical Center Library as a Nuclear Age Memory

Philip Montgomery
Texas Medical Center Library

***NOTES**: The following text is an edited and shortened version of a keynote speech given in Tokyo at Gakushuin University on November 20, 2011. The symposium titled "Memory and Records of the Nuclear Age" was sponsored by Gakushuin University and the Gakushuin Graduate Course in Archival Science, and funded by a grant from the Japan Society for the Promotion of Science. The symposium focused on the records of the Atomic Bomb Casualty Commission (ABCC), which was created in 1946 to study the after effects of the atomic bombings of Hiroshima and Nagasaki. There are three major sources of information about the ABCC. Those sources are the National Academy of Sciences archives, which houses the official ABCC records; the Radiation Effects Research Foundation in Hiroshima, which is closed to the public; and the Texas Medical Center Library, which houses personal papers and is the largest collection open to the public. Smaller scattered collections are also located in Japan.*

As an aside, my sponsors in Japan made very clear to me that the invitation to speak was a response to the Fukushima Daiichi nuclear disaster. Although the audience at the symposium was limited to about 100 people, representatives from a Fukushima grass-roots group and TEPCO, the nuclear plant owners, physicists, geneticists, and graduate archival students were present. It was a passionate symposium and quite emotional.

I am here to talk about the Atomic Bomb Casualty Commission (ABCC) collection in the Texas Medical Center Library. Also, I am here to learn from my Japanese colleagues and to see for myself the work that has been done to preserve and make accessible to the public the papers and records related to the bombings of Hiroshima and Nagasaki and the aftermath. In truth, the documents in the Atomic Bomb Casualty Commission collection belong to everyone in the world, not just to scholars in the United States or in Japan. As the archivist for the Texas Medical Center Library, I am merely the caretaker of the ABCC papers. I am responsible for preserving these papers so future generations will have the information they need to understand the past.

The Japanese people know better than any nation in the world the devastating results of radiation, whether those effects are caused by nuclear war or by natural catastrophes. I have enormous admiration for the resiliency, courage, and steadfastness that the people of Japan have shown in the face of great adversity. The rest of the world has much to learn from you.

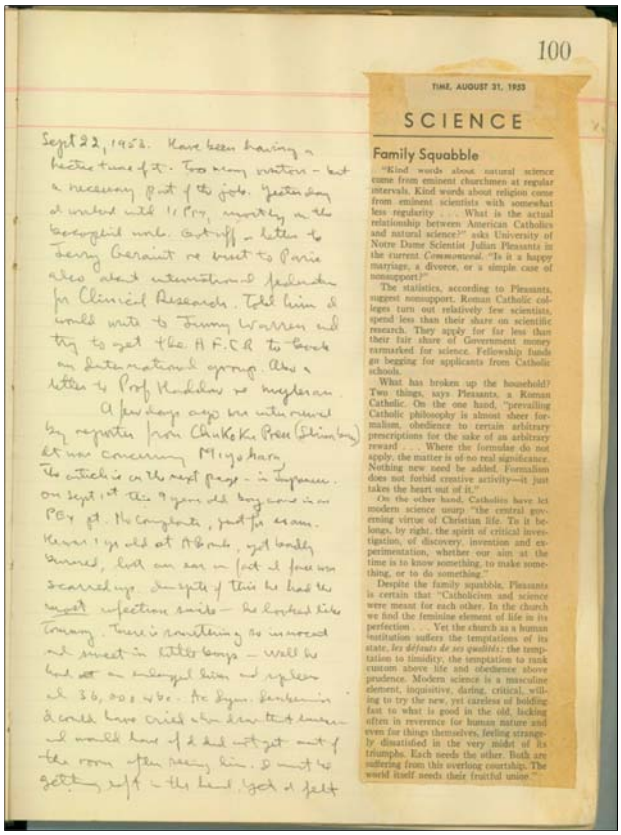
Today, I am here to talk about memory and records in a nuclear era. I have been asking myself, "what are records?" and "what is memory?" To answer that question, I have fallen back on an archivist's definition. Richard Pearce-Moses, in his 2005 book *A Glossary of Archival & Records Terminology,* says a record is "data or information in a fixed form that is

created or received in the course of individual or institutional activity and set aside (preserved) as evidence of that activity for future reference." Pearce-Moses says memory is "the knowledge of events, people, places, and other things of the past; an individual's knowledge of the past, or a specific recollection of something in the past." In her novel *Animal Dreams*, Barbara Kingsolver writes, "memory is a complicated thing, a relative to truth, but not its twin."

