University of Nebraska - Lincoln DigitalCommons@University of Nebraska - Lincoln

Faculty Publications, UNL Libraries

Libraries at University of Nebraska-Lincoln

Spring 2012

Preservation and Access in an Age of E-Science and Electronic Records: Sharing the Problem and Discovering Common Solutions

Claire Stewart Northwestern University, cstewart@unl.edu

Follow this and additional works at: https://digitalcommons.unl.edu/libraryscience Part of the <u>Archival Science Commons</u>, and the <u>Cataloging and Metadata Commons</u>

Stewart, Claire, "Preservation and Access in an Age of E-Science and Electronic Records: Sharing the Problem and Discovering Common Solutions" (2012). *Faculty Publications, UNL Libraries.* 389. https://digitalcommons.unl.edu/libraryscience/389

This Article is brought to you for free and open access by the Libraries at University of Nebraska-Lincoln at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Faculty Publications, UNL Libraries by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



Published in *Journal of Library Administration* 52:3–4 (Spring 2012), pp. 265–278; doi: 10.1080/01930826 .2012.684505

Copyright © 2012 Claire Stewart. Published by Routledge/Taylor & Francis. Used by permission. Published online July 2, 2012.

Preservation and Access in an Age of E-Science and Electronic Records: Sharing the Problem and Discovering Common Solutions

Claire Stewart

Northwestern University, Evanston, Illinois, USA

Corresponding author – Claire Stewart, Digital Collections and Scholarly Communication Services, Northwestern University, 1970 Campus Dr., Evanston, IL 60208, USA, email <u>claire-stewart@northwestern.edu</u>

Abstract

As academic libraries grapple with the challenge of preserving their own digitized special collections, intensification of interest in preserving other electronic content may present opportunities to collaborate with organizations on campus. This article offers a brief introduction to some of the core issues in digital preservation and suggests an orientation to the problems that can be helpful in thinking about how to join forces with others on campus.

Keywords: preservation, curation, digital, collaboration, e-science

Introduction

Preservation is having a moment. From preserving digitized and born-digital special collections, to electronic records management, email archiving and e-discovery, to discussions about how to store and manage digital mountains of research data—some generated at incredibly rapid rates by networked sensors, always on, always collecting—the entire academic organization has become aware of the importance of having a good digital memory. For academic libraries, this dawning awareness presents both opportunity and risk. There is a valuable opportunity to more fully join library problems to the problems facing the parent institution, to distribute both costs and risks, and to share the burden of investing valuable staff time in another new enterprise as budgets contract. Libraries can share their special expertise and enjoy the benefits of economies of scale. But there is also risk that the library will either oversell or undersell its role as partner in solving these problems. Oversell, and the library may bear crippling infrastructure costs, drain staff resources away from other valued services, and run the risk of failure if unable to retool, retrain, and provide credible and sustainable service. Undersell, or shy away from accepting a role in meeting this new grand challenge, and the library may cease to be viewed as a key partner in memory services, becoming just another client of technology services that other units must provide. Actively seeking partners and seeking to share librarian and archivist expertise can be an effective strategy if the library can identify the shared problems and explain how involving itself in solving these problems is a natural extension of its mission.

