



## NORTHEAST DOCUMENT CONSERVATION CENTER

### **Surveying Digital Collections: A “Trends” Report**

By Tom Clareson, PALINET, and Liz Bishoff, The Bishoff Group

The “Digital Preservation Survey Project” was developed by the Northeast Document Conservation Center (NEDCC), with funding from a 2004 Institute of Museum and Library Services (IMLS) National Leadership Grant. The purpose of the project was to develop a Digital Preservation Readiness Assessment that could be used to determine the state of digital preservation readiness of museums, libraries, or other cultural heritage institutions. The assessments were conducted by a team of consultants experienced in digital libraries, preservation and conservation, and digital preservation.

The IMLS project began with a Web-based survey, conducted in spring 2005, that assessed current digital preservation practices of cultural heritage institutions. In July 2005, a colloquium of museum decision-makers, digital library leaders, preservation and digital preservation experts, and digital content users was held to review the survey results.<sup>1</sup> Based on the recommendations of the colloquium attendees, a team of consultants began conducting tests of a “Digital Preservation Readiness Assessment.”

Through June 2006, eight museums and libraries of varying size and focus were surveyed. These surveys went beyond digital preservation to look at all facets of digital collection activity in the institutions, including digital collection administration and management, selection and acquisition, access, preservation, and digital rights issues. The project continued in the first half of 2007 when two organizations with long-term digital preservation programs were surveyed to determine if they had moved further in digital preservation policy and practice. A third institution was interviewed that had outsourced its long-term solution to a not-for-profit long-term digital preservation program. Below are some trends from the eleven survey visits.

1. Mission: Digitization is incorporated into the organizational missions of the two organizations operating long term digital preservation solutions. Other institutions are either in the process of modifying their mission or completely disregard digital initiatives within their mission.

Regardless of the number of years that an institution has been digitizing its collections or acquiring digital content, the majority of the visited sites are still at

---

<sup>1</sup> A report of the Web survey and colloquium results appeared in the February 15, 2006, *RLG Digi-News* e-newsletter.

- the digital *project* stage, not having moved to the *program* stage. An institutional approach to digital initiatives does not exist, nor do the majority of these institutions have coordinated efforts in the area of standards, access, promotion, and preservation of digitized materials.
2. Management: The organizations with the strongest digital programs have built strong awareness and support from their administration and board. More than 50% of the funds for digitization initiatives come from the institution's operating budgets, followed by grant funds. The least likely sources of funding are fees derived from digital collections.
  3. Evaluation: The visited institutions have conducted limited evaluation of the use of the digital collections by their constituents/audiences, including staff. The evaluation that *has* been done focuses more on the interface, screen manipulation, and search techniques, rather than the usefulness of the collection to the user.
  4. Digital Preservation: Except in the case of the organizations that operate long-term preservation solutions and the institution that outsources its preservation solution, even organizations with major digital initiatives have only rudimentary digital preservation programs.
    - For most of the sites visited, digital preservation is on their radar; they “know that they should be paying attention to it.”
    - A general lack of understanding exists of even the most rudimentary aspects of digital preservation.
    - Education in digital preservation is a key component before any of these institutions start “doing” anything.
    - Knowledge of digital preservation activities varies across the organizations, from very limited to sophisticated.
    - Staff members responsible for traditional preservation/conservation are generally not involved in the digital initiatives at their institutions.
    - CDs and DVDs are used in the majority of the projects and seem to be the preferred digital content storage media for the visited sites, with a few institutions beginning to move to networked storage.
    - Refreshing data is a strategy used by a few institutions for maintaining digital data. However, rather long periods of time elapse for checking authenticity of the data.
  5. Quality control of master images is occasionally done at production, but generally not done as part of the maintenance of the digital object. Quality control standards are lax, being seen as too labor intensive and costly. At some institutions, quality control is dependent upon automation or the skill of the photographer/scan technician.
  6. Most of the institutions still believe backing up is digital preservation.

7. Disaster Planning: None of the institutions has incorporated digital content into their disaster plans. None of the institutions has incorporated digital content creation into their existing disaster plans.
8. Documentation: Documentation of workflow and practices in digitization and digital preservation are needed at all of the surveyed institutions, except the contemporary art museum and the two long-term preservation sites.
9. Digital Asset Management Systems: Some staff believe that digital preservation would be accomplished through acquiring a digital asset management system. A general lack of information exists about the relationship between the DAM and a long-term digital preservation solution.
10. Standards: Generally, formal cross-institutional standards in digital imaging and workflow are lacking within these programs. One of the historical societies surveyed, because of its integration of multiple digital programs, has made the most significant progress in uniformity of standards.
11. Born Digital: The issues related to born-digital material are not being addressed by most of the institutions. Museums and libraries are acquiring, through donation, collections that are only available in digital format. They have not established policies or procedures for curating these collections, from their deeds of gift to collection processing, presentation and dissemination, and preservation. The contemporary art museum offers an approach that many art museums are implementing: they regularly “reinstall” performance art, providing a means of interacting with the artist/creator to address preservation issues through a “team approach.” It addressed issues associated with long-term preservation of the instructions for reinstallation, as well as the digital art itself.
12. Digital Survey Questionnaire: All of the institutions have been very positive about the two-step survey process, which includes completing a pre-survey questionnaire prior to the site visit, the visit itself, and a written report that summarizes findings from both the questionnaire and on-site visit.
13. At the two institutions with long-term digital preservation programs surveyed in early 2007, some of the best practices observed include collaborative approaches to digital preservation, strong documentation and policies, addressing the need for financial viability and sustainability of the digital preservation effort, utilization of the OAIS model, and implementation of both dark archives and LOCKSS model open-source solutions.
14. Dark archives are no longer seen as the only option. An increasing number of institutions are looking to offer continuous access, or “light archiving,” for their digital collections.

Evaluation of the program was performed by an independent project consultant. The consultant reviewed the project process (the pre-visit questionnaire, on-site visit, and report), and administered both an evaluation within several week of the final report and a follow-up report several months after the final report. In the evaluation process, key areas studied were usefulness of the project, progress related to expected project outcomes at all institutions, and progress related to the consultants' specific recommendations to each institution.

In both the immediate and later evaluations, every institution reported making progress in developing and implementing digital preservation policies and practices. Additionally, institutions used nearly 50% of the specific recommendations from the reports for direct planning and action.

The evaluation process found that the site visit was the most useful component of the project, mainly because of its interactive nature. Additionally, there were uniformly positive responses on the project's role in improving sensitivity of staff to how the institutions managed their digital collections and to the interdepartmental nature of the discussion with the consultants. The process helped staff articulate their digital preservation needs, and led one institution to start a new Digital Images Group. Many surveyed institutions used the report to "force attention to the issue."

NEDCC has developed reports and written tools related to the survey process. As a result of the project, NEDCC has developed an intensive, policy-based national workshop series on digital preservation, *Stewardship of Digital Assets*, and is also looking at methods to train additional surveyors.

Possible long-term outcomes of this project could include the establishment of documentation and training for digital surveys similar to the IMLS/Heritage Preservation Collection Assessment Program (CAP), and potential funding of additional digital preservation site surveys through the National Endowment for the Humanities' Preservation Assistance Grants or other sources.