



Digital Preservation in State Government Best Practices Exchange 2006

<http://statelibrary.dcr.state.nc.us/digidocs/bestpractices/index.html>

Summary of the “Authentication of Digital Assets” Sessions Compiled by Kelly Eubank, Electronic Records Archivist

When participants met to discuss authenticity and authentication of electronic records, they focused on a range of issues surrounding authentication – from the security it provides to the difficulty in determining the official record copy versus the “authentic” copy. Contributors from Washington State Archives, the Illinois State Library, Microsoft Corporation, Library of Congress, North Carolina Secretary of State’s Office, the American Association of Law Libraries and the Government Printing Office explored issues related to technology tools currently available to identify assets as authentic and the need to establish and document workflow processes regarding digital material. Participants also defined a best practice of keeping multiple copies of a record to ensure longevity and data integrity. Unresolved issues surrounded the idea of the “official copy” versus the “authentic copy” and how to create standards to bridge this gap.

During the three exchanges on Authentication, discussions focused on tools available that help to establish the authenticity of a document and the issues those tools raise. Additionally, all agreed that institutions should assess and document workflow in conjunction with these digital tools in order to augment them. At this point in time, technology has advanced to offer a technological solution for electronically “proving” a document is authentic and unchanged. A major theme surrounding these tools is the authority conveyed on the record once it has been authenticated. A government entity, through these tools, verifies the record as official and that it is what it purports to be. Both Adam Jansen of Washington State Archives and Andy Pitman of Microsoft Corporation discussed some of the tools, open-source and proprietary, that can be employed to essentially “freeze” an object. Andy Pitman discussed digital signatures as a methodology for a signer of a document to employ in order to affix their “signature” on a document. Adam Jansen discussed the MD-5 hashing algorithm that Washington State Archive uses to assign a mathematical “stamp” on a file. Should the content of the file change either through media break down or tampering of a file, that change will be reflected in the hash when it is re-run against the object. According to many of the attendees, this ability provides a level of security with electronic files that does not exist with paper files. The same holds true for the digital signature. If someone attempts to change it, it automatically becomes invalid.

Interestingly enough, these new technologies raised several questions about data integrity and the need for the workflow to be documented in order to reinforce the “authenticity” stamp. This documentation forms the framework for the Archives and Libraries roles as custodians of the official, “preservation master” copy. Additionally, many of the participants felt that multiple copies of a record leads to data integrity. Many of the contributors and attendees agreed that multiple copies of material make it much harder to alter the material. Multiple copies means that a person who wishes to alter material would

have to find and then alter several copies of a record in order to make that change permanent.

Participants also discussed the idea of the official copy versus the authentic copy. In other words, what the content creator considers the official copy may be different from the authentic copy simply because the authentic copy has gone through this “hashing” or digital signature process. Butch Lazorchak, from Library of Congress (LC), detailed an effort in which the Library of Congress, with participation from other institutions, attempted to take records out of one system and “ingest” them into another system. LC discovered that different software versions and different standards such as METS yielded different results and which copy is then the authentic copy? This raises a host of concerns as institutions look toward migration as a preservation strategy.

Certain themes were universal across all exchange sessions, regardless of the topic. One such theme was the necessity of sustainable funding and the continual challenge of carrying on digital efforts without it. Attendees also mentioned the importance of education at all levels within state government both inside and outside one’s institution. Advocacy was identified as particularly critical at the legislative level in order to encourage the creation of legislative mandates for the long-term maintenance and preservation of digital assets. Another popular theme was the creation of standards. Contributors noted that standards allow for consistent processes and routines as well as enhance interoperability among systems. Finally, concern was expressed over unnecessary duplication of effort. With so many states getting involved in digital management and preservation, and so little communication between practitioners, it is likely that redundancies will occur. This concern emphasized more than ever the importance of open communication between digital practitioners within the community of practice.