

State of State Documents

by **Susanne Caro** (Government Information Librarian, Mansfield Library, University of Montana)
<susanne.caro@umontana.edu>

In 1870, New Mexico's territorial governor, **William A. Pile**, wanted to use a room that housed Spanish, Mexican, and U.S. territorial archives in the historic Palace of the Governors in Santa Fe. **Pile** instructed the territorial librarian, **Ira Bond**, to clean out the room. **Bond** did this by selling the documents to local shopkeepers for wrapping paper, giving them to the prison, or tossing them out of a window. **Bond's** actions drew outrage in 1870, caused legal issues in the years that followed, and left modern researchers wondering what documents were lost.¹ Today, his actions seem extreme, but this was not the first or only time unique state history had been lost.

State documents that are deemed to be of long-term value must be organized, stored, made accessible, and the spaces housing these collections must be maintained. Changing technologies and limited space pose similar challenges with electronic materials. Over all of this hangs the constant threat of budget and administrative cuts by individuals with divergent goals.

What Should Be Kept?

State and local government materials are subject to retention schedules that determine the length of time a record should be kept, and if it must be destroyed. Montana's General Records Retention Schedules direct that materials of historical value be sent to the state archives. Sensitive records, such as patient medical files, must be shredded after ten years. Reports, budgets, and other documents intended to inform the public are sent to archives and libraries where they are evaluated for condition, and uniqueness, and to determine the best way to care for the materials. Preservation often starts with condition. Is it in a stable medium? Is it the only copy? What is the file format? Many state libraries have collections of images, cassettes, VHS tapes, or documents on acidic paper that are starting to crumble. These types of materials have a limited lifespan and there is little that can be done to stop the inevitable decay beyond digitization. Digitizing a document creates a surrogate version to preserve the content while also increasing the accessibility of the content.

Documents that are unique should be preserved. The more local the information is, then the less likely it is that any other institutions will have a copy. It is important to save internal reports, policies, records, and materials showing the work of an agency because those may be the only copies. If there is only one copy, it must be kept safe.

Saving Tangible Materials

One danger for historical, print documents like those lost in New Mexico is that they are often the only copy. Before technology made creating multiple copies easier, documents were written by hand or printed in limited numbers. A courthouse fire or a broken pipe

could wipe out years of history. For this reason, it has been important for libraries to collect and disseminate copies. Depending on the state, agencies may be required to follow archival retention schedules for documents or send copies to their state library, where documents are then cataloged, organized, and properly stored.

State libraries with depository programs typically send document copies to designated collections and archives across their state. Spreading documents to geographically diverse locations provides access to users and reduces the likelihood of losing all copies to natural disasters. Libraries in a depository or distribution program are usually a mix of public and academic institutions that keep or discard materials based on the needs of their user communities.

Preservation requirements for paper documents involve keeping materials safe from moisture, insects, fires, administrators, and users. In 2013, a new court clerk in Franklin County, North Carolina, investigated the basement of the courthouse and started a very necessary cleaning project. The basement had been used to store broken furniture, old equipment, boxes of legislative documents, court records, photos, letters, financial records, and more dating back to the 1840s. The local historical society was engaged to inventory the material, but due to mold contamination, concern regarding confidential documents, and miscommunication, all of the records were sent to the local humane society and incinerated.² The historical society's treasure became the county's trash.

Users can also cause loss of materials. In June 2017, the *Medford Mail Tribune* reported that employees for an oil company removed materials relating to the Jordan Cove Energy Project from libraries along a proposed pipeline route.³ At the **Coos Bay Public Library**, an oil employee had asked a librarian for permission to take documents provided by the company. The librarian asked that they wait for staff to pull materials from the collection, which included state and federal documents relating to the pipeline. Rather than wait for the company's materials to be identified the employee took everything, leaving five bare shelves.

The Digital Switch

The growth of digital content removes some pressure libraries concerned with space issues. The number of documents that only exist in an electronic format has increased

sharply over the last fifteen years. **Kristin Martin** and **Jan Regan's** 2003 study, *North Carolina State Government Information: Realities and Possibilities*, found that 50 percent of North Carolina's state documents were born digital.⁴ **Jennifer Davison**, head of content management and access at the **State Library of North Carolina**, estimates that the current born-digital content is nearly 95 percent of all public content. To preserve these materials, plans and policies must be in place for the collection, storage, and accessibility of electronic formats.

