

Preserving and ensuring long-term access to digitally born legal information

Sarah Rhodes^a and Dana Neacsu^{b*}

^a*Georgetown Law Library, Washington, DC, USA;* ^b*Columbia Law School Library, New York, NY, USA*

Written laws, records and legal materials form the very foundation of a democratic society. Lawmakers, legal scholars and everyday citizens alike need, and are entitled, to access the current and historic materials that comprise, explain, define, critique and contextualize their laws and legal institutions. The preservation of legal information in all formats is imperative. Thus far, the twenty-first century has witnessed unprecedented mass-scale acceptance and adoption of digital culture, which has resulted in an explosion in digital information. However, digitally born materials, especially those that are published directly and independently to the Web, are presently at an extremely high risk of permanent loss. Our legal heritage is no exception to this phenomenon, and efforts must be put forth to ensure that our current body of digital legal information is not lost. The authors explored the role of the United States law library community in the preservation of digital legal information. Through an online survey of state and academic law library directors, it was determined that those represented in the sample recognize that digitally born legal materials are at high risk for loss, yet their own digital preservation projects have primarily focused upon the preservation of digitized print materials, rather than digitally born materials. Digital preservation activities among surveyed libraries have been largely limited by a lack of funding, staffing and expertise; however, these barriers could be overcome by collaboration with other institutions, as well as participation in a large-scale regional or national digital preservation movement, which would allow for resource-sharing among participants. One such collaborative digital preservation program, the Chesapeake Project, is profiled in the article and explored as a collaborative effort that may be expanded upon or replicated by other institutions and libraries tackling the challenges of digital preservation.

Keywords: right to information; digital legal information; digital preservation; legal archives

Introduction

Legal information and risk: history and background

Nearly two centuries ago, American Founding Father James Madison wrote:

[A] popular Government, without popular information, or the means of acquiring it, is but a Prologue to a Farce or a Tragedy; or, perhaps both. Knowledge will forever govern ignorance: And a people who mean to be their own Governors, must arm themselves with the power which knowledge gives. (Madison, 1910, p. 103)

*Corresponding author. Email: dneacs@law.columbia.edu

Madison helped author the United States Constitution and draft the Bill of Rights – two historic legal documents that continue to influence conceptions of basic human rights throughout the world. As his statement affirms, among those rights, access to information is fundamental; it is the cornerstone of a society built upon ideals of egalitarianism and popular governance. This notion is as true today as it was in the early days of the American Republic. The body of current and historic legal works that comprise, explain, define, critique and contextualize society's laws, legal institutions and systems of justice represent a vital portion of the human record. These materials stand as the underpinnings of democracy, and access to legal information is essential for law practitioners, lawmakers, jurists, legal scholars and ordinary citizens alike. To ensure that this information remains accessible to current and future generations, it must be preserved. In a day and age characterized by the proliferation of information in digital formats, this task requires new strategies and new ways of thinking about preservation.

The written record, including our laws and legal documentation, has always been in danger of destruction, whether by natural or human means (Deegan & Tanner, 2006). The British burning of the Capitol building in 1814, for example, destroyed the original 3000 volumes of the Library of Congress (*International Review*, 1878). In 1966, the Arno River escaped its banks in the city of Florence, destroying masterworks of art as well as thousands of books, manuscripts and legal records in the collections of the Biblioteca Nazionale and Archivio di Stato (*Burlington Magazine*, 1967). More recently, the war in Iraq resulted in the tragic burning and looting of the National Library of Iraq (Kniffel, 2003), and in New Orleans, state and local officials continue to wrestle with the recovery and rescue of government records and documents following Hurricane Katrina in 2005 (NASCIO, 2007).

Human pursuit of profit has also put printed information at risk; in the latter part of the twentieth century, it was discovered that books published from the mid-1800s onward were deteriorating at an alarmingly rapid rate, the result of inexpensive mass-printing processes using highly acidic paper. To ensure that these materials and their content would be available for future access, libraries took action to preserve them, including treatments using acid-free materials, controlled environmental conditions and conversion to archival microforms (Anon., 1980).

At the present time, a new threat to our written heritage has surfaced, demanding immediate attention and action. The twenty-first century has witnessed not only the emergence of information created and disseminated in digital formats, but also an unprecedented mass acceptance and adoption of digital technologies, resulting in a flood of digital information that is expanding exponentially. In fact, for the first time in history, the amount of digital information 'created, captured or replicated' in the year 2007 alone, which equals about 281 billion gigabytes, surpassed the world's existing electronic storage capacity (International Data Corporation, 2008). Within a few short years, by 2011, the International Data Corporation expects the digital universe to increase 10 times over.

Not surprisingly, important legal materials are increasingly being digitally born and then distributed online rather than published on paper. Since the mid-1990s, for example, the number of government documents distributed in digital, as opposed to print, formats by the United States Government Printing Office (GPO) has ballooned. Federal and state agencies over the past decade have likewise produced a growing number of digitally born documents and reports, which have been posted directly to the Web (Lyons, 2006). Court opinions are now being published online,

and legal scholarship increasingly relies on digitally born sources, identified only by a Uniform Resource Locator (URL), or website address, directing one to an online document (Rumsey, 2002; Neacsu, 2007). In fact, what has been called the 'disintermediation of legal scholarship' through collaborative and open-access Web-based publishing is having a noticeable impact on the practice and study of law in the United States (Solum, 2006, p. 1071). Articles and commentary posted on legal Web logs (blogs, or 'blawgs'), for example, have been cited in many prestigious law reviews as well as in cases argued before state and federal courts, including the United States Supreme Court (as cited in Solum, 2006).

Exploring the risk of digitally born legal information published online

The appeal of digital formats and online publishing to entities responsible for creating and distributing legal information is understandable. Digital materials are more compactly stored, easily transportable, widely distributable and instantly accessible than information produced in any previous format or medium. However, the proliferation of digital information poses an extraordinary challenge for professionals in the fields of law and legal informatics who are concerned about the long-term preservation of legal materials.

Although printed materials are produced in limited quantities and vulnerable to fires, floods and physical deterioration, digitally born materials without printed counterparts are, in fact, at tremendous risk of permanent loss. While the content of a printed item can be read directly by the human eye, accessing materials in digital formats is an indirect process requiring sophisticated retrieval technology. As older technology is rendered obsolete by the emergence of new technology, so, too, are the corresponding older digital formats. Thus, digital information existing in obsolete formats, without the appropriate preservation treatment, can be lost forever (Deegan & Tanner, 2006).

Additionally, a large store of knowledge has accumulated over the years about the lifespan and degradation of printed materials. As a result, archival specialists are able to safeguard both the intellectual content and physical packaging of printed items. The longevity of digital media, however, remains uncertain. For example, compact disks, by one estimation, have a physical lifespan of anywhere between five and 59 years, and a given format's descent into obsolescence, on average, can be expected to occur within five to 20 years (Rothenberg, 1999; Holdsworth, 2006). Whereas materials printed on acidic paper in the nineteenth and twentieth centuries were threatened by an abbreviated lifespan of less than 50 years, the twenty-first century's digitally born information could be rendered inaccessible within five short years of its creation.

Within the realm of digital information, the transient quality of legal information published directly to the free Web (as opposed to within subscription databases), often by government and independent entities, is troubling. Documents, reports and other legal information published online can be unexpectedly and permanently lost as files are removed and URLs are changed or inactivated through routine and seemingly innocuous website maintenance activities. A study by Robert Lopresti and Marcia Gorin (2002) found that among a sample of government agency documents removed from the Web, roughly half had been replaced by newer versions or volumes, while the rest had vanished completely. The non-profit Internet Archive (n.d.), creator of the Wayback Machine Web archive, estimates the average lifespan

of a webpage to be between 44 and 75 days. Yet, the challenges of preserving and providing sustainable access to digitally born legal information are not insurmountable. Tools and standards for best practices in digital preservation have been developed. Digital repositories and Web harvesters are available for a fee or independent development using open-source software. Many libraries have experimented with and implemented this technology within their own institutions.

Statement of the problem and research objective

This article aims to advance the dialogue about the need for successful, sustainable digital preservation programs to ensure ongoing access to our legal heritage, and to advocate for increased collaboration among law libraries in digital preservation efforts. At present, the status of digital preservation activities at state and academic law libraries is unclear. Of particular concern are projects aimed at preserving digitally born legal information published on the free Web – materials that have been demonstrated to be among the most ephemeral and at-risk component of the United States' legal heritage. The objective of the present project is twofold: first, to survey the United States state and academic law library community about its digital preservation efforts, and second, to profile a collaborative, multi-institution project currently underway to preserve digitally born legal information. The digital preservation survey seeks to answer the following questions:

- (1) To what extent are law libraries participating in digital preservation activities?
- (2) What are the law library community's attitudes and perceptions about digital preservation and the vulnerability of digitally born legal materials?
- (3) To what extent are law libraries collaborating in digital preservation projects and programs?
- (4) What factors are limiting or fostering law libraries' participation in digital preservation activities?

In addition, a case study is presented of the Chesapeake Project: a collaborative pilot digital preservation program being implemented by the State Law Library of Maryland, State Law Library of Virginia and the Georgetown Law Library, under the auspices of the Legal Information Preservation Alliance (LIPA). The pilot project is focused solely on harvesting and preserving legal information currently available in digital formats on the free Web, with the goal of evolving into a nationwide preservation program for digitally born legal materials. The Chesapeake Project is explored as a collaborative effort that may be replicated by other institutions and libraries tackling the challenges of digital preservation.