The preservation dilemma for library- and archive-trained specialists is becoming acute: they must continue to maintain existing practices, including caring for sizable paper-based collections, while at the same time acquiring the skills and tools to effectively transition to a new digital environment. Libraries have to find ways to preserve digital content while keeping the "old world" safe (Gracy & Kahn, 2012). This dilemma is not unique to preservation activities: even as libraries embrace electronic journals and ebooks, shift ever greater percentages of collections budgets to electronic materials and embrace mass digitization, they continue to shoulder the substantial cost of maintaining analog services: space for physical collections and human expertise to connect readers to information. But physical space is a comparatively stable consumer of resources: heat, electricity, maintenance, and insurance all demand a continuing chunk of budgets, to be sure. Imagine instead that the library building were to became obsolete every 4-5 years and require total replacement: if all of the books in it had to be taken out, checked for missing pages, and moved to a new location. Imagine if the only way to be sure the books were truly safe was to pay to keep a second copy of each one at another location, or even a second and a third copy, preferably at a geographic remove as a safeguard against catastrophic events. Actions like these are required to preserve content in the new digital infrastructure libraries and universities are building. Server hardware becomes obsolete and must be replaced. Regular file integrity checks, migration, replication, and mirroring have become accepted digital preservation strategies, but the vast majority of libraries do not have systems and programs in place to adequately carry them out (Skinner & Schultz, 2010). Libraries have struggled to move much beyond basic description and storage for digital content. Providing true digital preservation, and thinking about how to abstract and extend its services to extend storage, description, and management services to the technical environments that are in use and most useful to other campus content stewards, is still impractical for most libraries. However, these need not be problems that libraries must solve on their own.

The Rise of Electronic Content on Campus: What Are These New Things and Why Are They Special?

Academic libraries, particularly through special collections and archives units, have devoted effort to collecting and maintaining rare, unique, and valuable materials. Although libraries consider special collections to be repositories that serve an international community, most also consider the needs of their local communities of scholars in building special collections, and archives often serve as memory for the parent institution itself. Libraries are also disposed to view the digital output of their own local communities as valuable institutional assets, worthy of the effort they already devote to other types of special collections (McMillan, Schultz, & Skinner, 2011). Other units on campus are also thinking hard about the challenge of safeguarding locally produced information.

Electronic Records

Some libraries already serve as the official repository for the records of the institution and are starting the transition to maintaining these records in electronic form. At other colleges and universities, deposit with the archives is optional, or there may be an electronic records program that is separately administered. Electronic records are also common across all sectors of society: every business, bank, state or local government, and doctor's office is generating records in some electronic format. This has led to an explosion of demand for electronic systems to reliably track human resources, invoices and payroll, procurement and purchasing, medical transactions, and the like. There is a professional community of records managers, with professional associations, certification programs, and standards. In part driven by concerns about compliance and litigation, electronic records managers and systems have proliferated (Scanlan, 2011).

Electronic records systems already exist in colleges and universities too, though they are likely viewed primarily as administrative systems: records are often transactional, the information may be carefully guarded for privacy and confidentiality reasons, and data is likely siloed. There is growing recognition of the need to maintain these records in electronic form over a longer time period, presenting new opportunities for libraries to seek mutually beneficial partnerships with records managers on campus. Email archives, document management systems, course management systems, and enterprise resource planning (ERP) systems contain information with historical and research value, but they are stored in systems optimized for rapid and efficient transactions, not for long-term preservation. Many campus units make important information such as directories and course catalogs searchable via dynamic Web sites, but may have done little to prepare for longterm preservation of these assets. Archivists with records management responsibilities have so far been able to make little headway in establishing viable electronic records programs (Davis, 2008; Zach & Peri, 2010; Brantley, 2010); these efforts have struggled to achieve buy-in from records managers or to attract sufficient IT support. Rather than indicating the fruitlessness of collaboration, however, this struggle to stabilize electronic records underscores the absolute necessity of close working relationships with both groups: digital preservation can only succeed in an environment where it is a consideration from the beginning of the content's life cycle, and this can happen only when technical administrators, records managers, archivists, and digital preservationists are aligned.