Collection

As agencies transition from paper to digital publications, many consider electronic versions exempt from existing policies requiring submission to library or archive programs. In some cases, the agency may believe their obligations to the public are met by having a document available on their website. Capturing this content takes time, infrastructure, and personnel. Web harvesting tools can crawl websites to find environmental reports and videos, but this method can miss materials deeply hidden in a website. A 2009 study by **Claudene Sproles** and **Angel Clemons**, *Permanent Electronic Access to Government Information: A Study of Federal, State, and Local Documents*, found that, compared to federal and local documents, state documents were the most at risk with an estimated 12 percent of digital documents lost or moved to new URL.⁵

The 2016 election in the United States drew more interest in efforts to collect and save digital content from government websites, but this is an endeavor that librarians and archivists have been engaged in for years. Librarians search state pages, especially before changes in administration, to capture materials before they are deleted. Captured pages are archived in local servers and made available through services including the **Wayback Machine**. State digital preservationists have made similar efforts to collect state pages before a transition.

Beyond administrative changes, a document may be removed because of the content. This removal may be for legitimate reasons, such as a report on a state park providing information on archaeological sites which should not be made public, or for inaccuracies like the advanced placement statistics on the

continued on page 20

California Department of Education's website that were removed in May 2017.

Other changes are political. For example, the Wisconsin Department of Natural Resources replaced the content on the page "The Great Lakes and a Changing World." The content changed from discussing climate change to content claiming that causes of climate change are still being debated, implying that the observed effects could be due to a natural cycle. The older version of the site is available for comparison because it was captured by the **Wisconsin Historical Society** as part of their web archives hosted by the **Internet Archive**, a non-profit organization that digitizes, stores, maintains, and provides access to a large collection of materials from libraries, museums, organizations, and individuals.

Storage

There are two popular options for storing electronic content: onsite servers and cloud storage. The practice of sending documents to multiple libraries in different geographies is mirrored on the digital landscape through a philosophy and practice known as Lots Of Copies Keeps Stuff Safe or LOCKSS. Instead of library shelves the electronic documents are sent to multiple servers in different geographical locations. This redundancy prevents loss if one server should fall to floods, fire, or human error. In Arizona, this started with the Persistent Digital Archives and Library System, or PeDALS, which became a collaboration among seven state archives, including Arizona State Library Archives, New Mexico State Records Center, and New York State Archives to store and preserve digital state documents. This duplication can prevent the loss of content.

Loss can be purely accidental as in 2007 when an error by a technical worker deleted 800,000 images from Alaska's Department of Revenue. According to a March 2007 *Seattle Times* article, three hundred boxes of deleted information had to be rescanned after it was found that all the backup measures to protect the digital files had failed.⁶ This was a case where a large amount of content was lost and the error was reported by the media. It is nearly impossible to know how many times smaller amounts of content have been deleted.

Preservation

Once a library has captured born-digital content, or digitized print content to create digital surrogates, these files must be maintained. Digital documents have their own unique preservation requirements and challenges. Having files on a server is not enough. Modern technology changes at a much faster rate: think of old floppy disks, CD-ROMs, and VHS tapes. Magnetic tape lasts about 30 years, CDs last about ten, and the lifespan of digital materials is not known. Moreover, even if the information is available in a specific medium, the technology needed to access

the contents may not be available. Preserving digital materials requires an assessment of the materials and their formats, file monitoring, migration to updated formats, and use of emulation software, such as a program that acts like Windows 98 to access and migrate files. Even if the URL for a document is stored, changes to a website could make a link obsolete. To counter this problem, libraries and archives create persistent URLs, or PURLs. Ideally, PURLs should link to a document in a stable location such as a database.