Literature review

Introduction

To provide a cursory overview of current digital preservation strategies, and also to determine which entities in the United States have taken responsibility for the preservation of various categories of digitally born legal information, a comprehensive review of both legal and non-legal literature, including the library and information science literature, was conducted. Searches using academic resources

and finding aids (such as library catalogs and article indexes), full-text database searches and search-engine queries were conducted. In addition to these sources, articles and reports of interest appearing on subscription electronic mailing lists and online news sites, which linked to primary and secondary source materials, were also tracked, allowing for the analysis of works authored by private and governmental institutions.

Digital preservation overview

A number of digital preservation strategies have been introduced and explored since the issue emerged. Kevin Bradley (2007), in a comprehensive article exploring contemporary notions of digital sustainability, affirms that migration and emulation continue to represent two of the most viable digital preservation strategies. Migration is a process by which information is successively copied from old formats onto different or newer formats, thereby ensuring the content remains accessible by staying one step ahead of technology obsolescence. Emulation is a preservation strategy involving the development of new programs and applications that emulate, or mimic, the functionality of outdated programs and applications. Information stored in obsolete formats can thus be retrieved using these programs, theoretically allowing the user to access the digital item in its original format, with its original significant properties intact.

Today, most experts in the field have shifted the thrust of their efforts away from mulling over specific digital preservation strategies and toward the active development and implementation of sustainable 'metadata, standards and architecture' (Bradley, 2007, p. 161). Undoubtedly, migration, emulation or other to-be-developed preservation actions will be taken to ensure ongoing access to preserved items. However, many unforeseen factors and technologies will impact these future decisions. In the meantime, digital items are being archived in standard-compliant digital repositories alongside the appropriate technical, structural, administrative and descriptive preservation metadata, which will guide the extraction and rendering of the archived digital objects for whatever future preservation action is taken.

Legal information and collaborative digital preservation efforts

Archiving and preserving digitally born legal materials for sustainable, permanent access requires a commitment of staffing, infrastructure and budget monies. It is important to understand that the successful preservation of these materials depends upon the stable, ongoing commitment of law libraries and institutions themselves, just as much it depends upon their technological assets and digital repository systems. In fact, one of the core requirements for digital archives put forth by the Center for Research Libraries (2007) is a 'demonstrat[ed] organizational fitness (including financial, staffing structure and processes) to fulfill its commitment'. The costs associated with digital preservation are high, which can threaten the stability and long-term viability of a digital archiving program. As such, it is worthwhile for law libraries to collaborate in digital preservation activities in order to share costs, workloads and overall responsibility for sustainable access to digitally born legal materials.

There has been a movement to encourage governmental, organizational and institutional collaboration in the preservation of legal materials in digital formats. In 2000, the United States Congress established the National Digital Information

Infrastructure and Preservation Program (NDIIPP), which is administered by the Library of Congress, and authorized up to US\$100 million to fund the initiative, with the goal of establishing a collaborative 'national network of entities committed to digital preservation and that are linked through a shared technical framework' (LeFurgy, 2005, p. 164). In 2003, the Legal Information Preservation Alliance, formed under the auspices of the American Association of Law Libraries, was established to begin to address the need for a collaborative national agenda for the preservation of legal literature, in both print and digital formats, by supporting and advancing such efforts (Legal Information Preservation Alliance, n.d.b).

Yet, despite these strides toward collaborative digital preservation programs, two recently published studies, one by NASCIO (2007) and the other sponsored by the Center for Technology and Government (Pardo et al., 2006), found discouraging levels of fragmentation, inconsistency and lack of standardization in the digital preservation of state government information. Each report advocated increased institutional collaboration, on a regional and a nationwide scale, as a possible solution to ensure the sustainability of digital preservation programs.

Who is preserving digitally born legal information?

The professional literature reflects a variety of digital preservation projects, varying in size and approach, aimed at providing sustainable access to various types of legal information. Articles discovered were reviewed to discern the type of legal materials preserved (primary or secondary), the type of institution preserving the material and the preservation strategy implemented. Special attention was also paid to project-oriented collaborative efforts involving federal government institutions, academic libraries, law libraries, and state libraries and archives.

Regarding the preservation of legal publications produced by the United States Government, the Government Printing Office (GPO) is required by Title 44 of the United States Code to provide permanent public access to information and documents created and funded by the federal government. In an electronic world, achieving such permanent public access necessitates digital preservation, and Title 44 specifically mandates that the GPO maintain a directory, a system of access and an electronic storage facility for federal electronic information (Lyons, 2006). In 2003, the GPO and the National Archives and Records Administration (NARA) entered into an agreement whereby the GPO was made a NARA archival affiliate, designating NARA as the legal custodian of the content made available through GPO Access, the GPO's free public website, while conferring physical custody, permanent public access and preservation responsibilities upon the GPO (US Government Printing Office, 2003). However, a preservation challenge for the GPO persists in the form of fugitive documents, elusive government publications published directly to the Web, which have not been cataloged, indexed or archived by the GPO. As of 2003, fugitive documents were conservatively estimated to comprise 'about 50% of the universe of Federal printing' (Baldwin, 2003). Currently, the GPO is working toward the release of the Federal Digital System (FDsys), an impressive new initiative that 'will allow federal agencies to easily create and submit content that can then be preserved, authenticated, managed and delivered upon request' (US Government Printing Office, n.d.). This system is expected to minimize the problem of fugitive documents and is being released in increments, with its first public release planned for late 2008.

The GPO has also entered into formal agreements with libraries and government agencies to ensure permanent public access to electronic publications (Kumar, 2006). For example, the GPO, in collaboration with the United States Commission on Civil Rights, has conferred upon the University of Maryland's Thurgood Marshall Law Library the responsibility to store and ensure permanent public access via the Internet to an electronic collection of historic United States Commission on Civil Rights documents held in the law library's collections (Commission on Civil Rights, 2007). A different type of collaborative effort between the GPO and an academic library is described by Atifa Rawan and Cheryl Knott Malone (2006) in their account of a University of Arizona Library project implemented in conjunction with the GPO's Library Programs Service, which involved the identification of digital government documents by way of a 'virtual depository model'. The GPO acted as a partner in the program by providing legal and organizational support, as well as indexing services for the project, while resolving issues relating to fugitive documents and broken permanent URL (PURL) links discovered by University of Arizona Library.

In early 2006, Yale University Library's Government Documents and Information Center (GDIC) librarians initiated an independent pilot project to migrate at-risk government information recorded on more than 3000 CD-ROMs acquired through the GPO's Federal Depository Library Program (FDLP) to a stable server environment, along with preservation of metadata records (Gano & Linden, 2007). The migration process was found to be costly and time-consuming, resulting in the successful migration of only 13 CD-ROMs. Project coordinators concluded that addressing the challenge of ensuring the preservation and long-term accessibility of government information distributed on legacy digital formats could not be addressed by a single institution alone. However, through a collaborative effort, FDLP libraries could take responsibility for the preservation of materials 'suited to each institution's interest and expertise' while avoiding duplication of efforts (Gano & Linden, 2007).

The University of North Texas Libraries have played an active role in the preservation of and provision of access to federal and state government information, notably with the CyberCemetery collection, which provides permanent public access to the websites and publications of defunct United States Government agencies and commissions (Glenn, 2007; Murray & Phillips, 2007; University of North Texas Libraries, n.d.a). As part of the Federal Depository Library Program, the University of North Texas Libraries created this archive in partnership with the GPO, and it has been cited in the literature as a successful effort in the area of preserving and ensuring access to online government information that would have otherwise been lost (Glenn, 2007; Murray & Phillips, 2007). Another worthy permanent public access project is the University of North Texas Libraries' Congressional Research Service (CRS) Reports collection, comprising CRS reports, which are not made readily available to American citizens by CRS. The reports in this collection are harvested from the Web, archived and then made freely accessible via the University of North Texas website (University of North Texas Libraries, n.d.b).

The Library of Congress-administered National Digital Information Infrastructure and Preservation Program (NDIIPP) has established projects in collaboration with university libraries, state libraries and archives, consortia, nonprofit electronic archiving services and educational organizations to address the preservation needs of a number of priority, at-risk digital materials, including digital geospatial data produced by state and local government entities, political and cultural websites, noncommercial foreign- and American-based public television programming, social

science datasets, electronic scholarly journals, historic archival materials and state government information (LeFurgy, 2005; Library of Congress, n.d.a).

Valerie Glenn (2007) describes the NDIIPP-funded Web-at-Risk project, which involves harvesting and then preserving digitally born legal information published online. The project is led by the California Digital Library, in collaboration with partners at the University of North Texas and New York University, and is developing and providing tools and services allowing librarians to 'capture, curate and preserve Web-based government and political information' (Glenn, 2007; California Digital Library, n.d.; Web-at-Risk, n.d.). NDIPP has also launched a Multistate Government Digital Information Program as a collaborative 'effort to facilitate the development of digital preservation partnerships among states and territories', involving libraries, archives, records management units and other information agencies (Pardo et al., 2006, p. 1; Library of Congress, n.d.b).