New Challenges: e-Research and e-Science

One of the most rapidly growing problems for research institutions is sustaining and safeguarding the digital products of research. Particularly in the sciences, research has become a computationally intensive activity, generating large numbers of files, files that are very large in size, or both (Hey, Tansley, & Tolle, 2009). Funded projects may last a few years and then cease, or they may continue over many years and be funded through a series of grants. Researchers need guaranteed and reliable storage for their digital data. There is also growing interest in making data available outside of the research group or university: as the accompaniment to a research article, to promote transparency through reproducibility, to avoid wasting time by sharing negative results, and to reduce redundancy and waste caused by funding the same research more than once. Data sharing demands clear description through consistent, and preferably standardized, metadata; repositories that can provide adequate network infrastructure for uploading and downloading, and tools that are easy for researchers in multiple disciplines to use. Colleges and universities are not very well prepared to develop stable systems for reliable long-term preservation of these data, but must share the responsibility for helping to identify solutions (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010a). Because these materials are part of the original scholarly output of universities, it is consistent with the mission of their libraries that they should also seek to help solve the problems of research data management and preservation. Library experience identifying and implementing metadata standards, understanding the value of unique and stable identifiers, experience with versioning and citation all represent valuable contributions to the e-science data management problem. Helping to discover solutions to these, and to the data integrity, normalization, and migration challenges will also aid libraries seeking to solve digital preservation problems for their digital collections.

The Library's Experience and Readiness

Any library that has engaged in digital archiving or digital reformatting of its collections is aware that these activities can incur considerable costs, both short- but particularly longterm, often costs that libraries have struggled to meet. Digital storage may be cheap and getting cheaper, but effectively managing it is just as expensive as ever. Problems of frailty in physical media have been replaced with the more worrying problem of the frailty of digital content: digital media obsolescence, bit rot, or disk failure (Wright, Miller, & Addis, 2009), file format obsolescence or dependence on proprietary software, and bedeviling challenges around authenticity (Cullen, Hirtle, Levy, Lynch, & Rothenberg, 2000) and provenance. Some strategies for detecting alterations to paper records are centuries old, but digital forensics is still a relatively young field (Kirschenbaum, Ovendon, & Redwine, 2010), and the ease with which digital content can be altered or destroyed is a looming threat to the information steward who seeks to preserve authentic digital records, and to assure future researchers that the bits have not been altered since they were first born.

Context, particularly as expressed through metadata, the key to so many information management problems, is doubly important in a digital environment. In addition to describing intellectual content, metadata tracks changes and other events in the life cycle of content (PREMIS Editorial Committee, 2011), captures its digital characteristics, and provides important information about its connections to other content: a scan of a book page is a child of a full digital book file; the first draft and the second draft of a prominent scholar's manuscript are intellectually connected; a day's worth of data about astronomical observations is the child of the week data, the month data, the year data. These are relationships that must be preserved as a digital object moves through digital storage and access systems. For librarians, preserving both essence and many different forms of "aboutness" is instinctive: understanding the importance of the one to the other is one of the core values of the profession. Librarians have worn grooves in their brains with years of painstaking MARC record and archival finding aid creation, and, for that matter, information literacy training, and selecting appropriate materials for library collections. The library's counterparts in teaching and research support, information technology, and academic and administrative offices across the institution are also critically dependent on a close connection between keeping the stuff safe and keeping the stuff about the stuff safe, but they may not know it (yet). As EMC researchers observed in their 2011 report on the state of the digital universe: "Big data will be a fountain of big value only if it can speak to you through metadata" (Gantz & Reinsel, 2011, p. 11).

The opportunity and challenge for the library is to effectively partner with key campus groups, extend traditional library expertise to identify common problems and work together to solve them while simultaneously redefining what preservation means for its own collections and programs. Libraries' understanding of the value of both metadata and preservation will be among their most valuable contributions to the new grand challenge facing knowledge creators (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010b).

Defining Digital Preservation and Preparing for Collaboration

For libraries, a digital preservation service is often limited in definition to tools that preserve bits; description, discovery, and access may be handled by other systems. The library may have a digital repository, either for its own digitized collections or to collect the scholarly output of the institution, and may only lack preservation services to guarantee the health of the bits in the system. In the case of other campus content producers, however, solutions to long-term management of both context and bits, supporting both access and preservation, may be lacking. From a coalition-building standpoint, making a case for pure preservation, without access, may be a nonstarter, as potential stakeholders may fail to find value in a major infrastructure investment that does not address a broader spectrum of needs, discovery and use paramount among them. Therefore, an important establishing topic will be exploring the relationship between needs for a digital preservation solution and affiliated needs from an access system.

