Maine Policy Review

Volume 24
Issue 1 *Humanities and Policy*

2015

Data Curation Needed to Avoid a Digital Dark Age

Desiree Butterfield-Nagy
desiree.butterfield@maine.edu

Data Curation Needed to Avoid a Digital Dark Age

by Desirée Butterfield-Nagy

Few of us have escaped the inadvertent loss of data, whether through a computer crash, the evaporation of several hours' worth of typing due to a power outage or an accidental keystroke, or discovering that files are trapped on storage formats no longer compatible with current systems. While the ephemeral nature of digital files may be no secret, most individuals and organizations could still do considerably more to better manage and preserve their digital resources.

For several years, scholars such as Jerome P. McDonough of the University of Illinois have been warning us of the potential for a "Digital Dark Age," or the growing concern that unless more careful measures are taken, vast amounts of early twenty-first century information— indeed an era characterized by an ever-growing information-based economy—may be lost. Even important national data sets have been overlooked until it was nearly too late. At one point, the 1960 U.S. Census could only be read by two machines, one in Japan and the other at the Smithsonian Institution. Some data from NASA's 1976 Viking landing on Mars were stored on magnetic tape in an unreadable format, and the original programmers had either died or left NASA.

A culture's emphasis on preservation, some scholars suggest, and the selection of durable formats and building materials, may go a long way toward explaining why we know relatively more about some ancient cultures, such as Egypt, compared to more recent civilizations where little may remain. As McDonough (2008) pointed out, "If we can't keep today's information alive for future generations, we will lose a lot of our culture."

Unlike paper or clay records, which have been known to last without intervention for hundreds or even thousands of years, most recommendations are to test digital files every two to three years. Testing even 10 years can prove to be a significant threat to unmanaged files.

Archivists and librarians have long been entrusted with selecting documents thought to have enduring value and taking steps toward ensuring their preservation. An increasingly recognized need for digital curation, or the active selection of digital files and taking steps toward preserving them, has been a natural evolution in a contemporary environment where vast amounts of intellectual and cultural content is born digital and may not be represented in tangible form. While a first impulse may be to encourage the printing of copies, and even if physical space could be found to house a reasonable selection, this effort can fail to capture layers of code and properties that can greatly add to the understanding and value of complex digital objects, whether images, email messages, or relational databases.

Preservation has always required active, managed care, and

FIGURE 1: **University of Maine Computer Center, 1965**

many of the principles that apply to paper-based formats still apply to digital objects. Selecting and providing ideal storage containers and environmental conditions apply to digital materials, but additional measures are needed. A variety of approaches has been suggested for the active management of digital content over time. Some efforts essentially recreate the interface and the original technical environment, while others stress the need to continually migrate information to keep up with evolving file formats. More nuanced preservation programs maintain multiple copies in multiple locations and launch systems that automatically alert digital curators when copies no longer match one another, when file structure has changed, or when formats are at risk of becoming obsolete.

FIGURE 2: **University of Maine Computer Center, 1980**

Now that many agreed-upon practices have emerged, perhaps the phrase Digital Dark Age most appropriately refers to the period between society's widespread adoption of digital technologies and the development of techniques and infrastructure that have allowed the material to be kept safe for the long term. As noted in a 2012 article by Stuart Jeffrey, while many trusted repositories have emerged, barriers to digitally archiving outputs have "moved from the realm of the technical to the realm of the financial, political and practical" (2012: 554). Some estimates suggest that the cost of preserving electronic records will be as much as five times what has been needed for paper-based materials.

Companies, governments, educational institutions, and individuals all invest a tremendous amount of time and resources in creating digital content. Actively thinking about how best to select, manage, and preserve this information is an important part of making sure that the work and intellectual endeavors of today will be available for others to examine and build upon in the future.

REFERENCES

Jeffrey, Stuart. 2012. "A New Digital Dark Age? Collaborative Web Tools, Social Media and Long Term Preservation." World Archaeology 44(4): 553–570. http://dx.doi.org/10.1080/00438243.2012.737579

McDonough, Jerome P. 2008. "'Digital Dark Age' May Doom Some Data." ScienceDaily (October 29). http://www.sciencedaily.com/releases/2008/10/081027174646.htm

Desirée Butterfield-Nagy is an archivist in the Special Collections Department at the University of Maine's Raymond H. Fogler Library. She has coordinated several digitization projects including those providing access to early university publications, photographs, postcards, and diaries through the DigitalCommons@UMaine institutional repository. She also teaches in the university's digital curation certificate program offered through the New Media Department.