

**Spartan Archive:
An Electronic Records Archive at Michigan State University
NHPRC Project #RE-10025-10**

NHPRC Interim Narrative Progress Report, October 1, 2011–March 31, 2012

This report addresses activities undertaken on “Spartan Archive: An Electronic Records Archive at Michigan State University,” a National Historical Publications and Records Commission (NHPRC)-funded project. The Michigan State University Archives & Historical Collections (UAHC) is using a proof-of-concept approach to build and test an archival solution for the long-term access and preservation of four large electronic records series produced by MSU’s Office of the Registrar (RO). The original grant proposal and supporting documents for the project may be found at <http://spartanarchive.wordpress.com/project-documentation/>.

Project Administration

- **Office of the Registrar staffing change**
As Damion Gadson has left MSU, Kristin Schuette, associate registrar for technology, will represent the RO’s interests in the project.

Design, Development, and Implementation of Preservation Environment

- **Implementation of Archivists’ Toolkit (AT)**
Data cleanup of UAHC’s Archivists’ Toolkit (AT) implementation continues under the direction of Ed Busch, electronic records archivist. This includes conversion of subject and name headings into valid Library of Congress and local headings. Nearly all of the imported and manually entered accessions have been converted. UAHC staff and part-time workers are helping with the cleanup, which will likely run through 2012.

UAHC has been able to use AT for collection management, but the conversion process, which includes automating the transfer of hundreds of paper-based accession records, has proved time-consuming. The AT implementation was modified for records management, and UAHC has used it to create accession records for the RO files and other digital material.

AT content has been modified to work with MARC and EAD exports. All finding aids in MS Word format have been imported into AT and are fully searchable. Word-based box and shelf lists have also been imported— including lists of yearbooks, photograph collections, atlases, and more. Preliminary workflow has been identified for moving AT collection-level MARC catalog records into the MSU Libraries OPAC and OCLC.

Usage has revealed the need for further improvements and modifications. Rich Burgis, Spartan Archive project information technologist, has created a number of custom reports, and Busch continues to explore options for automating the creation of EAD finding aids.

- **Repository development**

Lisa Schmidt, electronic records archivist/project manager, drafted a system requirements document for Michigan State University's Spartan Archive electronic records preservation environment. The document, which covers the requirements for building Spartan Archive per the grant as well as for expanding it to include other MSU born digital records with permanent retention, is available on the project website at

http://spartanarchive.files.wordpress.com/2011/04/spartan-archive-system-requirements-document_5.pdf.

Burgis continues to refine the ingest process for Spartan Archive. In addition, he has developed an access interface prototype at <http://www.spartanarchive.msu.edu>. Burgis has also drafted documentation defining the Fedora objects in Spartan Archive, as well as an updated ingest process and graphical depictions of the ingest and access processes. These documents may be found on the project website at <http://spartanarchive.wordpress.com/project-deliverables/>.

Busch took the lead on creating a framework for a test plan based on the system requirements. Schmidt and Burgis will help Busch finalize and carry out the plan during Q2 2012. Burgis is also drafting systems documentation for Spartan Archive. Test plan and systems documentation drafts are available on the Spartan Archive project website at

<http://spartanarchive.wordpress.com/project-deliverables/>.

In early March, Schmidt invited three MSU reference librarians to the Archives to test the usability of the access interface prototype. Burgis made changes to the interface based on the librarians' feedback. An MSU web designer will create a more user-friendly interface to Spartan Archive during Q2 2012.

- **Data extraction and transmission**

Burgis continued to work with the RO to test and revise the database models and XML schemas for all four data series. These latest revisions are available on the project website at

<http://spartanarchive.wordpress.com/project-deliverables/>. More than 160 files were transferred in the first set of transmissions, 20 years of data.

After facing some challenges in building a sustainable infrastructure through various proposed partnerships with MSU's IT department, UAHC has secured file space for transfers and storage. Two terabytes of storage space has been allocated for use by UAHC to date, with the expectation that more space will be added over time.