The Consultative Committee for Space Data Systems' Reference Model for an Open Archival Information System (2002) specification document offers a useful framework for talking about the phases in the lifecycle of digital content as it is passed between systems, the actions performed on it from creation onwards. Other organizations have developed data lifecycle models (Ball, 2012) that will also be helpful in making explicit the different stages, considering what pieces of information about the content and the circumstances of its creation are needed to make it usable in the future.

Another valuable first discussion will be to establish a baseline definition for "adequate" digital preservation. The DRAMBORA (Digital Curation Centre & Digital Preservation Europe, 2007) and TRAC (Consultative Committee for Space Data Systems, 2011) projects have developed instruments to evaluate the trustworthiness of digital systems and to attempt to quantify the likelihood of information damage or loss. They guide system managers through a process that captures and evaluates the organizational and technical setting. The DRAMBORA toolkit attempts to capture and assign a numerical value to problems that could occur, for example, "Negligible impact, results in isolated but fully recoverable loss of digital object authenticity and understandability" is given a risk score of 2, but "Cataclysmic impact, results in total and unrecoverable loss of digital object authenticity and understandability" achieves a risk score of 7 (Digital Curation Centre & Digital Preservation Europe, 2007, p. 89). In addition to helping libraries understand and measure the risks they face, these tools can be valuable for starting conversations with other groups on campus. Groups outside the library will readily understand concepts such as continuity planning, organizational stability, and service level agreements. They also introduce ideas that may be new to groups whose main challenges lie elsewhere: an IT administrator will appreciate the value of metadata that can clearly identify content owners and provide a clear record of changes over time. Understanding where an organization falls on the preparedness spectrum is an important first step in digital preservation planning. The metrics expressed in tools such as TRAC and DRAMBORA can be used in conjunction with other technical infrastructure planning activities that may be the domain of others on campus, such as predicting growth in digital storage needs, evolving identity management services for authentication and authorization, and maintaining security requirements for data in keeping with evolving approaches to architecture, such as cloud services.

Identify Common Problems

One way to find common ground with potential collaborators is to articulate shared problems and to use them as a starting point for conversations about shared solutions. Among the needs or problems likely to be common across different information management groups are:

Scalable massive storage

A common core problem for a data-intensive science group or a collections reformatter will be acquiring sufficient digital storage space. Long before there is any thought of metadata or integrity checks, creators will be concerned with keeping the creation pipeline supplied with disk. Leaving aside need differentiators such as the speed of the network leading to the disk or frequency of changes, appetite for storage will be a common problem and an excellent starting point for collaboration discussions.

Unique identifiers for people and objects

University human resources systems, document management and email systems, research administration staff, and authors of scholarly papers all deal in information about people. Libraries, of course, pay very close attention to author, creator, or contributor names and roles. Author disambiguation is a major challenge in research publication systems. Administrative and academic computing systems across campus need to be able to rely either on a single source for identity management, or a federated set of identity services to track content ownership and other information. These efforts will also benefit from the library's decades of experience with standardized identification schemas, and with national and international efforts (ORCID, ISNI, etc.) to develop shared personal and corporate identity systems that can be used across many domains. Beyond identifiers for people, these groups will also find value in schemes to create stable and universal identifiers and URNs for objects, to support citation and accurate version referencing.

Reliably track provenance and change

Digital preservation systems must preserve a trustworthy record of the events in the lifecycle of an object. Administrative record stewards and research scientists will benefit from a system capable of logging changes and tracking versions. Version tracking will support recovery in case of accidental alteration or deletion. A log or other record of events, such as integrity checks, file format transformations, or record edits, will make actions in the system transparent to future users. Even minor changes may have unintended consequences, so systems designed to track and, to the extent possible, make them reversible may be of tremendous value to several key campus communities. Although there is much to be gained from this type of provenance tracking, there are also costs associated with it: versioning very large files will rapidly consume scarce disk resources, and maintaining a complete record of all changes, and the persons who perform them, in frequently updated systems may slow them to the point of inoperability. Therefore, decisions about provenance tracking must include careful assessment of costs and benefits.