Memory is complicated and often distorted over time. For example, my mother grew up on a farm near a small community called Star City, Indiana. She recalls her childhood during the Great Depression of the 1930s in the United States as a time of great hardship and privation. She does not share happy memories of those days. I was raised on stories of privation about the Great Depression. I also heard the stories of my mother's older sister: my aunt grew up on the same farm during the same period and remembers those years as a hard time, but she also had fond and happy memories of those days. I honor the memories of my mother and my aunt. But which is closer to the truth? Where does the truth lie? For my family, the truth is shrouded in oral stories. My family has no written memories or farm or banking records to illuminate the truth.

For the history of the ABCC, there are numerous written memories and official records. The ABCC papers located at the Texas Medical Center Library are important because so many of the collections contain "grey literature," documents outside of normal publication channels, and papers that reveal personal viewpoints rather than the stories expressed in official ABCC records. These non-official records and personal papers provide insight into the culture and the bureaucracy of the ABCC not found in official records. They show how the forces of culture, society, and politics, both internal and external, affected the ABCC.

Since the 1980s, the staff of the McGovern Historical Collections has solicited and preserved the documents of the Atomic Bomb Casualty Commission. This Collection comprises manuscripts and other records donated by former ABCC members throughout the United States. There are about twenty collections with nearly 200 cubic feet of records. The individual collections offer personal insight into the workings of the ABCC. The entire collection offers a comprehensive view of the attitudes, goals, and activities of the Commission from the late 1940s through its evolution into the Radiation Effects Research Foundation. These records provide more than just published materials. These papers augment the official records and provide the personal perspectives of the ABCC staff. The ABCC materials consist of correspondence; memos; official publications; diaries; scrapbooks; ephemera such as brochures and items printed or created for special and short-term purposes; and objects associated with daily business. The collections also include graphic materials, such as photographs, art prints, drawings, film, and video.


100

TIME, AUGUST 31, 1953

SCIENCE

Family Squabble

"Kind words about natural science come from eminent churchmen at regular intervals. Kind words about religion come from eminent scientists with somewhat less regularity . . . What is the actual relationship between American Catholics and natural science?" asks University of Notre Dame Scientist Julian Pleasants in the current *Commonweal*. "Is it a happy marriage, a divorce, or a simple case of nonsupport?"

The statistics, according to Pleasants, suggest nonsupport. Roman Catholic colleges turn out relatively few scientists, spend less than their share on scientific research. They apply for far less than their fair share of Government money earmarked for science. Fellowship funds go begging for applicants from Catholic schools.

What has broken up the household? Two things, says Pleasants, a Roman Catholic. On the one hand, "prevailing Catholic philosophy is almost sheer formalism, obedience to certain arbitrary prescriptions for the sake of an arbitrary reward . . . Where the formulae do not apply, the matter is of no real significance. Nothing new need be added. Formalism does not forbid creative activity—it just takes the heart out of it."

On the other hand, Catholics have let modern science usurp "the central governing virtue of Christian life. To it belongs, by right, the spirit of critical investigation, of discovery, invention and experimentation, whether our aim at the time is to know something, to make something, or to do something."

Despite the family squabble, Pleasants is certain that "Catholicism and science were meant for each other. In the church we find the feminine element of life in its perfection . . . Yet the church as a human institution suffers the temptations of its state, *les défauts de ses qualités*: the temptation to timidity, the temptation to rank custom above life and obedience above prudence. Modern science is a masculine element, inquisitive, daring, critical, willing to try the new, yet careless of holding fast to what is good in the old, lacking often in reverence for human nature and even for things themselves, feeling strangely dissatisfied in the very midst of its triumphs. Each needs the other. Both are suffering from this overlong courtship. The world itself needs their fruitful union."

Journal of Dr. William Moloney, MS 73 William C. Moloney, MD, papers; 1952-1954.

I want to talk briefly about some of the interesting items in the collection. First is the journal of Dr. William C. Moloney, a hematologist who served with the ABCC from 1952 to 1954. During that time, he wrote a journal about his experiences. The journal is preserved in the collection called MS 73 William C. Moloney, MD, papers; 1952-1954.

