November 2013

People Profile: Richard Pearce-Moses

Editor

Follow this and additional works at: http://docs.lib.purdue.edu/atg

Part of the Library and Information Science Commons

Recommended Citation

DOI: https://doi.org/10.7771/2380-176X.2561

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
often kept on tape to reduce the cost of online storage. These systems may be marketed as a means to store records for a long time. However, within IT, ten years is often considered a very long time, hardly the same time frame in an archivist’s mind.

Possibly the first roadblock to building a digital archive is to ensure that IT professionals on the project understand what archivists mean by a long time. Examples help. The National Archives and Records Administration (NARA) holds the Constitution, a document that has been in use for centuries. Land records are kept permanently to ensure clear title to deeds. Birth records remain in use for decades, throughout an individual’s life, and beyond death for historical and genealogical purposes.

**Fixity and Integrity**

Another characteristic of archival records is the unchanging nature of the information they contain. The records serve as a reliable voice from the past, and that reliability is based on the stability of their form and content. Records may suffer some degradation over time without seriously affecting their reliability as evidence of the past. Paper may yellow and ink may fade, but the record remains readable.

Demonstrating the integrity of electronic records is more challenging. Current methods use a hash value that can detect a change to an individual bit. Unfortunately, those tests cannot indicate whether the change has a significant impact on the content. A single flipped bit might look like a typo or speck in an image. But if information is encoded as a binary zero (no) or one (yes), a flipped bit could completely reverse meaning.

Archivists must find systems that can spot changes resulting from degradation and correct those errors. Current practice keeps two copies of every file. The system constantly checks for degradation, and replaces a corrupted version with the second — presumably — correct copy.

**Preservation of Unique Records**

Archival records are also distinguished by the fact that they are typically unique. Loss of a single publication distributed in even a modest run is mitigated by the availability of other copies. If a copy is destroyed in a disaster at one repository, other copies are likely available at other repositories.

Because records have no redundant copy, archivists take exceptional care to protect their holdings. A second copy of records kept as a check against loss of integrity can be stored offsite, eliminating the risk of losing unique copies. The ease of duplicating digital records and transferring them for offsite storage is one of the greatest benefits of digital records over paper records.

**Authenticity**

Another important characteristic of archives is the need to ensure the authenticity of the records. The archives must be able to demonstrate that the records are what they purport to be, that they are genuine and not fakes or forgeries. As records are often used in litigation, archivists must be able to demonstrate that the records meet the requirements of authenticity outlined in a court’s rules of evidence. When used for historical research, the authenticity of the records serves as a foundation for understanding the past and is essential for factual scholarship.

**Cost**

Even in the best of times, archives seldom have adequate resources. Like libraries, archives are now being asked to do double duty, working parallel in tangible and digital universes, and often with little or no additional funds. Unfortunately, investments in information technology can be very expensive. Some commercial systems designed to meet archival requirements may be prohibitively expensive for even medium and large sized archives.

**Why LOCKSS?**

Archivists and IT professionals must work together to find solutions that can keep archival records accessible for a very long time. They must ensure that the records’ integrity is preserved, that the records are protected against disaster, and that their authenticity can be demonstrated. All this must be done within a limited budget. Unfortunately, because digital preservation is so new, there are no time-tested best practices.

The Arizona State Library, Archives and Public Records is the official archives for the State of Arizona and also serves as the custodian for local governments’ archives. The agency also serves as the official depository of state agency publications. The agency has not yet allowed archival records to be deposited in digital format, but it is under increasing pressure to do so. The agency has effectively been forced to accept digital publications, as many of those documents are never printed. Arizona needed a solution.

Over the past several years, staff has taken the first step of creating secure storage for digital records and publications. However, that system failed to address all the archival requirements for a robust digital repository described here. Commercial vendors often failed to understand the particular needs of the system, especially the need to build a system that could support permanent retention. Vendor systems required both a large up-front investment plus significant ongoing costs in personnel for support.

While LOCKSS was originally conceived as a system for serial publications, a certain parallelism suggested that the technology might be adapted to archives. Where serials have a publisher, possibly with many titles, and many issues within a title, archival records have a provenance, possibly with many series, and many records within those series. On further investigation, LOCKSS clearly addressed the