Processes for making preservation and removal decisions

Some content will be subject to clearly articulated retention and deletion requirements, but in many cases, a framework must be designed to determine when and on what criteria objects will be selected for long-term preservation. This need aligns closely with the provenance need. Libraries must determine whether they should retain all versions of their digitally reformatted collections, or only high-resolution copies. Research groups must decide whether all versions of a data set are retained, all raw data or only processed data, only those containing significant results, or only those directly linked to published research. Records managers may have specific requirements for retention of email and other correspondence. Some born-digital documents will have future research value, and others will not. Collectively, the organizations with a stake in long-term digital preservation can assist each other in developing frameworks for preservation and removal decisions. Partners need not accept the burden of making decisions that cannot be revised, however: "A decision to preserve now need not be thought of as a permanent or open-ended commitment of resources over time. In cases where future value is uncertain, choosing to preserve assets at low levels of curation can postpone ultimate decisions about long-term retention and quality of curation until such time as value and use become apparent" (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010a, p. 2).

Automated feature extraction and description

Libraries have struggled to maintain good descriptive metadata for large collections of objects, evolving practices for collection-level description and other strategies to streamline metadata creation. In the digital realm, and as faculty researchers and others become responsible for describing digital content destined for a preservation system, a balance must again be struck to adequately describe large collections of content without overwhelming available human resources. One strategy will be to implement systems that can automatically extract and track features such as file size, file type, digital signature, and even keywords when appropriate. Likewise, inferring information about authors and authority based on the source of records or other information may provide sufficient description for a preservation system (Pearce-Moses, 2009).

Ensuring against accidental or malicious damage to bits and files

All potential partners will want a digital preservation system to have the capability to guard against, and recover from, accidental or intentional damage. A basic feature of a digital preservation system is periodic integrity checking, with some mechanism to recover in case of data loss, such as by recovering from an earlier backup or healing by rebuilding a file from mirrored copies. This form of bit-level preservation is file-type and use agnostic, though it may face special challenges in content with very large files or complex internal structures, such as relational databases. In these cases, the decision processes discussed earlier for making preservation decisions will be crucial.

Support for complex objects

In addition to complex objects such as databases, many other types of internally complex content will need to be preserved. Digitized library collections will have internal structures: multipage books, photographs in albums, etc. Likewise, important administrative or research objects may have structure. A faculty member who has collected video recordings and produced written works that refer to these media objects will want a preservation system to be able to maintain the connections between pieces of digital information.

Security and compliance with regulations and mandates

Some data will be subject to specific retention and destruction requirements. Research data, particularly when human subjects or health information are concerned, may have to be anonymized or safeguarded in very specific ways. Library collections may be subject to donor or privacy restrictions. It is likely that every potential partner in a digital preservation program will have some specific requirements for security and control, and may be required to demonstrate the system's efficacy to internal or external auditors.

Integration with external systems: Accept/ingest and discover/publish/ share

As the volume of digital content grows, depositing content in a digital preservation system manually—that is, by human intervention—will become unsustainable. A shared digital preservation program and its systems must evolve capabilities so that other systems and software can connect to it to query or display its contents. These systems, which may include a wide variety of administrative information systems, small researcher lab software, or distributed content management systems, must also be able to publish to the digital preservation system new content or changes to existing content. Preservation actions may be based on business rules, and not on human actions, so systems must evolve to accommodate this type of submission management.