Dr. Moloney's journal appears to have been written with no intention of publication. His diary sheds light on his feelings. His words are heartfelt expressions of his thoughts as he touched pen to paper. Most of his entries are mundane and describe social activities such as tennis matches, family issues, visitors to the Hijiyama facility outside Hiroshima,

and concerns about the quality of the work of the ABCC. On page 18 in the entry for January 9, 1953, Dr. Moloney says

> We had a dull day at ABCC; only 2 out of 9 patients scheduled showed up. ... The attrition is really serious and Grant [Taylor] is beginning to worry about it. He cut out my comments on it in the semiannual report -- said it belonged under biostatistics. Hope it will not get completely emasculated. I am afraid Lowell [Woodbury] will tend to minimize it.

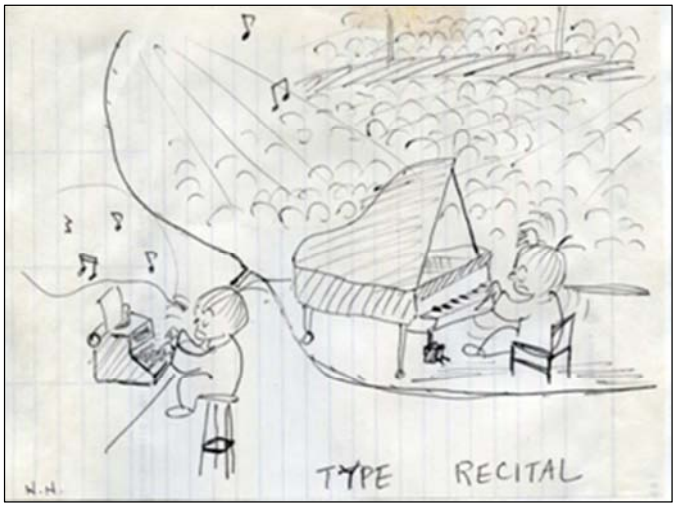
For a historian, this small entry sheds some light on office politics behind the official documents. Personal papers serve to reveal new interpretations of official records.

Dr. Moloney also talks about the frustration of being a doctor and not being allowed to treat patients and not having a cure for leukemia. Dr. Moloney was known for his life-long research to understand leukemia. On pages 100 to 102 of his journal in an entry dated September 22, 1953, Dr. Moloney writes with emotion about battling the disease of leukemia and about his reason for coming to Japan:

> On Sept. 1st this 9 year old boy came in.... No complaints, just for exam. He was 1 year old at A Bomb, got badly burned, lost an ear ... face was scarred up. Despite ... this he had a [beautiful] smile -- - he looked like Tommy. [Moloney's son.] Well, he had an enlarged liver and spleen of 36,000 wbc [white blood cells]. ... Leukemia. I could have cried ... and would have if I did not get out of the room after seeing him. ... Yes I felt so frustrated and impotent -- but I knew why I came here. All the time now I know this was right, that I can and am doing something worthwhile. ... If I can add one little detail to the knowledge that will [defeat] this disease, I'll feel justified forever.

Personal papers can reveal a great deal about life in the ABCC. Look at the cartoons of Dr. Akio Awa, one of the world's foremost cytogeneticists, who worked at the ABCC during the tenure of Dr. Howard B. Hamilton. Dr. Hamilton was the chief of clinical laboratories for the Commission from 1956 to 1975.

Dr. Awa spent nearly 30 years performing radiation cytogenetic population studies of the effects of the atomic bomb on the people of Hiroshima. He was also a cartoonist who was known for salting staff meetings with his cartoons and using humor to communicate his feelings. Dr. Hamilton collected the discarded cartoons from 1967 to 1970 and pasted them into four notebooks, which he donated to the Texas Medical Center archives. Dr. Awa's cartoons are insightful, clever, well drawn and a great commentary on life in a bureaucracy. It is impossible to capture Dr. Awa's humor in words, but his cartoons enrich our understanding of staff life in the ABCC in a way that "official" records can never do.