Workflows, policies, and procedures

- **Metadata definitions**

Schmidt drafted metadata recommendations, including modified Dublin Core and parts of PREMIS, for sample collections and instances of RO data files. MODS was examined for use in Spartan Archive, but it was determined that although MODS provides more descriptive capabilities than Dublin Core, its complexity outweighs its usefulness. The project team decided that Dublin Core will meet the needs of Spartan Archive and be easier to use. Burgis is working with the recommendations, which will be refined as needed. In addition, PREMIS event metadata will be more comprehensively defined. The draft metadata recommendations

may be found on the project website at

<http://spartanarchive.files.wordpress.com/2011/04/metadata.pdf>.

- **Electronic records processing workflow**

In Q4 2011, a workstation dedicated to processing electronic records was procured and set up. The electronic records processing workflow task force of UAHC director/Spartan Archive project director Cynthia Ghering, Schmidt, Busch, and Burgis met regularly, and made the decision to separate accessioning from the processing of electronic records, as is done frequently with analog (paper) records. To that end, Schmidt and Busch developed a workflow for accessioning records to UAHC's "digital shelf" space. They also began developing a processing methodology. The draft workflow is available on the project website at <http://spartanarchive.wordpress.com/project-deliverables/>.

Schmidt experimented with adding disk imaging to the workflow. Disk images are exact copies of all original files on their originating media, and it was thought that such a copy might be preserved alongside original, preservation, and access copies of archived records. Disk imaging is not practical for use in the routine accessioning of institutional records, however, as it would double storage requirements and add more steps to the workflow. Consideration will be given to incorporating disk imaging into the workflow for processing manuscript and other special electronic records collections when the need arises.

In terms of normalizing files for preservation, Schmidt proposed that campus offices be educated on converting Word documents to PDF/A format rather than PDF before transfer to UAHC. She prepared a conversion guide for MS Office for Windows 2007 users based on a similar reference on the University of Michigan's Deep Blue repository website. UAHC is currently piloting this process with MSU administrative offices. The conversion guide is available at <http://spartanarchive.files.wordpress.com/2011/04/how-to-create-pdf-a-from-word.pdf>.

The 2 TB of server space allocated for Spartan Archive is also available for general accessioning of electronic records until Fedora content models may be developed for them. This will include digital photos and video from MSU's University Relations as well as digital records from other units. Schmidt and Busch are in the process of moving the digital content currently on the electronic records or "digital shelf" drive to this new space designated as the "Digital Vault," as well as the contents of optical media discs currently stored in a file drawer. Keeping the affordances of the new storage space in mind, Busch and Schmidt will continue developing accessioning guidelines and processes.

iRODS Training

- **2012 iRODS workshop**

At the request of CIC archivists who attended the technical MSU iRODS workshop with Reagan Moore in April 2011, UAHC planned a second workshop focused on the practical aspects of implementing iRODS from an archival perspective. Jill Sexton, systems librarian at the University of North Carolina and project manager for the Carolina Digital Repository (CDR), was engaged to lead a half-day workshop, "Archivists, Technologists, and iRODS" on April 21, immediately following the annual Midwest Archives Conference meeting in Grand Rapids. Topics include the use of iRODS and Fedora in the CDR as well as communication strategies for archivists working with technology teams.

All MAC attendees and MSU technical staff, as well as CIC archivists and librarians, were invited to the workshop. Registration stands at 40 participants from a wide variety of institutions, including corporations and religious orders as well as higher education and state and local governments. This workshop takes the place of the iRODS consultant visit articulated in the original Spartan Archive grant proposal. Information about the workshop may be found at <http://msuirods2012.wordpress.com/>.

Project Reporting

- **Best Practices Exchange 2011 Presentation**

On October 20, Ghering delivered the presentation “Spartan Archive: Archiving Institutional Data” at the Best Practices Exchange conference. The presentation was well received, generating good discussion with attending representatives of NARA’s Electronic Records Archive (ERA), Tufts University, the North Carolina State Archives, and other organizations building electronic records archives. This discussion led directly to Ghering’s decision to develop systems requirements documentation for Spartan Archive. The presentation is available on the project website at http://spartanarchive.files.wordpress.com/2011/05/bpe_2011_cynthia_ghering.pdf.