Automated transformation for normalization, standardization, and forward migration

Many stakeholders will also want to maximize preservability by adherence, whenever possible, to file formats and descriptive standards that best lend themselves to long-term preservation. Although it will not be appropriate for all content types, a digital preservation system should be capable of performing some transformations to render content in its most preservable format. Publications produced in Adobe Illustrator can be exported to PDF/A. Tabular data can be retained in Microsoft Excel format but also exported to tab- or comma-delimited plain text or crosswalked to XML to improve the chances of long-term preservation. Images in proprietary formats may be converted to more standard file formats. In all cases, these actions must be taken with full awareness of any loss of quality or provenance information, and care taken to also preserve a copy of the original file, even when there is little hope that the format will be readable in the future. Identifying an appropriate crosswalk from the native content format to a preservable format may be one of the most challenging parts of the process; even email, arguably the most ubiquitous, plentiful, and simple technology deployed in an academic setting, is difficult to capture and preserve systematically, largely due to the flexibility of configuration and degree of user autonomy (Prom, 2011). As technology evolves, forward migration into new file formats will also be a valuable service, ensuring that documents and other files can continue to be accessed with widely available software.

Desire to collaborate with others beyond the campus: Shared standards, shared infrastructure, and shared content

A major point of interest for many organizations on campus with an interest in digital preservation will be an accompanying interest in collaboration beyond the institution. This may take many forms, including an interest in shared infrastructure, participating in or adopting standards developed externally, and the like. Many libraries with experience with LOCKSS or Private LOCKSS Networks (PLN) have enjoyed a level of success with distributed preservation solutions (Skinner & Schultz, 2010). Researchers are very likely to have collaborators at other institutions, and will need to share data. Their funders, publishers, or disciplinary communities may develop services for data deposit, which may displace need and free up valuable resources for other users. Such a development may, however, make the researcher's needs more complex and challenging, requiring synchronization with external systems, or support for new identity or metadata schemes. This may

also reduce the interest of an influential group of campus partners, who may care less about robust and scalable local systems. Libraries will want to make their collections findable and usable beyond campus borders, but will share an interest in a robust access management system for collections restricted due to copyright status, donor intent or for other reasons. Preservation and access, as in the analog realm, are complementary activities.

Evolve Sustainability and Governance Models

Key concerns in implementing a successful digital preservation system will include questions of cost, how costs will be distributed across partners, and how shared programs will be governed. Studies have shown that sustainability planning and exploration of economic models must be a part of the conversation from the beginning (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010a) (Eakin, Friedlander, Schonfeld, & Choudhury, 2006). The Blue Ribbon Task Force outlined five conditions required for sustainable digital preservation: "recognition of the benefits of preservation by decision makers; a process for selecting digital materials with long-term value; incentives for decision makers to preserve in the public interest; appropriate organization and governance of digital preservation activities; and mechanisms to secure an ongoing, efficient allocation of resources to digital preservation activities" (Blue Ribbon Task Force on Sustainable Digital Preservation and Access, 2010a, p. 12). Many organizations have an interest in sustainable preservation for digital research data: the scientists themselves, the funders, and several groups within the university including IT, research administration, and the library. Some digital content will be more clearly linked to the university than others. Administrative and business records and digitized library collections are clearly assets of the University. Although the university will feel a strong commitment to faculty research in active phases, there may arise questions about the cost of maintaining data after projects have completed, or after the researcher moves to another university. These may be reasons to consider hybrid approaches to locating preservation services, and also to distributing costs. Universities are experimenting with different funding models: charging researchers an up-front fee and guaranteeing forever access (Goldstein & Hillegas, 2011) or building costs into new funding proposals (Johns Hopkins University, n.d.). Researchers will also have varying ability to contribute financially: science and medicine faculty may expect that their work will be supported through large grants, but social science and humanities faculty may rely on smaller awards or internal funding to support their activities. A key to evolving a shared approach to preservation will be to begin the exploration of funding early on, to make it a transparent part of the planning process, and to consider making adjustments over time or to have a variety of possible levels of participation and support.