Cartoon by Dr. Akio Awa

The ABCC records in Houston are owned and housed by the Texas Medical Center Library, but in truth, the ABCC records are as much yours as mine. I believe nuclear tragedies will continue to happen, despite the best efforts of mankind to control the atom. Uncontrolled nuclear radiation, whether caused by accident, disaster, or war, remains a problem for the entire world. The records compiled by the Commission are more important than ever before. Today all of us can take steps to ensure that the knowledge acquired through the ABCC survives far into the future. Wisdom is acquired through poor judgment: the future must be informed by the past.

The knowledge gained from the bombing of Hiroshima and Nagasaki must be available to the world. To make the information available, we need cooperation between archives, institutions, peoples and cultures. It is our destiny to do our best to preserve these papers and make them accessible to the world. As the archivist for the Texas Medical Center Library, I can speak for the library when I say that our commitment to preserving the papers and records of the ABCC is strong. We are committed to making the papers accessible to the world. It is my honor to oversee this priceless collection and to stand before you and share an overview of the ABCC papers in my care.

(Continued)

SOURCES

Pearce-Moses, Richard. (2005). *A Glossary of Archival & Records Terminology*. The Society of American Archivists. Chicago, Il.
Kingsolver, Barbara. (October 13, 2009). *Animal Dreams*. HarperCollins e-books. HarperCollins Publishers.

Reitz, Joan M. (2010). "Grey literature." *Online Dictionary of Library and Information Science*. Retrieved on Oct. 3, 2011 from http://www.abc-clio.com/ODLIS/searchODLIS.aspx.

The William C. Moloney papers; MS 73; John P. McGovern Historical Collections and Research Center, Houston Academy of Medicine-Texas Medical Center Library.

The Howard B. Hamilton, MD, papers; MS 66; John P. McGovern Historical Collections and Research Center, Houston Academy of Medicine-Texas Medical Center Library.

Motion Picture Digitization at the National Library of Medicine

John P. Rees
National Library of Medicine

Introduction

In September 2010, the National Library of Medicine (NLM) launched *Digital Collections*, a Fedora-based repository which allows rich access to and preservation of digital content from NLM's History of Medicine Division (HMD). Included in this initial launch were eleven digitized motion pictures, now grown to seventy titles. This article examines the content modeling, selection, workflows, and software used to produce repository content and access.

Background

HMD's Historic Audiovisual Program (HAV) collects, preserves, and makes available motion pictures, films, and videos produced or created up to 1970; NLM's General Collection is responsible for post-1970 content. HAV currently manages over 5,000 titles, much of it in the public domain. Providing online content to portions of these films has been a long-time goal, but technology, policies, and manpower all conspired against these efforts. In 2009, development work began in earnest to build a "digital repository" and motion pictures were to be one of the pilot formats.

With a clear mandate and additional resources at our disposal there were still many hurdles to overcome. Most important among these is the Department of Health and Human Service's strict Section 508 Accessibility policy covering online video: one hundred percent accurate captioning must accompany any online video content. Digitizing video content is the easy part; creating transcripts and captions and providing access to the complete package was much more challenging. Luckily, our repository developers already had a side project to develop a video player with search capabilities for the NLM Director that we were able to incorporate into our technology stack.

Content Selection

As a member of NLM's Digital Repository Implementation Group, I serve on the architecture and metadata subgroups, lead the preservation subgroup, and have responsibility for guiding video content development. For our pilot test, I selected the eleven oldest cataloged motion pictures that were in the public domain. These were mainly World War II public health films produced by the U.S. Navy and Army. Why not start with the most difficult case study first, right? 1940s audio fidelity was likely to be poor and image resolution likely to be grainy, but on the plus side the content was copyright-free and was some of our most curatorially engaging. Who wouldn't like to see a young, uncredited Gene Kelly playing the part of a shell-shocked sailor in *Combat Fatigue Irritability* or learn good battlefield dental hygiene techniques in *Dental Health*?

Curated content currently drives our selection model, rather than large-scale or mass digitization methodologies. Titles are drawn from existing

subject guides or other thematic content and pass a content, quality, audience, and restrictions review. For example, our two recent projects were drawn from an in-house DVD education module: Public Health Films Go to War and Guide to Tropical Disease Motion Pictures (TDM). Two titles proposed for TDM were rejected on grounds laid out by our Access to Personally Identifiable Health Information policy. Both films were created by their donor and we had the creator's copyright permission; however, each film contained scenes diagnosing specific diseases using full-body images of individuals manifesting these diseases. We do not have the technology to obscure their faces and have no releases from these individuals. These two titles will also be removed from general circulation and restriction notices applied to their catalog records.