In addition to its administration of NDIIPP projects, the Library of Congress has also been involved in the development of a related series of thematic Web Capture collections, involving the acquisition and preservation of websites of historical significance (Library of Congress, n.d.c). Many of the Library's Web Capture collections are highly relevant to legal scholars; these include collections of archived legal blogs, election websites, a 11 September Web archive, websites of members of Congress and congressional committees, online reports relating to the crisis in Darfur, websites relating to the nomination and appointment of Supreme Court justices, and online reports and commentary relating to Hurricane Katrina.

Among academic law libraries, the Cornell Law Library has played an active role in the preservation of website-based legal information. Claire Germain (2002) advocates the use of Web mirror sites as a relatively inexpensive option for the preservation of and long-term access to official legal information published online. In collaboration with producing organizations and governmental entities, entire collections can be loaded onto library servers and archived at regular intervals. The Cornell Law Library has used this strategy to create open-access Web mirror sites for the International Labour Organization and International Court of Justice, which issue court decisions, treaties and other primary documents (Germain, 2002).

Many digital preservation projects involving legal materials focus on the preservation of primary and government resources. In response to this trend, some law librarians have called for academic law libraries to embark upon preserving secondary sources in the form of digitally born legal scholarship – especially online sources cited by scholars and identified only by URLs in the legal literature. Dana Neacsu (2007), building upon her previous work, as well as other librarians' scholarship on the new challenges posed by Web citations, has offered up potential solutions to this issue, which involve preservation and access systems (Neacsu, 2002; Rumsey, 2002). Although librarians and the law library literature are increasingly advocating the preservation of digitally born sources, academic law libraries themselves appear hesitant to devote institutional resources to such projects involving the preservation of this body of legal scholarship and secondary sources of law.

Survey

Introduction and background

In an effort to gain insight into law libraries' digital preservation activities, preservation priorities, attitudes and perceptions about digital preservation, and

collaborative digital preservation efforts, the authors devised an online survey and distributed it to selected members of the law library community. While the literature demonstrates a general awareness among the library community of digitally born materials' heightened risk of loss, as well as the need for preservation strategies to ensure sustainable access to digital materials, there remains some uncertainty as to how individual law libraries are addressing these challenges, what factors have impacted their activities and what they perceive their roles to be in this area.

In 2005, the Legal Information Preservation Alliance (LIPA) sponsored a survey of the state and academic law library community about their digital and analog preservation efforts. Of 41 respondents, only 19 had been involved in preservation projects involving legal materials in any format. Six respondents indicated that their institutions were involved in archiving digital publications that had been downloaded from the Internet. However, in describing their preservation strategy, five indicated that these digital documents were downloaded, printed, bound and added to their physical collections rather than stored in a digital archiving system (Breeze, 2005).

Several issues have emerged as factors inhibiting libraries' efforts in digital preservation, but primary among these is a lack of resources – both human and financial. The Center for Technology in Government in 2006 sponsored a baseline survey on the status of institutional efforts to preserve digital state government information, and found that the greatest barrier to participation in digital preservation activities was a lack of staffing and funding (Pardo et al., 2006). Beyond the law library community, a digital preservation survey by the Museums, Libraries and Archives Council found that 'no long-term funding is being provided for the management of digital material' in institutions based in the United Kingdom (Simpson, 2005, p. 1). A similar study of libraries, archives and museums in the United States, conducted by the Northeast Document Conservation Center, found that a large number of surveyed institutions had 'no or low levels of institutional funds allocated for creation, acquisition, management or sustainability of digital collections' (Clareson, 2006). In addition to lack of funding and staffing, these surveys have demonstrated that a dearth of expert personnel who are trained in digital preservation has limited institutions' ability to participate digital preservation activities (Clareson, 2005; Simpson, 2005; Kenney & Buckley, 2005; Pardo et al., 2006). Moreover, it has been shown that, in some cases, digital preservation is neglected because it simply is not an institutional priority (Kenney & Buckley, 2005).

When it comes to the issue of preserving legal information specifically, there is a broad perception that the government, rather than law libraries, should be involved in such projects; after all, it is the government's mandate to publish the law and make it available to the public in a liberal democracy, where the rule of law requires the public to be informed (Pardo et al., 2006). The GPO has been involved with impressive efforts to ensure perpetual access to federal government information in the United States. Yet, the library's traditional role encompasses collecting information, making that information accessible to users and preserving the information in its collections for future access. As traditional stewards of information, the Library of Congress and other libraries throughout the United States have implemented some of their own projects in protecting primary as well as some secondary sources. Yet, how exactly do state and academic law libraries view their roles in this enterprise? The preservation of which materials, primary or secondary, if any, and in which formats, print or digitally born, do state and academic law libraries believe to be within their institutional role? The authors

surmised that law library digital preservation projects have been more focused on the preservation of digitized print materials, as opposed to digitally born materials, and that few law libraries were actively harvesting content from the Web. Moreover, it was suspected that the type of material (primary, as opposed to secondary legal sources) made a policy difference in terms of digital preservation priorities.

Methodology

Research questions

The survey sought to address the following research questions:

- (1) To what extent are law libraries participating in digital preservation activities?
- (2) What are the law library community's attitudes and perceptions about digital preservation and the vulnerability of digitally born legal materials?
- (3) To what extent are law libraries collaborating in digital preservation projects and programs?
- (4) What factors are limiting or fostering law libraries' participation in digital preservation activities?

Survey instrument

The authors developed a rather detailed, mid-length, Web-based questionnaire (see Appendix). It included 45 questions, which were designed to gather data related to law libraries' digital preservation activities, policies and priorities. The survey, titled 'Law Libraries & Digital Preservation: A Survey', was divided into the following five sections:

- (1) Welcome (a brief introduction and instructions)
- (2) Demographic information (4 questions)
- (3) Digital-preservation activities (23 questions)
- (4) Perceptions and attitudes toward digital preservation (8 questions)
- (5) Copyright and access policies for archived digital materials (10 questions)

The survey was created and made available online using Survey Monkey: survey software that was also used by the Center for Technology in Government's baseline survey of the preservation of state government information (Pardo et al., 2006). This survey development and distribution service was chosen after experimenting with alternate survey formats, including a PDF form designed to be submitted electronically, which was determined by a trial user in the target population to be too unwieldy and difficult to submit.

An assortment of question types were used in the survey: four multi-item Likert questions, 18 close-ended questions and 23 open-ended questions. It is worth noting that participants were not required to complete all survey questions, and that many of the open-ended questions included instructions asking for brief responses. Most open-ended questions were optional and designed to give respondents an opportunity to explain their answers or reactions to the preceding question. Before

issuing the survey and publishing it online, the authors sought feedback on its content, design and usability from other law library professionals. An academic law library director and a law library government affairs representative reviewed the survey instrument, and their recommendations were incorporated into the final survey prior to distribution.

Survey distribution

State and academic law library directors comprised the survey's target population. This group was targeted in an effort to obtain the perspective of upper-level administrators and individuals with insight into the budgetary, organizational and policy issues that impact digital preservation decisions. Moreover, the authors sought to avoid duplication of survey responses, in which more than one respondent reports on the activities of a single law library. A brief introductory message containing the Web survey's hyperlinked URL was distributed via two separate electronic mailing lists, or listservs: the State Law Librarians Roundtable (scclslr@aallnet.org) and LawLibDir (lawlibdir@lists.washlaw.edu). The State Law Librarians Roundtable reaches 42 state law library directors who are members of the American Association of Law Libraries' Special Interest Section of State, Court, & County Law Libraries. LawLibDir, operated by Washburn University Law School, reaches a subscription base of 230 academic law library directors. The authors did not distribute the surveys directly; rather, they enlisted the assistance of colleagues who were listserv subscribers to post the survey on their behalf. The survey was published online and available for respondents to complete between the dates of 29 February 2008, and 20 March 2008. It was posted to the State Law Librarians Roundtable on 29 February 2008, and to LawLibDir on 4 March 2008.

Response rate

The survey was thus made available online to 272 recipients. A total of 39 recipients accessed the survey (i.e., they followed the URL to the survey's location on the Web); of these, however, 37 submitted surveys were determined to be valid. As a result, the survey response rate was low (13.6%). Thus, though the sample size is a weakness of the study, it is worth noting that a similar survey of digital preservation conducted via e-mail by the Northeast Document Conservation Center reported a comparable response rate of 12.5%.

Among submitted surveys, the types of law libraries represented were equally split: 18 academic law libraries and 18 state law libraries (see Figure 1). One additional respondent was identified as an 'other' law library, and the respondent declined to give additional information regarding his/her institution. Considering each listserv's subscription numbers, the academic law library response rate translates to 7.8%, while the state law library response rate stands at 42.9%.

Results

Law libraries and digital preservation activities

Based on the reports of 21 survey respondents, representing 11 academic law libraries, nine state law libraries and one 'other' library, a cumulative total of about

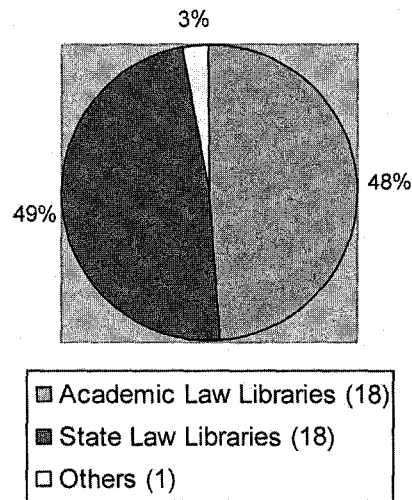


Figure 1. Survey respondents by library type.