In close conjunction with economic models, governance models must be developed to make the administrative structure and decision-making authorities clear. Participants must clearly understand their roles. As others (Berman, McDonald, Schottlaender, & Kozbial, 2007) have noted, formalizing trust in design of organization and governance models is a crucial step in expressing these roles and in helping partners to understand their commitments: "Working across institutional and organizational boundaries is one

step toward developing the necessary shared datacyberinfrastructure. Solidifying this process using formalized trust mechanisms is crucial to long-term sustainability" (Berman, McDonald, Schottlaender, & Kozbial, 2007, p. 5).

Be Not Naïve: Barriers to Successful Collaboration in Digital Preservation

Although, as discussed above, the library's needs for a digital preservation solution will have much in common with the needs of other groups, forging campus partnerships to implement solutions will not be easy. A review of the organizational challenges typically faced in cross-organizational collaborations is outside of the scope of this article, but it should be expected that not all organizations will be ready to collaborate. Should one group desire to have a greater share of decision-making, or prefer to work only with collaborators from the same area of the institution, partnering will be difficult. Some units will have very legitimate concerns about security, confidentiality, or privacy and may be reluctant to use shared systems to manage their content, particularly if administrative access is shared with staff outside of their group. They may be subject to particular regulatory requirements, such as HIPAA, for ensuring data privacy. For such groups, distributed storage and administration may be antithetical to their goals (Skinner & Schultz, 2010), although an option to enforce more strict separation between dark-archived digital preservation services and access systems may mitigate some of these concerns. Likewise, some emerging commercial solutions, for example storage solutions, may embed regulatory compliance (Gantz & Reinsel, 2011) in the underlying architecture, so partners must be sensitive to these opportunities. There may be significant gaps in readiness and staff skill across potential partners. The library will have a strong interest in digital preservation, but its relevant expertise may be limited to a very few staff. The same is likely to be true for other units. Readiness, particularly readiness to seek common ground, transcend disciplinary practices and norms, may also be limited, or may vary across partner groups. Finally, it is very likely that a shared digital preservation system will not be a one-size-fits-all solution. It may be a set of affiliated systems, each tuned to a specific purpose, and few utilized by all partners. Solutions may have to grow and evolve over time. The library can benefit from collaboration on digital preservation programs, and can contribute valuable expertise to the process, but must be aware that the problem is difficult, collaboration itself takes effort, and planning for sustainability must begin immediately.

References

- Ball, A. (2012). Review of data management lifecycle models. Retrieved from http://opus.bath.ac.uk/ 28587/
- Berman, F., McDonald, R., Schottlaender, B., & Kozbial, A. (2007). The need for formalized trust in digital repository collaborative infrastructure. Presented at the NSF/JISC Repositories Workshop. Retrieved from http://www.sis.pitt.edu/~repwkshop/papers/berman_schottlaender.html
- Blue Ribbon Task Force on Sustainable Digital Preservation and Access. (2010a). *Sustainable economics for a digital planet*. Retrieved from http://brtf.sdsc.edu/biblio/BRTF_Final_Report.pdf