Content Modeling

No matter the software platform, we would need to decide what kinds of objects were needed to satisfy a variety of use cases:

- Native playback of video and captions
- Playback of video and/or captions on non-native devices, e.g. local desktop applications, handheld devices
- Export/download of master files and selected user access formats for external re-use
- Data Harvesting/data mining
- Preservation (?)

Fedora is not exactly a repository, but is instead a service architecture

> for storing, managing, and accessing digital content in the form of *digital objects* inspired by the Kahn and Wilensky[1] Framework. Fedora defines a set of abstractions for expressing digital objects, asserting relationships among digital objects, and linking "behaviors" (i.e., services) to digital objects.[2]

Fedora's architecture is articulated through "content models" – a data model or a profile for a particular "genre" of digital object. The basic components of a Fedora digital object are:

- **PID:** A persistent, unique identifier for the object.
- **Object Properties:** A set of system-defined descriptive properties that are necessary to manage and track the object in the repository.
- **Datastream(s):** The element in a Fedora digital object that represents a content item (the bits).[3]

Working off some common assumptions made for a concomitant book digitization pilot project, our video content model reads like an overwhelming laundry list of files:

- MARC XML (derived from ILS; definitive descriptive metadata store)
- DMDINDEX (local XML descriptive metadata transformed from MARC XML, consumed by the digital repository's SOLR index for search and display of content)
- Dublin Core (used for OAI-PMH harvesting)
- Master MPEG2 (.mpeg)
- Access MOV Quicktime derivative (.mov)
- Access Windows Media derivative (.wmv)
- Full size access H.264 derivative (.m4v)
- Video Player, compressed, H.264 derivative (.m4v)
- Clip H.264 derivative (.m4v)
- iPhone H.264 derivative (.m4v)
- Quicktime SMIL caption file (.smil)
- Quicktime SMIL transcript file (.txt)
- Adobe Flash DFXP caption file (.xml)
- Magpie caption master file (.magpie)
- Plain text transcript file (.txt)
- Large poster image (.jpg)
- Medium poster image (.jpg)
- Small poster image (.jpg)

Digitization and Derivative Production

You may have noticed my question mark alongside the Preservation use case bullet point above—at this point I'm hesitant to characterize these digitization activities "preservation- oriented." It is more closely defined as an "access project with some preservation considerations." All our work uses standard commercial off-the-shelf or open source tools.

[1] http://www.cnri.reston.va.us/k-w.html, A Framework for Distributed Digital Object Services, accessed June 20, 2012.

[2] http://www.fedora-commons.org/about. Fedora Commons website accessed June 15, 2012.

[3] https://wiki.duraspace.org/display/FEDORA34/Fedora+Digital+Object+Model. Fedora version 3.4 documentation accessed June 15, 2012

Best practices in film preservation still have not coalesced around any preservation-worthy file formats, although the National Archives creates uncompressed video in an AVI (Audio Video Interleaved) wrapper for its Standard Definition video preservation masters. Their use case statement, "this file is at an appropriate information capture level to substitute for the original record if the original record copy is damaged, destroyed or not retained," demonstrates a fundamental preservation condition—the copy is suitable for substitution of the original.[4] Alternatively, the Library of Congress' Packard Campus for Audio-Visual Conservation has chosen losslessly-compressed JPEG 2000 encoded video wrapped in MXF as its preservation format.[5] Oftentimes the exact file type decision is dependent on an institution's software abilities. For instance at NARA, both their manual and SAMMA robotic video transfer tools natively capture AVI.

At NLM we exert little control over creating our "master" digital format, MPEG2. We have no sophisticated production software or hardware. A DVD is derived from the current analog preservation best practice BetacamSP copy, which itself is not the highest-quality imaging as compared to the original film. Most DVD technology wraps video codecs in the MPEG2 standard, which is a relatively non-proprietary, high bit-rate, yet lossy compression format; the video is very high quality, yet the lossiness is a negative criterion for a preservation use case.