59 digital preservation projects have been planned or implemented by respondents' libraries within the past five years. Responses varied, with nine respondents, representing four academic law libraries, three state law libraries and two 'other' libraries, indicating that their institutions had been involved in no digital preservation activities at all. Four academic law libraries indicated their involvement in one project, while four state law libraries and one academic law library indicated involvement in two digital preservation projects during the past five years. An academic law library and a state law library reported participation in three and five projects, respectively, and another state law library respondent indicated involvement in 12 projects over the stated period. Finally, an academic law library respondent reported participation in between 20 and 30 digital preservation projects (which was estimated to be 25 projects for the purpose of this study).

In assessing the extent to which respondents' efforts focused on the preservation of digitized, or scanned, print materials, as opposed to digitally born materials with no print original, it was determined that 55 of the 59 digital preservation projects involved digitized print materials. Only four projects involved digitally born materials, and out of 15 respondents, 13 indicated that they were not harvesting and preserving Web-published materials. Table 1 provides an overview of respondents' digital preservation projects involving digitized print and digitally born materials.

Of the total number of 18 state law library respondents, only two reported participation in a total of three projects to preserve digitally born materials, while five (27.8%) respondents indicated their involvement in a total of 22 projects to preserve digitized print materials. Of the 18 academic law library respondents who submitted responses to the survey, only one reported participation in a single project to preserve digitally born materials, while eight (44.4%) respondents indicated their involvement in a total of 33 projects to preserve digitized print materials.

With regard to law libraries' emphasis on the preservation of primary versus secondary legal materials, three academic law library respondents indicated that their institutions preserved primary domestic law, while one preserved both primary and secondary materials, whether domestic, foreign or international. Among state

Table 1. State and academic law library respondents' digital preservation projects involving digitized print and digitally born materials.

	Total projects	Projects by academic law libraries	Projects by state law libraries
Total number of digital preservation projects	59	34	25
Digital preservation projects involving digitized (scanned) print materials	55	33	22
Digital preservation projects involving digitally born materials	4	1	3

law library respondents, three respondents, representing all who answered the question, reported preserving domestic law materials in digital formats.

Regarding copyright and access issues, a total of eight academic law library respondents answered the questions on this topic. Of these respondents, two reported that their institutions did not preserve materials in the public domain and therefore always sought copyright permission before including protected materials in digital preservation projects. Two respondents indicated that they sometimes obtained permission, and four respondents never sought such permission, even though one respondent was involved in the preservation of non-public domain materials. Of the six state law library respondents involved in digital preservation projects, five responded to the questions on copyright and access policies. Of these respondents, two indicated that copyright permission to include protected materials in digital preservation projects was obtained, while three indicated that it was not. Four respondents indicated that their institutions preserved materials that were in the public domain.

Respondents report different approaches to providing access to preserved digital materials in their collections. Out of seven academic law library respondents, four libraries make their archives accessible online to authenticated patrons, two provide onsite-only access to library patrons and one has no current access system but is working on an open-access Web interface. One academic law library respondent distinguishes between the methods of access for copyrighted and public-access materials, allowing online access to the latter. Three state law library respondents reported making preserved digital material fully accessible to the public via the Web, while two respondents provide access onsite to in-library patrons only.

Attitudes and perceptions about digital preservation in the law library community

Survey respondents were asked to indicate their level of agreement on a Likert scale to a series of four statements regarding the types of digital legal materials that law libraries should be preserving. Table 2 provides respondents' rankings for each item. Among eleven academic law library respondents, four strongly agreed that law libraries should be involved in the preservation of law-related information published on the Web, with one respondent adding government information as well as information published on the Web and cited in law review articles. Three academic law library respondents agreed that law-related information published on the Web should be preserved, while two agreed that law libraries should be involved in the preservation of law review articles published digitally within subscription databases. One respondent expressed disagreement with the statement that law libraries should

be involved in the preservation of digitally born government information, and one respondent also did not agree that law libraries should be responsible for articles and materials made available through subscription databases.

Among seven state law library respondents, all seven agreed or strongly agreed that law libraries should be involved in the preservation of law-related information published on the Web. Six respondents agreed or strongly agreed that law libraries should be involved in the preservation of law review articles published digitally within subscription databases, as well as the preservation of digitally born government information and Web-published legal information cited in law review articles.

Although the vast majority of digital projects implemented by survey respondents involved the preservation of digitized print, as opposed to digitally born, materials, respondents indicated by a two-to-one margin that they believed that digitally born materials were more in need of digital preservation efforts than print materials (see Table 3). Respondents advocating for the preservation of

Table 2. State and academic law library respondents indicated their level of agreement to a series of statements regarding the types of digital legal materials that law libraries should be preserving (% in parentheses).

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Response count
<i>Law libraries should be involved in preventing the loss of law-related information published to the Web</i>						
Total	0 (0.0)	0 (0.0)	4 (22.2)	8 (44.4)	7 (38.9)	18
Academic	0 (0.0)	0 (0.0)	4 (36.4)	4 (36.4)	4 (36.4)	10
State	0 (0.0)	0 (0.0)	0 (0.0)	4 (57.1)	3 (42.9)	7
<i>Law libraries should be involved in the preservation of digitally born government information</i>						
Total	0 (0.0)	1 (5.9)	3 (17.6)	9 (52.9)	4 (23.5)	17
Academic	0 (0.0)	1 (9.1)	3 (27.3)	6 (54.5)	1 (9.1)	10
State	0 (0.0)	0 (0.0)	1 (14.3)	3 (42.9)	3 (42.9)	7
<i>Law libraries should be involved in preventing the loss of information published on the Web and cited within law review articles</i>						
Total	0 (0.0)	0 (0.0)	6 (33.3)	8 (44.4)	4 (22.2)	18
Academic	0 (0.0)	0 (0.0)	5 (45.5)	4 (36.4)	2 (18.2)	10
State	0 (0.0)	0 (0.0)	1 (14.3)	4 (57.1)	2 (28.6)	7
<i>Law libraries should be involved in the long-term preservation of and sustained access to law review articles and other legal materials published digitally within subscription databases (HeinOnline, LexisNexis, Westlaw, etc.)</i>						
Total	0 (0.0)	1 (5.6)	4 (22.2)	9 (50.0)	4 (22.2)	18
Academic	0 (0.0)	1 (10.0)	2 (20.0)	5 (50.0)	2 (20.0)	11
State	0 (0.0)	0 (0.0)	1 (14.3)	4 (57.1)	2 (28.6)	7

Table 3. State and academic law library respondents' answers to the question: 'Which materials, in your opinion, deserve more attention when it comes to preservation?'

Which materials are in greater need of preservation?	Total	Academic	State
Print materials	6	4	2
Digitally born materials	12	6	6

digitally born materials overwhelmingly expressed, in response to an open-ended question, an understanding of the risk posed to these materials due to obsolescence.

Digital preservation and collaboration in the law library community

Regarding survey respondents' levels of participation in collaborative digital preservation projects, none indicated that they always collaborated with digital preservation partners; eight respondents indicated that they sometimes or almost always collaborated with other libraries and institutions; three indicated that they almost always collaborated with nonprofit partners; and four indicated that they sometimes collaborated with corporate or for-profit partners. Five respondents indicated that they rarely or never collaborated with other libraries or institutions, and nine and eight respondents, respectively, answered that they never collaborated with nonprofit or for-profit partners. Survey responses are available in Table 4.

In response to an open-ended question about which institutions and organizations, specifically, respondents had partnered with in digital preservation activities, nine state and academic law library respondents listed the following as collaborators: seven libraries and academic institutions, five for-profit companies, four nonprofit organizations and two state government entities (excluding state libraries and archives). The list of collaborators named by respondents, classified by type, is available in Table 5.

Factors impacting law libraries' digital preservation activities

Survey respondents were asked to indicate their level of agreement, on a Likert scale, as to the extent to which various factors have impacted digital preservation activities at their respective libraries. The first series of Likert items assessed factors limiting libraries' involvement in digital preservation activities (see Table 6). The second series of Likert items assessed factors that would encourage libraries' involvement in digital preservation activities (see Table 7). Lack of funding, staffing shortages, lack

Table 4. State and academic law library respondents' answers to the question: 'To what extent do you collaborate with other institutions and/or nonprofit/for-profit partners in developing your digital projects?' (% in parentheses).

	Never collaborate	Rarely collaborate	Sometimes collaborate	Almost always collaborate	Always collaborate	Response count
<i>Collaborate with libraries/institutions</i>						
Total	4 (30.8)	1 (7.7)	6 (46.2)	2 (15.4)	0 (0.0)	13
Academic	4 (57.1)	1 (14.3)	1 (14.3)	1 (14.3)	0 (0.0)	7
State	0 (0.0)	0 (0.0)	4 (80.0)	1 (20.0)	0 (0.0)	5
<i>Collaborate with nonprofit partners</i>						
Total	9 (75.0)	0 (0.0)	0 (0.0)	3 (25.0)	0 (0.0)	12
Academic	5 (88.3)	0 (0.0)	0 (0.0)	1 (16.7)	0 (0.0)	6
State	3 (60.0)	0 (0.0)	0 (0.0)	2 (40.)	0 (0.0)	5
<i>Collaborate with for-profit partners</i>						
Total	8 (66.7)	0 (0.0)	4 (33.3)	0 (0.0)	0 (0.0)	12
Academic	3 (50.0)	0 (0.0)	3 (50.0)	0 (0.0)	0 (0.0)	6
State	4 (80.0)	0 (0.0)	1 (20.0)	0 (0.0)	0 (0.0)	5

Table 5. Entities that have collaborated in the digital preservation initiatives of nine survey respondents.