Managing electronic records also requires checking files for changes that indicate data loss. Tools called checksums assign alphanumeric codes to files based on the content. If the code changes, this indicates changes in the file, which alerts managers to the issue. This technology still requires personnel to verify a problem and upload an uncompromised file. To maintain digital records, local governments in New York share a service called the Digital Towpath (<http://digitaltowpath.org>). This project was spearheaded by the **Center for Technology in Government** to help local governments manage electronic records and to comply with records retention policies and management laws. Record maintenance expenses are reduced by sharing the service. This is not a true archive for managing material for posterity but the model shows promise and meets the needs of the community, including deleting records in accordance with the law.

The Greatest Threat: Funding

When legislators are unaware of the importance of state library preservation efforts, libraries often seem like good places to save money. At the state level, budgets can be cut by the legislature or governor and are rarely restored. For example, in Michigan the entire collection of the state library was nearly dismantled by **Governor Jennifer Granholm's** Executive Order No. 2009-36 in 2009. The library had been a stand-alone agency in 2000 when it had a budget of \$35 million. When the governor's executive order transferred the agency to the Department of Education, the budget was cut by \$24 million. In 2009, the state appropriated ten million for the library, and by 2010 the appropriations were cut by another million. The 2017 budget for the library returned to a higher level, but is still significantly lower than before the recession. Many state budgets cut during the recession have not recovered. An Institute of Museum and Library Services (IMLS) survey found that most state libraries that suffered budget cuts during the 2009 recession have not had their budgets restored as of 2014.⁷ Positions were lost and departments cut or consolidated, so the work of preservation falls to fewer hands. Physical items are less likely to be repaired or stored in a way to prevent further damage. If information technology personnel are lost, digital files may not be properly managed and programs to collect born-digital items may be halted.

In the summer of 2017, the state of Montana experienced a revenue shortfall

resulting in significant budget cuts. For the state library, this amounted to over \$666,000 for fiscal year 2017, and more the following year. The library lost a quarter of its staff. To help reduce the financial loss and continue to provide vital services, the state librarian made the difficult choice to discard a large, historic collection of federal documents in order to rent the space to another agency. Thankfully, however, most of the state documents are still accessible, as the extensive collection had already been digitized through a long-term project with the **Internet Archive**.

Modern-day libraries are unlikely to burn, but they may fade away as library workers are laid off and servers are not properly maintained. A legislator may decide that digital archives are not needed or have other uses for collection space. In a time when people assume that everything is online we lose more content each day. To prevent the loss of our local histories, state libraries, archives, and local governments must work together to develop solutions and plans to mitigate physical and financial disasters. 🍁

Endnotes

1. **J.J. Bowden**, "Estancia Grant," *Office of the State Historian*, accessed November 21, 2017. <http://newmexicohistory.org/people/estancia-grant>
2. "Historians Lament Destruction of Franklin County Records," WRAL.com, last modified December 18, 2013. <http://www.wral.com/historians-lament-destruction-of-franklin-county-records/1322992/>
3. **Vickie Aldous**, "Company Removed LNG Pipeline Documents from Public Libraries," *Medford Mail Tribune*, last modified June 27, 2017. <http://portlandtribune.com/sl/364455-245514-company-removed-lng-pipeline-documents-from-public-libraries>
4. **Kristin Martin and Jan Regan**, *North Carolina State Government Information: Realities and Possibilities*, State Library of North Carolina, 2003. <http://digital.ncdcr.gov/cdm/ref/collection/p249901coll122/id/498567>
5. **Claudene Sproles and Angel Clemons**, "Permanent Electronic Access to Government Information: A Study of Federal, State, and Local Documents." *E-JASL: The Electronic Journal of Academic and Special Librarianship* 10, no. 2 (2009). http://southernlibrarianship.icaap.org/content/v10n02/sproles_c01.html
6. **Ann Sutton**, "Alaska Fund's \$38 Billion Disappearing Act." *The Seattle Times*, last modified March 21, 2007. <https://www.seattletimes.com/nation-world/alaska-funds-38-billion-disappearing-act/>
7. *State Library Administrative Agencies Survey: Fiscal Year 2014*, (Washington, DC: **Institute of Museum and Library Services**, August 2016). <https://www.ims.gov/publications/state-library-administrative-agencies-survey-fiscal-year-2014>