July 2014

Digitization, Can it Play a Role in Disaster Preparedness?: Notes from the Field

Gayle Palmer
OCLC Western

Digitization, Can it Play a Role in Disaster Preparedness?:

Notes from the Field

by Gayle Palmer
Digital and Preservation Program Manager,
OCLC Western

Like most librarians and archivists I have listened to recent news stories about natural disasters and their impact on library and archival collections. Collection materials in all formats have been damaged by catastrophic hurricanes on the Gulf Coast and flooding in the Midwest. Recently I have spent a lot of my time developing training material to introduce librarians and archivists to preservation planning for digital collections. Among the topics addressed in the training are business resumption techniques and storage of digital files. Both topics are important to disaster preparedness. Disaster preparedness education is designed to help all of us plan for quick and efficient response to a "disaster event" whether it is caused by disruptive weather or by more mundane events like leaks from construction or remodeling damage.

So, what role can digitization play in disaster preparedness? When this topic was posed to me my first reaction was that digitization is not a preservation technique, rather it is a convenient delivery and access option. However, I came to realize that digital collection items, like all collection materials, must be protected from disaster caused by technical or organizational failure as well as from disruption events. The good news is that principles of disaster preparedness and preservation apply to all collection formats.

The role that digitization does play in disaster preparedness can be condensed into these few topics.

Replication, lots of copies keep stuff safe, is a common rubric of preservation. Replication of analog collection items through digitization can be an effective method to maintain and distribute copies of collection materials. Whether the materials are born digital, or reformatted from paper or analog media collections, the digital versions are easily duplicated as exact copies with no loss of information or degradation of the file. Distribution of digital copies to other socially trusted repositories can help keep collections safe and accessible during disasters. The distributed LOCKSS system at Stanford University (http://www.lockss.org/lockss/How_It_Works) is based on this idea. The act of digitization assumes multiple versions of the digital object will be created to suit particular needs. Thus for every digitized version of a photograph there is at minimum an uncompressed "master" file, a compressed "access" derivative for Web display, and often there is an even more compressed "thumbnail" or reference version. Protection of the "master" archival version of the digital object becomes a primary preservation technique and usually occurs when the digital item is created. This technique is similar to the method used to create archival microfilm versions of collections. The idea among digital collection curators is that derivative versions of a "master" digital file would be used to create "on demand" delivery files formatted to suit the digital access needs of the user. The surrogate version of the digital image or audio file is delivered to a user's e-mail account or sent to a cell phone. Leaving aside, for now, considerations of provenance, copyright, digital rights management and version control, the ease of digital replication is a great advantage for disaster recovery and ensuring continuous access.

Abundant storage and redundant backup routines are a staple of network digital environments. Automatic tape drive backups can help preserve access to digital collections, while master digital copies can be easily duplicated on tape, DVDs or other optical media and stored off-site with

a partner institution. Consider the practice of establishing a mirror database site for digital collections that is operated off-site, that remains up and running during "disruptiveevents" with users unaware there is a problem. Such sites duplicate information on several servers that are separated geographically. All of these servers are connected to the Internet and, if one goes down, the others remain up and running. Redundant storage and multiple access portals are recognized as sound business resumption strategy for institutions addressing disaster preparedness planning. Most organizations with digital collections of significance include replication of collections, and creation of backups, as important tasks to be included in planning from the first. Many organizations are beginning to explore off-site storage of duplicate master images in their sustainability planning.

Planning is key to providing organization and preparation for a smooth and thorough recovery from the disaster event. In the technology domain, change and unpredictability are facts of life, and frequently represent opportunities rather than disasters for a well-planned project. Quick business resumption is possible if redundant systems can maintain access to collections replicated through digitization. It may not happen today, but every organization will experience a disaster event sometime. Pipes can break, electrical shorts can become fires, backup routines can go awry and computers have been known to crash. Because of this, digitization projects and their products should be included in traditional disaster plans. Project managers need to take special precautions to protect not only original collection resources but their digital surrogate master copies located in primary and secondary archival storage repositories.

Inventories of all collection copies must be part of the disaster plan. Check out dPlan: The Online-disaster Planning Tool (http://www.nedcc.org/service/disaster.dplan.php).

Policies and procedures for digitization of collections will play a vital role in disaster preparedness by providing the road map for resumption of collection work. They also provide the foundation for consistency and quality assurance for digital collection work. A common misconception is that there's going to be a technical solution to digital preservation methodology—that some software is going to provide the full solution. The most stable part of the digitization process will be the policies and procedures that assure consistent practice over time. For every organization, early action to establish processes and procedures ensures more flexibility in disaster response. Staff that have policies and procedures to refer to will resume a productive workflow more quickly.

Documentation of unique identifiers and file format naming protocols for digital files, network pathway URLs, and procedures for local and remote storage of digital files are essential. Replacement of the active use copies of digital objects located in a digital collection management system such as http://www.oclc.org/contentdm/default. htm or http://www.dspace.org/DSpace.org/ DSpace is routine and common for digital collection managers. The documentation of all collection copies and identification of all versions is usually contained within the full metadata record for a digital object. It is important to managing digital collections and is a vital component of disaster response planning. This data is just as important as the "traditional" shelf lists and catalog data for analog collections. Fortunately, it is relatively easy to export metadata records and

21

replicate the data in many repositories such as WorldCat, where it can be recovered if the original digital collection database becomes inoperable due to a disaster.

Training and awareness are the most important actions for gaining the skills to address collection preservation and disaster preparations in the digital environment. This issue of the *Oregon Library Association Quarterly* is devoted to training and awareness about disaster and emergency preparedness. The encouraging news is that members of the library and archival communities already know much of what they need to be successful in planning for preservation of digital materials. We can use our knowledge of preserving analog collections to address disaster preparations for digital collections, and then we will need to learn some new techniques that apply to the digital environment.

References

CONTENTdm digital collection management software, http://www.oclc.org/contentdm/default.htm

dPlan™: The Online Disaster-Planning Tool produced by the Northeast Document Conservation Center, http://www.nedcc.org/services/disaster.dplan.php

DSpace an open source solution for accessing, managing and preserving scholarly works, http://www.dspace.org/

LOCKSS software is based on Association of Computing Machinery (ACM) award-winning technology. It provides an OAIS-compliant, open source, peer-to-peer, decentralized digital preservation infrastructure. http://www.lockss.org/lockss/How_It_Works