Named collaborators	Total
<i>Libraries/academic partners</i>	7
Academic/law libraries	1
Library consortia	2
State libraries/archives	2
Law schools/universities	2
<i>For-profit partners</i>	5
Publishers	4
Google	1
<i>Nonprofit/organization partners</i>	4
Heritage/historical societies	2
Legal institutions	1
Professional organizations	1
<i>State government partners</i>	2
State committees	1
State supreme courts	1

of staff with expertise in digital preservation and institutional lack of interest in digital preservation were all cited as limiting factors, and, not surprisingly, additional funding, additional staff and the recruitment of staff with digital preservation expertise were all cited as factors that would encourage increased participation in digital preservation activities at the institutional level.

Although most respondents neither agreed nor disagreed that lack of opportunities for collaboration and lack of a large-scale digital preservation movement in which to participate has limited their activities, an increased number of respondents indicated agreement or strong agreement that increases in opportunities for collaboration and the emergence of a large-scale digital preservation movement in which to participate would encourage more involvement in digital preservation activities at their libraries.

Discussion

While the low rate of response precludes broad generalization of the data, these survey findings nonetheless build upon an emerging area of academic interest in the field law librarianship and are worth consideration. Based on survey responses, the majority of state and law libraries' digital preservation projects have involved the preservation of digitized, or scanned, print materials, as opposed to the preservation of materials that are digitally born; out of 59 reported digital preservation projects, only four involved digitally born items. Yet, paradoxically, when asked about which material types were in greater need of preservation, print materials or digitally born materials, respondents replied that digitally born materials were in more urgent need of preservation by a margin of 2 to 1.

In response to questions about issues impacting their institutions' level of involvement in digital preservation activities, lack of funding, staffing shortages and lack of staff with technological or digital preservation expertise were cited as limiting factors. Not surprisingly, it was widely reported that increased funding, staffing and

Table 6. State and academic law library respondents indicated their level of agreement as to the extent to which various factors have limited involvement in digital preservation activities at their respective institutions (% in parentheses).

My library's level of involvement in digital preservation activities has been limited by:	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Response count
<i>Lack of funding</i>						
Total	0 (0.0)	3 (16.7)	3 (16.7)	6 (33.3)	6 (33.3)	18
Academic	0 (0.0)	3 (27.3)	1 (9.1)	3 (27.3)	4 (36.4)	11
State	0 (0.0)	0 (0.0)	2 (28.6)	3 (42.9)	2 (28.6)	7
<i>Concerns about technology/file format obsolescence</i>						
Total	1 (5.6)	4 (22.2)	5 (27.8)	7 (38.9)	1 (5.6)	18
Academic	0 (0.0)	4 (36.4)	2 (18.2)	4 (36.4)	1 (9.1)	11
State	1 (14.3)	0 (0.0)	3 (42.9)	3 (42.9)	0 (0.0)	7
<i>Staffing shortages</i>						
Total	0 (0.0)	2 (11.1)	1 (5.6)	10 (55.6)	5 (27.8)	18
Academic	0 (0.0)	2 (18.2)	1 (9.1)	6 (54.5)	2 (18.2)	11
State	0 (0.0)	0 (0.0)	0 (0.0)	4 (57.1)	3 (42.9)	7
<i>Lack of staff with digital preservation/technological expertise</i>						
Total	0 (0.0)	3 (16.7)	3 (16.7)	10 (55.6)	2 (11.1)	18
Academic	0 (0.0)	2 (18.2)	2 (18.2)	7 (63.6)	0 (0.0)	11
State	0 (0.0)	1 (14.3)	1 (14.3)	3 (42.9)	2 (28.6)	7
<i>Digital preservation is not an institutional priority</i>						
Total	1 (5.6)	4 (22.2)	4 (22.2)	8 (44.4)	1 (5.6)	18
Academic	0 (0.0)	1 (9.1)	3 (27.3)	6 (54.5)	1 (9.1)	11
State	1 (14.3)	3 (42.9)	1 (14.3)	2 (28.6)	0 (0.0)	7
<i>Lack of partners/opportunities to collaborate with other libraries and organizations in digital preservation activities</i>						
Total	1 (5.6)	5 (27.8)	10 (55.6)	2 (11.1)	0 (0.0)	18
Academic	0 (0.0)	3 (27.3)	7 (63.6)	1 (9.1)	0 (0.0)	11
State	1 (14.3)	2 (28.6)	3 (42.9)	1 (14.3)	0 (0.0)	7
<i>Lack of an organized statewide/nationwide/international digital preservation movement in which to participate</i>						
Total	1 (5.6)	5 (27.8)	7 (38.9)	3 (16.7)	2 (11.1)	18
Academic	0 (0.0)	5 (45.5)	6 (54.5)	0 (0.0)	0 (0.0)	11
State	1 (14.3)	0 (0.0)	1 (14.3)	3 (42.9)	2 (28.6)	7

recruitment or cultivation of well-trained staff would encourage greater involvement in digital preservation activities. While most respondents indicated that lack of collaborative opportunities or large-scale movements in which to participate had not necessarily limited their digital preservation activities, more than half of the respondents agreed that increased opportunities to collaborate with other institutions or within the context of a large-scale digital preservation movement would encourage more involvement in digital preservation activities at their libraries. Most likely, there is some recognition among law libraries that collaborative or large-scale digital preservation programs would allow for sharing of resources, including funds and workloads, among participating institutions, reducing the burden placed upon a single institution in implementing a digital preservation program.

Table 7. State and academic law library respondents indicated their level of agreement as to the extent to which various factors would encourage digital preservation activities at their respective institutions (% in parentheses).

The following would encourage greater involvement in digital preservation activities at my library:	Strongly disagree	Disagree	Neutral	Agree	Strongly agree	Response count
<i>Increased funding</i>						
Total	0 (0.0)	1 (5.6)	2 (11.1)	9 (50.0)	6 (33.3)	18
Academic	0 (0.0)	1 (9.1)	2 (18.2)	7 (63.6)	1 (9.1)	11
State	0 (0.0)	0 (0.0)	0 (0.0)	2 (28.6)	5 (71.4)	7
<i>Increased staffing</i>						
Total	0 (0.0)	1 (5.6)	2 (11.1)	8 (44.4)	7 (38.9)	18
Academic	0 (0.0)	1 (9.1)	2 (18.2)	6 (54.5)	2 (18.2)	11
State	0 (0.0)	0 (0.0)	0 (0.0)	2 (28.6)	5 (71.4)	7
<i>Recruitment/cultivation of staff with digital preservation/technological expertise</i>						
Total	0 (0.0)	0 (0.0)	4 (22.2)	9 (50.0)	5 (27.8)	18
Academic	0 (0.0)	0 (0.0)	3 (27.3)	7 (63.6)	1 (9.1)	11
State	0 (0.0)	0 (0.0)	1 (14.3)	2 (28.6)	4 (57.1)	7
<i>Increased opportunity to collaborate with other libraries and organizations in digital preservation activities</i>						
Total	0 (0.0)	1 (5.6)	7 (38.9)	7 (38.9)	4 (22.2)	18
Academic	0 (0.0)	1 (9.1)	7 (63.6)	3 (27.3)	1 (9.1)	11
State	0 (0.0)	0 (0.0)	0 (0.0)	4 (57.1)	3 (42.9)	7
<i>Establishment of an organized statewide/nationwide/international digital preservation movement in which to participate</i>						
Total	0 (0.0)	0 (0.0)	7 (38.9)	7 (38.9)	4 (22.2)	18
Academic	0 (0.0)	0 (0.0)	7 (63.6)	3 (27.3)	1 (9.1)	11
State	0 (0.0)	0 (0.0)	0 (0.0)	4 (57.1)	3 (42.9)	7

Case study

Introduction and overview

While law librarians, as an organized profession, have engaged in discussions and academic dialogues about the need for collaborative efforts to preserve digitally born legal materials, it is a pilot effort being presently implemented in the United States that may serve to provide unique insight into the practical experience of realizing such a program. Commonly in the medical and social sciences, the case study is used as a research tool to gather and present empirical data from quantitative and quasi-experimental designs, often for the purpose of establishing causal relationships; however, the case study that follows diverges from this design and borrows from the educational case studies found in the disciplines of law, business and public policy, which are presented with the intent of impacting public opinion, practice and policy development by describing and raising awareness of specific problem-solving strategies (Yin, 2003). Within the digital preservation community, this type of descriptive case study has been put forward to encourage, support and provide a framework for the establishment of digital preservation programs. For example, the Electronic Resource Preservation and Access Network (ERPANET), funded by the

European Commission and Swiss Confederation, is developing a series of sixty case studies to explore digital preservation programs in various institutions and commercial sectors (Ross, 2004). Presently, nearly forty such studies have been published on the ERPANET website (ERPANET, n.d.).

The study that follows provides an account of the Chesapeake Project, a collaborative pilot digital preservation project that began its archiving activities in March 2007, and investigates the following issues: project origin and background; mission and objectives; selection and collection scope; digital preservation strategies and tools; discovery and access; organizational framework and staffing; project and collection status; and post pilot-phase prospects.