- **Administrative Data Users Community (ADUC) presentation**

On February 20, Schmidt presented on Spartan Archive to the MSU Administrative Data Users Community (ADUC), a group of MSU staff “with a shared interest in the creation and use of institutional data and associated information systems.” This “Archiving Digital Records” presentation included a records management and e-mail segment covered by Tom Wellman, university records manager, as well as Schmidt’s segment on preservation of records with permanent retention. About 60 members of the MSU community attended. The presentation may be found at http://spartanarchive.files.wordpress.com/2011/05/aduc_archiving_digital_records_02-20-12-pptx.pdf.

- **Midwest Archives Conference (MAC) 2012 Session Proposals**

UAHC’s two session proposals on electronic record preservation were accepted for the 2012 Midwest Archives Conference (MAC) meeting. “Digital Preservation Comes of Age: Reports from the Field” will feature Schmidt presenting on the Spartan Archive project; Joanne Kaczmarek, of the University of Illinois will chair the session, and other participants include Mary Davison of American Girl, Mark Myers of the Kentucky Department for Libraries & Archives, and Caryn Wojcik of the State of Michigan.

A web archiving session will feature Busch, David McCartney of the University of Iowa, Benn Joseph of Northwestern University, and Olga Virakhovskaya of the Bentley Historical Library at the University of Michigan; Michael Shallcross of the Bentley will be moderator. This session will explore difficulties, obstacles, strategies, and successes experienced during the initiation and implementation of service-based web archiving programs.

- **Society of American Archivists (SAA) 2012 Round Table Proposal**

Schmidt's proposal to deliver a presentation on Spartan Archive to the Records Management Round Table during the SAA annual meeting was accepted. This presentation will showcase the new access interface and may include a demo.

Other Related Activities

- **Best Practices Exchange, October 20-22**

Ghering and Busch attended the Best Practices Exchange conference in Lexington, Kentucky. Themed "Distilling Digital Collections," the conference provided a good knowledge exchange on what state and local government institutions are doing to tackle various issues with electronic records. Presentations by Mike Wash and Meg Phillips, CIO and Electronics Records Lifecycle Manager of NARA respectively, helped to confirm that UAHC is taking the right approach in building the Spartan Archive. Wash emphasized the need for a comprehensive solution architecture, and Phillips noted that NARA should have first focused on digital preservation and building the repository rather than on business processes and workflows. Phillips also drew attention to NARA's Online Public Access (OPA) Catalog, which could be a good model for the Spartan Archive access interface. Veronica Martzahl of Tufts University presented on the TAPER project; UAHC may want to eventually use TAPER or a similar technology to automate records transmittals and integrate with Spartan Archive.

Busch also attended the annual Archive-It partner meeting the day before the regular conference began. Participants presented on lessons learned, shared experiences with other users, provided input on needed features, and learned about new developments planned for 2012.

- **ARL SPEC Kit: Managing born digital materials in ARL libraries**

Ghering and Schmidt are part of an ARL SPEC Kit team chartered to develop and analyze the results of a survey that addresses the challenges of managing born digital materials at ARL libraries. Spearheaded by Tim Pyatt, head of the special collections library at Pennsylvania State University, the team is comprised of members from the University of Michigan and Duke University as well as Michigan State and Penn State. This ARL SPEC Kit supports the overall ARL Strategic Directions, especially "Transforming Research Libraries: Outcomes and Strategies," in exploring best practices for managing digital content. Developed over the winter of 2011-2012, the survey was conducted in March. The team will gather at Penn State in May to analyze the results and complete the report by June.

- **MSU IT reorganization**

MSU's Office of the Vice Provost of Libraries, Computing, and Technology—which includes UAHC—has reorganized its technology units in an effort to provide consistent, better service and positive IT experiences across campus. The intent is "to change the central IT service, planning, and resource management model from separate departmental units with defined boundaries to a set of purpose-focused, highly coordinated, and flexible teams operating like a single unit."

UAHC director Cynthia Ghering is now director of the new Content and Collaboration unit, which includes web development, desktop, and collaborative software development services as well as UAHC. Moving forward, UAHC has access to more technical resources and the greater

unit is well positioned to lead repository development and other technology initiatives at MSU.
For more information, visit lct.msu.edu/itservices.



Cynthia A. Ghering
Project Director



Lisa M. Schmidt
Project Manager