We use the standard desktop software Roxio to rip MPEG2 from the DVD; however, there are no user controls to specify any output parameters. So while we create a fairly high-end copy and define it as our 'Master' (some may look twice at the file sizes--the MPEG2's file size is generally on the order of 2GB per 30 minutes of runtime), our use case more closely aligns with NARA's Video Median Capture – SD profile, that of a reproduction master copy but not suitable as a substitute for the original. NARA also specifies some very specific MPEG2 encoding parameters where we cannot.[6]

Quickmedia Converter[7], an open-source/freeware video transcoding software, is used to convert the MPEG2 to full-size (640x480 resolution) Windows Media (.wmv) and Quicktime (.mov) derivatives. Quickmedia Converter supports both asynchronous and synchronous operations: batch mode iterating through a list of files or processing one file at a time. Conversion rates generally run in the 1:2 range (half a video's runtime). File sizes for both derivatives are generally one-third that of the MPEG2 source file. We have also used VLC Player[8] to convert the actual DVD source Video_TS content when ripping with Roxio failed, and Roxio can also convert a Video_TS input source.

We then switch to Quicktime Pro[9] to create the H.264 derivatives and frame still thumbnail images (.jpg). Quick Media supports H.264 export, but without the ability to apply some specific conversion parameters needed to optimize our Video Player With Search user experience. The free Quicktime software already contains the Pro tool set, invoked by purchasing the $30 upgrade license key from the Apple Store: simply update the Preferences options and enter your registration key.

The full-size Quicktime .mov files are the input source used to create the H.264 transcodes. The .mov file is a bit smaller than the .wmv with equal image quality. The H.264 transcode takes significantly longer and is computationally expensive, so any file size savings affects the transcode time. It is preferable to use a dedicated CPU, as running a Quicktime job in the background while multi-tasking will be difficult without a machine with robust RAM (over 2GB). Transcoding the audio signal appears to be the principal reason for the increased conversion time.

Our first pass with Quicktime produces a high quality, full resolution (640x480) size H.264 codec wrapped as an .m4v file. It is about half again smaller than the size of the .mov source file. It also has high quality AAC stereo at 44.100 KHz. The total data rate of the output file averages 1.25 Mbps. One could easily output the same codec as an MP4, as it is basically interchangeable with .m4v and it would playback just easily across a variety of devices. However, we sometimes use the .m4v in MAGpie to create captions and MAGpie cannot consume an .mp4.

[4] http://www.archives.gov/preservation/products/products/vid-p1.html Video Maximum Manual Capture - SD, accessed June 18, 2012.

[5] http://www.digitalpreservation.gov/formats/fdd/fdd000059.shtml Digital Formats Web site, accessed June 18, 2012.

[6] http://www.archives.gov/preservation/products/products/vid-r1.html Video Median Capture – SD, accessed June 18, 2012.

[7] http://www.cocoonsoftware.com/, accessed June 20, 2012.

[8] http://www.videolan.org/vlc/index.html, accessed June 20, 2012.

[9] http://www.apple.com/quicktime/extending/, accessed June 20, 2012.

In the same transcode operation we also select the option to create an iPhone derivative (~10% smaller than the .mov file) and the frame still poster image (640x480 large size). This first pass can take at least as long as the film's runtime, if not a bit longer.

In our second pass with Quicktime, we use the iPhone derivative as our input to create the .m4v derivative used by the digital repository's Video Player with Search. This is a medium quality, medium resolution (480x360) H.264 codec with medium quality AAC stereo at 22.050 Khz. This file has a much lower video data rate and half the audio data rate of the H.264 full size file in order to progressively load quickly for the web user. Video data rate is either 300 Kbps or 500 Kbps. If in the reviewer's opinion the 300 Kbps rate produces too much pixelation, we will re-transcode at 500 Kbps; the file size and download time significantly increases, but the higher image quality video still loads within a reasonable time for low-bandwidth consumers. Audio data rate is 55-60 Kbps and the total data rate is about 375 or 575 Kbps.

Our third pass with Quicktime is to generate a short video clip that may or may not be used for our Coollris wall on the digital repository homepage. We use the iPhone derivative again to select a thirty-second portion of the film and export an .m4v with the same medium resolution encoding parameters used for the Video Player derivative.

Lastly, we use Quicktime to edit the large poster image into two smaller thumbnails also used in digital repository displays, one 320x240 and the smallest at 160x120.