Methodology

In researching the Chesapeake Project for the purpose of developing the present case study, a number of published and unpublished project-created policy documents were consulted and analyzed. In addition to these official project documents, other records consulted included project meeting materials, such as agendas and handouts, correspondence between collaborating project members and vendor documentation. Documents produced by the Legal Information Preservation Alliance, related to the early establishment of the Chesapeake Project, were also examined. It is important for readers to note that one of the two authors of this article is affiliated with an institution participating in the Chesapeake Project, and the other author is not. Working together, the authors have made every attempt to present an unbiased case study. Given this level of affiliation with the project, however, the authors were also privileged with extraordinary access to numerous materials and documents associated with the project, as well as the insight and direct observations of a project participant.

Project origin and background

The Chesapeake Project's origins can be traced to March 2003, when the conference 'Preserving Legal Information for the 21st Century: Toward a National Agenda' was held at the Georgetown University Law Center. The conference convened a select and strategic group of law library directors, law librarians, publishers and experts in the fields of information technology and digital preservation to discuss and set forth a national agenda to prevent the loss of legal information in both analog and digital formats (Legal Information Preservation Alliance, 2003).

In an effort to advance this national agenda, participants formed a new organization called the Legal Information Preservation Alliance (LIPA), with the mission 'to provide the leadership, the necessary organizational framework and the professional commitment necessary to preserve vital paper and electronic legal information by defining objectives, developing and/or adopting appropriate standards and models, creating networks, and fostering financial and political support for long term stability' (Legal Information Preservation Alliance n.d.b). LIPA's membership has increased steadily since its founding in 2003; by September 2005, the association had enlisted 36 institutions, and in January 2008, LIPA's membership had risen to 69, representing the American Association of Law Libraries (AALL) and state and academic libraries throughout the United States (Special Committee on Permanent Public Access to Legal Information, 2005; Legal Information Preservation Alliance, n.d.a).

In June 2006, LIPA finalized its Strategic Plan Outline, which articulated, among other things, the association's priorities in the area of preservation activities. First listed among these activities was the creation 'of a pilot project to preserve born digital materials' (Legal Information Preservation Alliance, 2006, p. 2). In an effort to accomplish this strategic objective, three LIPA-member institutions in the Chesapeake Bay region of the United States – the Georgetown Law Library, the Maryland State Law Library and the Virginia State Law Library – established the Chesapeake Project as a two-year pilot digital preservation program to address the challenge of preserving legal information published directly to the Web. Library directors at the Georgetown, Maryland State and Virginia State Law Libraries began organizing the project and evaluating Web harvesting, digital archiving tools and repository options in 2006. In March 2007, the institutions participating in the pilot began actively harvesting content from the Internet and preserving it within a shared digital repository. As a two-year pilot, the Chesapeake Project is slated to end its pilot phase in 2009. It is anticipated that it will establish a solid framework, body of policy documentation and support structure that will evolve into a nationwide digital preservation initiative by enlisting the participation of state and academic law libraries throughout the United States.

Project mission and objectives

The mission of the Chesapeake Project, as stated in project documentation, is

to successfully develop and implement a pilot program to stabilize, preserve and ensure permanent access to critical born-digital legal materials on the World Wide Web. The Chesapeake Project is working to establish the beginnings of a strong regional digital archive collection of US legal materials as well as a sound set of standards, policies and best practices that could potentially serve to guide the future realization of a nationwide preservation program. (Chesapeake Project, 2007, p. 2)

Beyond this broadly stated mission statement, participants in the Chesapeake Project, upon its inception, did not create a list of specific benchmarks or strategic objectives. As a pilot project, the Chesapeake Project is primarily an investigative effort to lay the foundation for a larger collaborative program. As stated in the Chesapeake Project's first-year evaluation document:

[P]roject participants have utilized the first year of the pilot to familiarize themselves with the digital archiving process, create shared documentation to guide project participation, assess digital-archiving costs and necessary staffing commitments, and develop reasonable expectations for progress in digital archiving and archive collection development. (Chesapeake Project, 2008, p. 4)

Although specific project benchmarks were not set for the project's first year, evaluation parameters for the project were established early in the pilot, and participants set up a formal project evaluation schedule, with evaluations being conducted at the pilot's one- and two-year marks. First-year evaluation measures included: a count of items and titles archived during the first year; access statistics; a test sample to determine the number of archived titles altered or removed from their original locations on the Web; and a qualitative analysis of the project's progress and challenges (Chesapeake Project, 2008). Second-year evaluation measures are expected to consider user feedback.

Selection and collection scope

Because there are no in-print lists or approval plans generated to guide the acquisition of legal content from the expansive independent publishing medium that is the Internet, selection of materials from the Web for long-term preservation can be a challenge. Due to the two-year time limitation of the pilot project, each library participating in the Chesapeake Project implemented a set of selection parameters in an effort to both guide and limit the scope of materials harvested and preserved, and to ensure the development of a cohesive collection of preserved legal materials. The selection of materials from the Web for preservation by institutions participating in the Chesapeake Project is driven by the following factors: the users of each library and their information needs; the institutional missions, mandates and priorities of each library; the perceived risk level of the material; and the collection scope and parameters determined, individually, by each institution.

The library patrons at each institution comprise the primary user group, which impacts the digital-archive selection decisions and collection scopes of participating libraries. These patron groups range from law students at Georgetown to state law practitioners in the states of Maryland and Virginia; they are described in project documentation as 'law practitioners, law faculty members, law students, justices and their staff members, judges and their staff members, and state government officials and their staff members' (Chesapeake Project, 2007, p. 3). The Chesapeake Project also aims to serve the legal information needs of a broader secondary user group comprising law students, scholars and practitioners who are not affiliated with the Georgetown, Maryland State or Virginia State Law Libraries, in addition to general public users with legal information needs.

The Maryland and Virginia State Law Libraries digital-archive collections consist primarily of state-issued materials, as well as some community- and organization-published reports and studies. The Maryland State Law Library specifically selects items 'that describe, analyze, document, propose, clarify or define public-policy and legal issues that affect the citizens of the state of Maryland' (Chesapeake Project, 2008, p. 3). Within the first year of its participation in the Chesapeake Project, the Maryland State Law Library has collected and preserved what may be the most comprehensive collection of Maryland General Assembly-mandated task force reports available online (Chesapeake Project, 2008). The digital-archive collections of the Virginia State Law Library represent the online publications of the state's judicial branch of government, including those of the Supreme Court of Virginia and the Judicial Council of Virginia. As an academic law library, the Georgetown Law Library's digital archive collections are largely thematic and include secondary legal materials based on scholarly areas of interest and the established legal research institutes at the Georgetown Law Center. Additionally, the library collects jurisdictional materials by and about local and neighboring government entities and a limited number of reports from federal commissions. The library also works with the Law Center's Office of Journal Administration to archive Web-based sources that are cited in legal journals and fit within the established project collection scope.

While some Web-harvesting projects focus on the capture and preservation of entire Web sites, the Chesapeake Project focuses upon the capture and preservation of discrete online publications. If multiple reports are posted on a single webpage, for example, the entire webpage is not harvested; rather, each report is harvested individually and preservation metadata is created to accompany each harvested title.

Although the process of harvesting and archiving individual publications one-by-one, as opposed to entire collections of publications at once, is considerably more time-consuming, project participants believe that this strategy is in the best interests of their users as individual titles can be cataloged and linked to single bibliographic records, facilitating user discovery. Entire websites, on the other hand, would require re-harvesting at regular intervals to capture newly posted content, and facilitating discovery of discrete reports embedded deeply within a harvested, content-rich website would pose a challenge.

Digital preservation strategies and tools

After considering various options for the storage, preservation and management of digital materials, including open-source options, the libraries participating in the Chesapeake Project selected the OCLC Digital Archive, operated and administered by the nonprofit Online Computer Library Center (OCLC). Although open-source digital repository systems represent a less expensive option, they require a staff with significant technological expertise and a designated storage site. As such, project participants chose to utilize a vendor-operated system. A number of factors influenced the choice of OCLC Digital Archive. The archive's storage system adhered to the ISO reference model for an Open Archival Information System (OAIS), which is the standard conceptual framework accepted by the digital preservation community for the permanent preservation of digital information. OCLC's prominence and stability in the library community also impacted this choice as the long-term viability of any digital preservation project demands a digital repository backed by a sound organizational structure. Moreover, OCLC was willing to work with the Chesapeake Project to negotiate a shared trial pricing structure for the term of the two-year pilot phase.

The Chesapeake Project utilized the OCLC Digital Archive's bit-level preservation services, ensuring that the digital files deposited into the archive remained uncorrupted and renderable in their original formats. OCLC's responsibilities included secure onsite storage of archived items at OCLC facilities, maintaining multiple copies of backup data and disaster tapes stored at an offsite facility, and a regular schedule of virus-checking, file format verification and fixity-checking using checksum algorithms. There was no explicit preservation action strategy, such as format migration, for items in the OCLC Digital Archive; however, customized preservation treatments would be implemented by OCLC, based on formats archived and institutional needs, to counteract future obsolescence.

