Project Proposal

For more information or guidance please see the host requirements, project examples and project characteristics as found on the NDSR website.

c/o George Coulbourne
ndsr@loc.gov | www.digitalpreservation.gov/ndsr

Date Submitted

2/5/2016

Submitted by:

Jordon Steele

Project Proposal Title

Large-Scale Digital Stewardship: Preserving Johns Hopkins University's Born-Digital Visual History

Project Summary
(provide 2-3 sentences summarizing the project)

Homewood Photography is Johns Hopkins University's professional photography services unit, which maintains a portfolio of Hopkins-related stock photography. This project will transfer a backlog of approximately 15 terabytes of digital images that Homewood Photography has generated since 2004 to the Sheridan Libraries control; ensure that the images are preserved and accessible through the libraries' digital repository; establish a mechanism and workflow for ingesting newly created images on a routine basis; and work with Homewood Photography to develop a metadata schema for describing future digital objects in the unit's digital asset management system. Time permitting, the resident will also use the skills learned in this project and his or her knowledge of the libraries' digital infrastructure to assist the Sheridan Libraries in its Trusted Digital Repository self-assessment.

Specific Goals and Objectives
(identify goals and objectives to be completed by end of residency)

• Goal 1: Select Homewood Photography born-digital images in accordance with University Archives collection development policy

o Objective 1.1: Development of appraisal criteria that aligns with University Archives collecting strategy

o Objective 1.2: Appraisal of about 50 terabytes of images to determine what selection should be permanently preserved by the Archives (expected yield is 10-15 terabytes, or a possible 30-50% reduction)

• Goal 2: Improve the long-term preservation of Homewood Photography's born-digital images

o Objective 2.1: Creation and application of unique identifiers that clearly map the metadata to the images in a machine-actionable way

o Objective 2.2: Migration of in-scope images from CDs, DVDs, servers, and other carrier media to networked storage managed by the Archives for long-term preservation

Timeframe and Deliverables
(provide a quarterly breakdown of assignments and deliverables necessary to complete objectives)

Q1: Archives orientation; meeting with Homewood Photography; training on Archives' born-digital management tools and disk-imaging workflow using test set of photographs; survey of images; appraisal and selection of images; determine preservation requirements for images

Q2: Regular expression training to assist in metadata clean-up; pre-ingest clean-up of metadata associated with images; transfer of selected images to Archives systems; post-ingest metadata clean-up in repository

Q3: Collaborate with Homewood Photography on DAMS-to-Archives workflow; test DAMS-to-Archives workflow; begin developing metadata schema for future photographs

Q4: Finalize DAMS-to-Archives workflow; document workflow; train others (including student workers) on the workflow; finalize metadata schema; close down project; compile project report and present findings to libraries staff, Homewood Photography, and interest groups external to the university; assist Archives in the libraries' Trusted Digital Repository self-assessment (time permitting)

Resources Required
(identify any resources necessary to complete the project)

- One resident
- Collaboration and support from libraries staff, including two mentors
- Storage space to serve as "scratch space" for working with images before permanent ingest
- Desktop computer and workspace
- Access to university computing environment (JHED ID, network privileges, etc.)
- Access to university archives' suite of tools used to ingest, manage and preserve born-digital content
- Support and guidance from mentors in identifying and accessing professional peers, networks and organizations

Context
(provide a narrative statement explaining the project – no more than 500 words)

As the official photographers for Johns Hopkins University, Homewood Photography creates images that provide an unparalleled visual record of the university's history. The Archives has previously accessioned and made available Homewood Photography's analog work, which has proven to be a rich source of historical images for the university community as well as the general public.

Since the mid-2000s, Homewood Photography has been using an exclusively digital workflow. Many of these digital images are stored on optical media which is beginning to fail. While Homewood Photography recognizes the historical importance of these images, they are unable to manage a strategy for long-term preservation and access. Metadata about these images is currently maintained in an Access database. However, these descriptions are inconsistently applied and sometimes incorrect.

Homewood Photography continues to produce digital images at a projected rate of 2TB per year, of which one-third to one-half would likely be appropriate for accessioning into the Archives based on a high-level understanding of the content and how it aligns with the Archives' collection development policy. In order to ensure preservation of these images while preventing the buildup of another backlog, a process needs to be developed for the routine and automatic transfer of these images to the Archives. Homewood Photography's transition to a hosted digital asset management system later this year provides the perfect opportunity for instituting such a process in partnership with the Archives.

A workflow featuring script-managed bulk actions will be necessary in order to accomplish the selection

Required Knowledge and Skills of Resident
(identify requirements necessary to successfully complete the project – technical skills, educational background, specialized experience, etc.)

- Demonstrated knowledge of current web technologies
- General knowledge of digital preservation and born-digital curation principles
- Ability to evaluate and implement a variety of software tools, including command-line tools
- Project management education or experience
- Knowledge of library and archival metadata standards including EAD, DACS and Dublin Core
- Demonstrated organizational, analytical and problem-solving skills
- Strong attention to detail
- Excellent written and verbal communication skills

Preferred Knowledge or Experience of Resident
(identify preferred knowledge or experience – technical skills, educational background, specialized experience, etc.)

- Experience with Archivists' Toolkit or ArchivesSpace
- Experience with forensic imaging hardware and software, such as BitCurator, Archivematica, or AccessFTK
- Experience with scripting languages
- Experience creating or maintaining process documentation and training materials