Transcriptions and Captioning

We assumed transcriptions and captioning would be performed in-house rather than contracted out. The digital repository's principal architects Ed Luczak and Doron Shalvi had conducted a detailed experiment on transcription techniques as part of their Video Player With Search development. Automated audio extraction software exists, such as Adobe Soundbooth, which might shorten our transcription production work. However, the experiments showed this to be unlikely as source speech recognition alone will not yield high accuracy (mean of 51%) due to poor fidelity, background music, or ambient noise common to 1940's era film. Accuracy improves to ~70% with high-fidelity audio as represented by NIH's modern digital videocast productions. At 50% accuracy editing the output takes more effort than manual transcription, but editing 70% accurate output would be cost effective.

We also experimented with Dragon Naturally Speaking software as an "echo recognition" technique. A transcriber listens to several seconds of audio, pauses the video, and repeats phrases into Dragon to generate a transcript. This technique worked fairly efficiently from a cost/benefit perspective, but does not scale very well with multiple users if budget restrictions prevent purchasing multiple copies of Dragon. Echo recognition workflow was fine for our eleven pilot videos, but not so well with eight transcribers working on fifty videos for the Tropical Disease Motion Picture project. In the end we dropped the Dragon technique and simply typed phrases using a text editor.

A quality assurance (QA) reviewer (someone other than the original transcriber) then reviews the transcript for errors while watching the film in real time. Marking up a hard copy and then editing the electronic document worked better than editing the text file in real time.

The transcript is created with certain parameters that enable it to be imported into the open source MAGpie captioning software.[10] MAGpie reads each paragraph break as a distinct caption area. In MAGpie the captioner then plays back the source video and adds start times for each caption—one keystroke performs this action. We then export several files from MAGpie: an official transcript as ASCII text; SMIL (*Synchronized Multimedia Integration Language*)[11] caption files; and a DFXP (Distribution Format Exchange Profile)[12] caption file used by our Video Player With Search. These processes are time consuming and present a significant hurdle for any online film digitization project. However, the access benefits, above and apart from satisfying the Section 508 mandate, clearly outweigh the production cost factors. In the end it is likely more cost effective to outsource transcription and captioning production. Our experiences produced a wide range of time/motion metrics. For our pilot, the fastest rate that any of us (relatively well-practiced producers) could transcribe/caption was 5:1 per hour of runtime: three hours to transcribe one hour of

[10] http://ncam.wgbh.org/invent_build/web_multimedia/tools-guidelines/magpie, accessed June 20, 2012.

[11] http://www.w3.org/TR/SMIL2/, accessed June 20, 2012

[12] http://www.w3.org/TR/2009/CR-ttaf1-dfxp-20090924/, accessed June 20, 2012

runtime; one hour QA review (error-free transcript, no editing required); and one hour captioning. This metric dropped significantly for the fifty video Tropical Disease Motion Pictures (TDM) project: the rate was closer to 8:1 per hour of runtime. Of course there were other variables, such as film content type; TDM had many multiple-hour interview films that naturally have more speech than a narrated training film. Staff motivation was another significant factor.

Future Directions

We anticipate changes to our derivative offerings and production workflows as our knowledge, skills, and budgets evolve. Extending access through mediums such as YouTube is a potential avenue now that NLM has its own official channel; YouTube has also developed some interesting auto-captioning services with transcript upload. We are already exploring Sorenson Squeeze as a server-side transcoding solution in use by other units within NLM. There are some exciting development projects in the Fedora context at the Rock and Roll Hall of Fame, the Variations on Video project, and at WGBH Broadcasting.

Conclusion

This article concisely describes our current policies, workflows, and software used to develop online access to digitized motion picture content. Online film service is still a complex prospect within our specific political and cultural milieu. Other institutions have taken other tacks, producing greater numbers of online film titles, but NLM is committed to providing value-added products to its customers. High standards require deep commitment to programs that we hope best serve our public mission.

About the Authors

Leigh Johnsen holds a doctorate in American history from the University of California, Riverside, and a master's degree in library science from the University of Illinois at Urbana-Champaign. He has worked as a historian and archivist at Claremont Graduate University, the University of California at Davis, and the University of the Pacific. He currently holds the position of archivist and librarian at the San Joaquin County Historical Society and Museum, in Lodi, California.