- Blue Ribbon Task Force on Sustainable Digital Preservation and Access. (2010b). *Grand challenge: Sustainable knowledge infrastructure*. Retrieved from http://brtf.sdsc.edu/biblio/Grand_Challenge_BRTF.pdf
- Brantley, S. A. (2010, July). Determining current practices for college and university electronic records management programs: Follow-up interviews. University of North Carolina, Chapel Hill. Retrieved from http://dc.lib.unc.edu/u?/s_papers,1268
- Consultative Committee for Space Data Systems. (2002). Reference model for an open archival information system (OAIS) (No. CCSDS 650.0-B-1). Washington, DC: National Aeronautics and Space Administration. Retrieved from http://public.ccsds.org/publications/archive/650×0b1.pdf
- Consultative Committee for Space Data Systems. (2011). Audit and certification of trustworthy digital repositories (No. CCSDS 652.0-M-1). Retrieved from http://public.ccsds.org/publications/archive/ 652×0m1.pdf
- Cullen, C. T., Hirtle, P., Levy, D., Lynch, C., & Rothenberg, J. (2000). Authenticity in a digital environment. Washington, DC: Council on Library and Information Resources (CLIR). Retrieved from http://www.clir.org/pubs/reports/pub92/pub92.pdf/view
- Davis, S. E. (2008). Electronic records planning in "collecting" repositories. *The American Archivist*, 71(1), 167–189.
- Digital Curation Centre, & DigitalPreservation Europe. (2007). Digital repository audit method based on risk assessment (DRAMBORA). Retrieved from http://www.repositoryaudit.eu/
- Eakin, L., Friedlander, A., Schonfeld, R. C., & Choudhury, S. (2008). A selective literature review on digital preservation sustainability. Retrieved from http://brtf.sdsc.edu/biblio/Cost_Literature _Review.pdf
- Gantz, J., & Reinsel, D. (2011). *Extracting value from chaos*. IDC IView. EMC Corporation. Retrieved from http://www.emc.com/collateral/demos/microsites/emcdigital-universe-2011/index.htm
- Goldstein, S. J., & Hillegas, C. (2011). DataSpace: A funding model for the preservation and sharing of research data. Retrieved from http://arks.princeton.edu/ark:/88435/dsp01w6634361k
- Gracy, K. F., & Kahn, M. B. (2012). Preservation in the digital age. Library Resources & Technical Services, 55(1), 25–43.
- Hey, T., Tansley, S., & Tolle, K. (Eds.). (2009). Jim Gray on eScience: A transformed scientific method. *The fourth paradigm: Data-intensive scientific discovery*. Redmond, WA: Microsoft Research.
- Johns Hopkins University. (n.d.). Data conservancy | A blueprint for research libraries. Retrieved from http://dataconservancy.org/
- Kirschenbaum, M. G., Ovendon, R., & Redwine, G. (2010). Digital forensics and born-digital content in cultural heritage collections (No. 149). Washington, DC: Council on Library and Information Resources (CLIR). Retrieved from http://www.clir.org/pubs/abstract/reports/pub149
- McMillan, G., Schultz, M., & Skinner, K. (2011). *Digital preservation* (No. 325). SPEC Kit. Washington, DC: Association of Research Libraries. Retrieved from http://www.arl.org/news/pr/spec325-25oct11.shtml
- Pearce-Moses, R. (2009, August 19). Finding a new way: Using automated business rules to process electronic records in the persistent digital archives and library system (PeDALS). Presented at the Society of American Archivists, Austin, TX. Retrieved from http://www.pedalspreservation.org/Resources/ SAA_PeDALS.mp3
- PREMIS Editorial Committee. (2011, January). PREMIS data dictionary for preservation metadata, version 2.1. Library of Congress. Retrieved from http://www.loc.gov/standards/premis/v2/ premis-2-1.pdf

- Prom, C. J. (2011). *Preserving email*. DPC Technology Watch Series. Digital Preservation Coalition. Retrieved from http://dx.doi.org/10.7207/twr11-01
- Scanlan, K. (2011). ARMA v. SAA: The history and heart of professional friction. *American Archivist*, 74(2), 428–450.
- Skinner, K., & Schultz, M. (Eds.). (2010). *A guide to distributed digital preservation*. Atlanta, GA: Educopia Institute. Retrieved from http://www.metaarchive.org/sites/default/files/GDDP_Educopia.pdf
- Wright, R., Miller, A., & Addis, M. (2009). The significance of storage in the "cost of risk" of digital preservation. *International Journal of Digital Curation*, 4(3). doi: 10.2218/ijdc.v4i3.125
- Zach, L., & Peri, M. (2010). Practices for college and university electronic records management (ERM) programs: Then and Now. *American Archivist*, 73(1), 105–128.