

September 29, 2000

Institute of Museum and Library Services
Office of Library Services
Program Office, National Leadership Awards
1100 Pennsylvania Avenue, NW, Room 802
Washington, DC 20506
Re: Interim Report

**Preserving Cornell's Digital Image Collections: Implementing an Archival Strategy
3rd Quarter Report, January-June 2000**

Project: 34774

PI: Anne R. Kenney
Department of Preservation and Conservation
Cornell University Library
Ithaca, NY 14853
607-255-6875, ark3@cornell.edu

The goal of Cornell's IMLS-funded project is to plan and implement an archiving strategy for its digital image collection, representing over a third of a terabyte of information. This report summarizes the achievements of the project team during the third six-month period (January-June 2000). The project Web site (<http://www.library.cornell.edu/imls/index.htm>) includes several reports and documents that complement this brief report.

January-June 2000

- Hosted John Price-Wilkin (Director, Digital Imaging Production Service, University of Michigan), the project consultant, at Cornell for a 2-day program that focused on comparing different image delivery systems and particularly the metadata that support the operation of these systems (especially structural metadata). The program, which was composed of several meetings and presentations, explored the relationship between the architecture and the long-term longevity problems and addressed how to successfully mainstream digital imaging and preservation initiatives to move them from a project to program mode.
- Prepared a preservation metadata proposal that identifies the important information elements needed to be recorded for the future digital imaging projects. Prospective project managers will be required to provide this information, which will be maintained centrally on a Web-based system. The goal of this database is to provide profiling information to support managing Cornell University Library's digital image collections (digital preservation administration). This approach tries to address the difficulty associated with gathering base-level information about our digital image collections.

- Evaluated the currently used TIFF2GIF on-the-fly conversion program and adjusted its settings to optimize image quality for access images. Started to explore the licensing issues pertaining to the use of this on-the-fly conversion software.
- Appointed a working group (Digital Preservation Policy Working Group) to develop a mainstreaming strategy for ensuring the longevity of Cornell's digital image collection. The goal is to create a permanent infrastructure to ensure that the concern for the longevity of our image collection will continue beyond the completion of this IMLS-funded project. The group represents different functional units of the library and aims to develop a plan to mainstream the implementation of digital preservation policies and procedures throughout the Cornell University Library, which is composed of 19 campus libraries. Building a consensus among Cornell colleagues is essential to ensure that the achievements of this project will have a long-lasting value for the library, rather than being a one-time effort. Co-chaired by Anne Kenney and Oya Rieger, the group plans to prepare the central deposit guidelines for Cornell's digital image materials by December 2000. The objectives of the group include:
 1. Develop requirements for deposit of digital imaging materials
 2. Plan a long-term maintenance and preservation strategy
 3. Identify resource requirements to support effective management and preservation
 4. Devise a strategy to promote the adoption and consistent use of the baseline standards throughout Cornell University Library
- Completed exporting the RDO files that were created by the proprietary XDOD system to Cornell Digital Library format (An RDO file contains information about the structure of an image document, as well as a file location pointer for each page image in that document). This migration process was essential to be able to amalgamate the collections in the SUN system. Also, a risk analysis study was conducted to determine the risks involved during this migration process (See the project Web site, "Risks Involved in Migrating RDO Files."). This risk assessment study was funded by a grant from the Council on Library and Information Resources.
- Started cleaning up the MARC records for two large digital image collections - the Making of America (MOA) and the Core Historical Literature of Agriculture (CHLA). Student employees were hired and trained and permanent employees were assigned to work on this clean-up, standardization project, which will be completed in October 2000. MARC records for the MOA materials were created when the first version of the system was made available. MARC records for CHLA titles were drafted when that system was being developed. Unfortunately, the cataloging of those records were quite dated because the standards for cataloging electronic resources and reproductions have changed considerably in the past five years. The goal is to use these MARC records as the basis for descriptive metadata to be used for long-term access to these digital materials. However, for this to be effective, radical changes needed to be made to the records. Among the changes are:

1. Creating and adding PURLs to the records
2. Verifying that access points were properly coded for automated transfer of descriptive elements from MARC to the DC-based structure
3. Modification of MARC record formats to reflect current standards for cataloging
4. Improving notes present in the record so they accurately reflect the history of the digitization projects

Future Tasks (July-December 2000)

The project, which will end on March 31, 2001, seeks to achieve the following goals by December 2000:

- Explore the potential of PNG file format in replacing GIF in image delivery and pros and cons of TIFF2PNG conversion utilities and potential of PNG file format in replacing GIF in image delivery. Identify and test a conversion software for creating on-the-fly access images in PNG file format to compare the functionality and image quality between TIFF2GIF and TIFF2PNG on-the-fly conversion software. Write a report on the findings of the TIF2GIF and TIF2PNG evaluation study (due October 2000).
- Complete the MARC record clean-up project, which started in May 2000 (due October 2000).
- Write a report on the storage requirements assessment that documents the storage system evaluation process (due October 2000).
- Design and implement the Web-based form that will support the recording and maintenance of a variety of preservation metadata in a centralized fashion (due December 2000).
- Present the draft depository guidelines prepared by the Digital Preservation Policy Working Group for discussion and approval by the Cornell University Library community (December 2000).
- Project costs associated with maintaining the digital image collection for the next ten years (December 2000).