Philip Montgomery is the archivist and assistant director of the John P. McGovern Historical Collections and Research Center at the Texas Medical Center Library in Houston. The McGovern Historical Collections is dedicated to preserving the history of medicine with a focus on Houston and Texas. He worked as an archivist and special collections librarian at Rice University's Woodson Research Center before taking his current position in 2010. He received his MLIS from the University of North Texas in Denton in 2005 and became a certified archivist by the Academy of Certified Archivists in 2006. He is currently working with archivists in the U.S. and Japan to create a consortium of archives dedicated to preserving and improving accessibility to Atomic Bomb Casualty Commission records.

Laura O'Hara has worked at the SLAC Archives and History Office for 15 years. Initially she was an itinerant archivist at Stanford combining part-time gigs at SLAC, Stanford Law School, and Stanford School of Medicine (only 2 at any given time), but she has been a full-time SLACer for almost 5 years now. Previously she worked at the Western Jewish History Center of the Magnes Museum and at the Bank of America Corporate Archives. She is active in SCA serving on and chairing a variety of committees and has served as membership director and is a past president.

John Rees is the Archivist and Digital Resources Manager for NLM's History of Medicine Division, Archives and Modern Manuscripts Program. John has led the program under a variety of titles since 2000. His position reflects the dual nature of his existing duties managing the manuscripts program and working on NLM's intra-library digital repository architecture development team. John received his MA in Southern Studies in 1992 from the University of Mississippi and his MLIS in 1997 from the University of Texas-Austin with a concentration in Archival Enterprise.

SAA Science, Technology Health Care Roundtable
Steering Committee Members (2011-2012)

John P. Rees (Co-Chair), National Library of Medicine, Bethesda, MD

Melanie Mueller (Co-Chair), Niels Bohr Library & Archives, American Institute of Physics, College Park, MD

Polina E. Ilieva (Web Liaison), Library and Center for Knowledge Management, University of California, San Francisco

Liz Phillips (Newsletter Editor), Special Collections, University of California, Davis

R. Joseph Anderson, American Institute of Physics, College Park, MD

Danielle Castronovo (Immediate Past Co-Chair), California Academy of Sciences, San Francisco, CA

Janice F. Goldblum, National Academy of Sciences, Washington, DC

Joan Echtenkamp Klein, Claude Moore Health Sciences Library, University of Virginia Health System, Charlottesville, VA

Jodi Koste, Tomkins-McCaw Library, Medical College, Virginia Commonwealth University, Richmond, VA

Nancy McCall, The Alan Mason Chesney Medical Archives, The Johns Hopkins Medical Institutions, Baltimore, MD

Charlene Nichols, Jet Propulsion Laboratory, Pasadena, CA

Stephen E. Novak, Augustus C. Long Health Sciences Library, Columbia University, New York, NY

Tim L. Pennycuff, Lister Hill Library of the Health Sciences, University of Alabama at Birmingham, Birmingham, AL

Tom Rosko, Institute Archives and Special Collections, Massachusetts Institute of Technology, Cambridge, MA

Paige Lindsey Smith, MDiv., MLIS, Birmingham, AL

Paul Theerman, National Library of Medicine, Bethesda, MD

Judith A. Wiener (Immediate Past Co-chair), Prior Health Sciences Library, The Ohio State University, Columbus, OH

Archival Elements is produced annually each summer. It is the official newsletter of the Science, Technology, and Health Care Roundtable of the Society of American Archivists. Please consider submitting an article to *Archival Elements*. For more information on submitting information or an article, please contact Liz Phillips (UC Davis).

Join the Society of American Archivists

Founded in 1936, the Society of American Archivists is North America's oldest and largest national archival professional association. SAA's mission is to serve the education and information needs of more than 5,500 individual and institutional members and to provide leadership to ensure the identification, preservation, and use of records of historical value. Membership includes those serving in government agencies, academic institutions, historical societies, businesses, museums, libraries, religious organizations, professional associations, and numerous other institutions in more than 60 countries. Through its publications, workshops, annual conference, and programs, SAA provides a means for contact, communication, and cooperation among archivists and archival institutions.

Society of American Archivists
17 North State Street
Suite 1425
Chicago, IL 60602-3315

(312) 606-0722
(or toll-free 866-722-7858)

info@archivists.org
www.archivists.org