All items harvested from the Internet and placed into the OCLC Digital Archive are accompanied by preservation metadata records, which contain information that will ultimately guide preservation action decisions and the future rendering of archived digital objects. The Digital Archive automatically generates and captures technical metadata about each item harvested, including the file format type, which is verified using the JSTOR/Harvard Object Validation Environment, or JHOVE. Project participants play an administrative and curatorial role in the creation of metadata records, manually entering descriptive and administrative metadata into the preservation records. As a quality control measure and to ensure consistence in metadata record creation, project participants consult a project metadata guide, which was developed at the start of the pilot.

Like all things in the digital environment, digital archives and repositories themselves are not immune to the advancement of technology and the threat of

obsolescence. Just as digital files require maintenance and migration, so too do digital archive systems. Within the first year of its pilot phase, the Chesapeake Project experienced this phenomenon first hand. Project participants were informed by OCLC in summer 2007 that the OCLC Digital Archive would be replaced by a new system. In April 2008, shortly after the project's the first-year mark, OCLC transitioned the Chesapeake Project's archived collections and metadata from the original OCLC Digital Archive to a more sophisticated, two-tiered digital-preservation and access system. Whereas the original OCLC Digital Archive acted as both an access and a preservation system, the new system separates access from preservation. Using the new system, two digital objects are created from the original item harvested from the Web: a master file and an access copy. The master file is stored in a dark digital archive, which is very similar to the previous OCLC Digital Archive, except that it is completely inaccessible to users. The derived access copy is imported into CONTENTdm, a customized storage and retrieval system, which makes archived collections accessible to users via a searchable Web interface.

Discovery and access

Discovery of and access to digital collections archived by the Chesapeake Project is made available through participating institutions' local OPACs, the open-access WorldCat.org system, subscription OCLC FirstSearch and WorldCat databases, and the Chesapeake Project's new CONTENTdm system. The bibliographic treatment of each item in the archive is vital to user access to and discovery of the Chesapeake Project's collections. In addition to archiving an item and generating a preservation metadata record, every archived title has a corresponding bibliographic MARC record, created in OCLC's shared global bibliographic database.

As a digital item is harvested from the Web and archived, it is assigned a unique URL that is hyperlinked to the archived access copy in the OCLC system. (Previous OCLC Digital Archive system URLs now resolve to CONTENTdm URLs.) This URL is added to local records and OCLC bibliographic records within an 856 field, alongside the original Web URL, and provides direct access to archived objects. If and when an object's original URL becomes inactive, the URL for the archived access copy will continue to provide access to the title. Any user with an Internet connection can discover these records through traditional catalog searching methods, using a library's OPAC or an OCLC database, and is provided with open access to archived resources via hyperlinked URLs placed prominently within the records. As a digital object is harvested from the Web, attached to a bibliographic record in OCLC and imported into CONTENTdm, so too is the item's bibliographic metadata, which is crosswalked from MARC format into a Qualified Dublin Core record in CONTENTdm. In addition to these Qualified Dublin Core records, the CONTENTdm system facilitates discovery through full-text PDF searchability.

Organizational framework and staffing

The three libraries participating in the Chesapeake Project vary considerably. In addition to the fact that project participants represent two state and one academic law library, each with different patron groups and mandates, the three libraries also differ significantly in terms of size. The Virginia State Law Library is operated by a

staff of five people. The Maryland State Law Library is larger, with a staff of 15, and the Georgetown Law Library, which consists of two separate law library buildings, has a staff of nearly 70. Given these differences, project structure, flexible policies and regular communication were required.

The director of each participating library assists with project planning, upper-level decision-making and strategy, and has appointed a staff librarian to coordinate the library's day-to-day participation in the project and manage project-related curatorial, cataloging and digital archiving tasks. Technical services and cataloging librarians at each institution also assist with the project as needed. The Georgetown Law Library hired a full-time librarian to manage the project, who devotes roughly 30 hours per week to project-related archiving, cataloging and coordination. Two librarians, a project coordinator and a cataloger at the Maryland State Law Library spend a combined amount of 12 hours per week on project-related tasks, and at the Virginia State Law Library, the project coordinator devotes about five hours per week to the project – down from 15 hours per week at the start of the project.

All libraries report that the most time-consuming task associated with the project is cataloging, largely because the majority of the items harvested and archived through the project are fugitive documents or gray literature, and as such, they require original cataloging. Other time-consuming tasks include Web harvesting, archiving and preservation metadata record creation.

A preservation metadata guide was developed early in the project to guide the creation of preservation metadata records for archived digital objects. Soon afterward, the libraries involved in the pilot approved a comprehensive collection plan, which laid out the project's mission and scope, methods of acquisition and selection, metadata policies, methods of access and preservation system. The structure of this collection plan was borrowed from the NDIIPP-sponsored Web-at-Risk project, which has developed and published online a flexible collection plan template to accommodate the various institutions participating in the Web-at-Risk (Web-at-Risk, n.d.). Project participants have continually convened to reassess established policies and update them, as needed, to address newly discovered challenges and the project's evolving circumstances. In addition to regular e-mail updates and discussions, project participants have implemented a formal schedule of quarterly meetings to facilitate communication, discuss project policy and share information. A conference call is set up to allow for the inclusion of any project participants who are unable to attend a quarterly meeting in person.

Project and collection status

In March 2008, at the time of the project's first-year evaluation, 2705 items, representing approximately 1270 titles, had been harvested from the Web and placed within the Chesapeake Project's shared digital archive (Chesapeake Project, 2008). The discrepancy between 'items' and 'titles' is largely due to serial publications, as well as some multi-part monographs, which require multiple, separate harvests, but, comprising a single 'title', are attached to a single corresponding bibliographic record.

An analysis of a random sample of 579 titles archived during the Chesapeake Project's first year demonstrated, with a confidence level of 95% and a confidence interval of ± 3 , that more than 95% of the titles in the archive were published in PDF format; 4% in the sample were published as X/HTML documents; and the

remaining titles were in either Microsoft Word format or multiple formats, such as an HTML publication with embedded supplements in PDF format (Chesapeake Project, 2008).

Project participants tested the same sample of 579 archived titles to determine the number of archived titles altered or removed from their original locations on the Web. This exercise demonstrated that more than 8% of titles harvested from the Web between the project's start date in 2007 to its first-year mark in March 2008 had inactive original URLs, meaning that these items had already been altered, removed from their original locations or deleted from the Web entirely.

The Chesapeake Project's first-year evaluation also included an analysis of access statistics for archived items. Although the project's first-year efforts were marketed to neither users nor other institutions, access figures showed a high level of archived item use, indicating that many users discovered and accessed archived titles through bibliographic records in participating institutions' OPACs, WorldCat.org, and other OCLC bibliographic databases. In the project's first year, archived items were accessed a total of 5317 times. Within this figure, project participants accessed their own archived items a total of 2267 times. Public users, users who accessed archived content through open-access means without first logging into an OCLC system, accounted for a surprisingly high 2528 instances of access, and authenticated OCLC-affiliated libraries and institutions, excluding those participating in the project, accounted for 522 instances of access, most probably occurring during the course of research, reference activities and adding OCLC bibliographic records with archived URLs to their own local catalogs.

Post pilot-phase prospects

The vision of the Chesapeake Project has been articulated as follows:

The Chesapeake Project aims to set a precedent for a national movement to prevent the widespread loss of legal information in digital formats, securing these materials for generations to come. Upon reaching the close of its two-year pilot phase in 2009, The Chesapeake Project hopes to help inspire, establish and galvanize widespread participation in a comprehensive, collaborative and nationwide preservation program for legal resources. (Chesapeake Project, 2007, p. 3)

It is important to remember that the Chesapeake Project is a two-year pilot, and it ultimately aspires to evolve into a much larger digital archive for legal materials, shared by law libraries throughout the United States. With the organization-wide support of the Legal Information Preservation Alliance and the American Association of Law Libraries, this vision is indeed within reach. Beyond the borders of the United States, the Chesapeake Project aims to inform the preservation initiatives of other organized groups of libraries, who may learn through its experiences, and to raise global awareness of the vulnerability of digitally born legal materials published on the Web.

Conclusion

The end of the first decade of the twenty-first century is in sight, and the law library community has reached a crossroads in determining its role as the steward of legal information in an increasingly digital world. The amount of information being

produced in digital formats and distributed via electronic media has exploded over the past decade. However, digitally born materials – especially those that are published directly and independently to the Web – are presently at an extremely high risk of permanent loss. Our legal heritage is no exception to this phenomenon, and efforts must be put forth to ensure that our current body digital legal information is not lost. Movements to preserve digitally born legal and government publications have been set in motion – most notably those implemented by the GPO and the NDIIPP projects administered by the Library of Congress, which have enlisted the assistance of libraries throughout the United States.

Where does the law library community stand when it comes to the active preservation of our digital legal heritage? Our survey findings indicate that law libraries represented in our sample recognize that digitally born legal materials are at high risk of loss, yet their own digital preservation projects have primarily focused upon the preservation of digitized print materials, rather than digitally born materials. Digital preservation activities among surveyed libraries have been largely limited by a lack of funding, staffing and expertise; however, these barriers could be overcome by collaboration with other institutions, as well as participation in a large-scale regional or national digital preservation movement, which would allow for resource sharing among participants.

One such collaborative digital preservation program has been initiated within the past year: the Chesapeake Project. The project, which is a collaborative effort between academic and state law libraries, is being implemented under the auspices of the Legal Information Preservation Alliance. The first year of the program has been shown to be successful, and the Chesapeake Project shows great promise in its goal to inspire a nationwide effort to prevent the loss of digital legal information. Tackling the challenges of digital preservation represents a means by which law libraries can reclaim their traditional roles as stewards of information in the digital sphere. More importantly, it would ensure that our contemporary legal heritage would be preserved for generations to come.

Acknowledgements

The authors would like to acknowledge the support and guidance of Kent McKeever, Steve Anderson, Mary Alice Baish and Janice Anderson.

References

- Anon. (1980). The preservation crisis. *Journal of Academic Librarianship*, 6(5), 290. Retrieved March 8, 2008 from <http://search.ebscohost.com/login.aspx?direct=true&db=aph&AN=11298426&site=ehost-live/>
- Baldwin, G. (2003). *Fugitive documents: On the loose or on the run*. Retrieved March 29, 2008, from http://www.access.gpo.gov/su_docs/fdlp/pubs/adnotes/ad081503.html#3/
- Bradley, K. (2007). Defining digital sustainability. *Library Trends*, 56(1), 148–163.
- Breeze, H. (2005). *Results of general survey*. Retrieved April 10, 2008, from <http://www.aallnet.org/committee/lipa/inventory/InventoryReportSectV.html/>
- Burlington Magazine*. (1967). The Florentine flood disaster. *Burlington Magazine*, 109(769), 192–194. Retrieved March 6, 2008 from <http://links.jstor.org/sici?sici=0007-6287%28196704%29109%3A769%3C192%3ATFFD%3E2.0.CO%3B2-%23/>
- California Digital Library. (n.d.). *The Web at risk: Preserving our nation's cultural heritage*. Retrieved March 29, 2008, from <http://www.cdlib.org/inside/projects/preservation/webatrisk/>

- Center for Research Libraries. (2007). *Core requirements for digital archives*. Retrieved March 29, 2008, from <http://www.crl.edu/content.asp?l1=13&l2=58&l3=162&l4=92/>
- Chesapeake Project. (2007). *Collection plan*. Retrieved March 8, 2008, from http://web3.unt.edu/webatrisk/cp_cur/CP/ChesapeakeCollectionPlan_July2Draft.pdf/
- Chesapeake Project. (2008). *First-year pilot project evaluation*. Unpublished manuscript.
- Clareson, T. (2006). NEDCC survey and colloquium explore digitization and digital preservation policies and practices. *RLG DigiNews*, 10(1). Retrieved April 10, 2008, from <http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000068991&reqid=112907#article1/>
- Commission on Civil Rights. (2007). *Reinvigorating the nation's civil rights debate: The strategic plan of the United States Commission on Civil Rights for fiscal years 2008–2013*. Retrieved April 13, 2008, from <http://www.usccr.gov/pubs/SPFY0813web.pdf/>
- Deegan, M., & Tanner, S. (2006). Key issues in digital preservation. In M. Deegan & S. Tanner (Eds.), *Digital preservation* (pp. 1–31). London: Facet.
- ERPANET (n.d.). *ErpaStudies*. Retrieved March 29, 2008, from <http://www.erpanet.org/studies/index.php/>
- Gano, G., & Linden, J. (2007). Government information in legacy formats: Scaling a pilot project to enable long-term access. *D-Lib Magazine*, 13(7/8). Retrieved March 12, 2008, from <http://www.dlib.org/dlib/july07/linden/07linden.html/>
- Germain, C.M. (2002). Web mirror sites: Creating the research library of the future, and more. *Legal Reference Services Quarterly*, 21(2/3), 87–104.
- Glenn, V.D. (2007). Preserving government and political information: The Web-at-Risk project. *First Monday*, 12(7). Retrieved March 12, 2008, from http://firstmonday.org/issues/issue12_7/glenn/index.html/
- Holdsworth, D. (2006). Strategies for digital preservation. In M. Deegan & S. Tanner (Eds.), *Digital preservation* (pp. 32–59). London: Facet.
- International Data Corporation. (2008). *The diverse and exploding digital universe: An updated forecast of worldwide information growth through 2011*. Retrieved March 8, 2008, from http://www.emc.com/digital_universe/
- International Review*. (1878). The government library at Washington. *International Review*, 5, 754–769. Retrieved March 6, 2008 from <http://0-proquest.umi.com.gull.georgetown.edu:80/pqdlink?did=409473881&sid=2&Fmt=2&clientId=5600&RQT=309&VName=HNP/>
- Internet Archive. (n.d.). *Web archiving services*. Retrieved March 12, 2008, from <http://www.archive.org/web/web.php/>
- Kenney, A., & Buckley, E. (2005). Developing digital preservation programs: The Cornell Survey of Institutional Readiness, 2003–2005. *RLG DigiNews*, 9(4). Retrieved March 12, 2008, from <http://digitalarchive.oclc.org/da/ViewObject.jsp?objid=0000068919&reqid=1642#article0%C2%A0/>
- Kniffel, L. (2003). Though devastating, Iraq library losses may be less than feared. *American Libraries*, 34(6), 40–41.
- Kumar, S.L. (2006). Providing perpetual access to government information. *Reference Librarian*, 45(94), 225–232.
- LeFurgy, W. (2005). Building preservation partnerships: The Library of Congress National Digital Information Infrastructure and Preservation Program. *Library Trends*, 54(1), 163–172.
- Legal Information Preservation Alliance. (2003). *Preserving legal information for the 21st century: Toward a national agenda, March 6–8, 2003*. Retrieved March 2, 2008, from http://www.aallnet.org/committee/lipa/LIPA_Conference_Report.pdf/
- Legal Information Preservation Alliance. (2006). *Legal Information Preservation Alliance strategic plan outline*. Retrieved March 8, 2008, from <http://www.aallnet.org/committee/lipa/StratPlanFinalDraft20060620.doc/>
- Legal Information Preservation Alliance. (n.d.a.). *LIPA member libraries*. Retrieved April 2, 2008, from <http://www.aallnet.org/committee/lipa/members.asp/>
- Legal Information Preservation Alliance. (n.d.b.). *LIPA's mission statement*. Retrieved March 2, 2008, from <http://www.aallnet.org/committee/lipa/mission.asp/>
- Library of Congress. (n.d.a.). *Preservation of state digital information*. Retrieved March 29, 2008, from <http://www.digitalpreservation.gov/partners/states.html/>
- Library of Congress. (n.d.b.). *Web capture*. Retrieved March 29, 2008, from <http://www.loc.gov/webcapture/index.html/>

- Library of Congress. (n.d.c.). *What is being saved*. Retrieved March 29, 2008, from <http://www.digitalpreservation.gov/library/collect.html/>
- Lopresti, R., & Gorin, M. (2002). The availability of US government depository publications on the World Wide Web. *Journal of Government Information*, 29(1), 17–29.
- Lyons, S. (2006). Preserving electronic government information: Looking back and looking forward. *Reference Librarian*, 94, 207–223.
- Madison, J. (1910). *The writings of James Madison: Comprising his public papers and his private correspondence, including numerous letters and documents now for the first time printed: Vol. 9. 1819–1836* (G. Hunt, Ed.). New York: G.P. Putnam's Sons. Retrieved March 13 2008, from <http://0-galenet.galegroup.com.gull.georgetown.edu:80/servlet/MOML?af=RN&ae=F150083426&srchtp=a&ste=14/>
- Murray, K., & Phillips, M. (2007). *Collaborations, best practices and collection development for born-digital and digitized materials*. Paper presented at DigCCurr 2007, An International Symposium in Digital Curation. Retrieved March 29, 2008, from http://ils.unc.edu/digccurr2007/papers/murrayPhillips_paper_9-3.pdf/
- NASCIO. (2007). *Electronic records management and digital preservation: Protecting the knowledge assets of the state government enterprise*. Retrieved March 8, 2008 from <http://www.nascio.org/publications/documents/NASCIO-RecordsManagement.pdf/>
- Neacsu, D. (2002). Legal scholarship and digital publishing: Has anything changed in the way we do legal research? *Legal Reference Services Quarterly*, 21(2/3), 105–122.
- Neacsu, D. (2007). *Google, legal citations and electronic fickleness: Legal scholarship in the digital environment*. Retrieved March 8, 2008, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=991190/
- Pardo, T.A., Burke, G.B., & Kwon, H. (2006). *Preserving state government digital information: A baseline report*. Retrieved March 9, 2008, from http://www.ctg.albany.edu/publications/reports/digital_preservation_baseline/
- Rawan, A., & Malone, C.K. (2006). A virtual depository: The Arizona project. *Reference Librarian*, 45(94), 5–18.
- Ross, S. (2004). The role of ERPANET in supporting digital curation and preservation in Eurpoe. *D-Lib Magazine*, 10(7/8). Retrieved March 29, 2008, from <http://www.dlib.org/dlib/july04/ross/07ross.html/>
- Rothenberg, J. (1999). *Ensuring the longevity of digital information*. Retrieved March 9, 2008, from <http://www.clir.org/pubs/archives/ensuring.pdf/>
- Rumsey, M. (2002). Runaway train: Problems of permanence, accessibility and stability in the use of Web sources in law review citations. *Law Library Journal*, 94(1), 27–39.
- Simpson, D. (2005). *Digital preservation in the regions: Sample survey of digital preservation preparedness and needs of organisations at local and regional levels*. Retrieved March 8, 2008, from http://www.mla.gov.uk/programmes/digital_initiatives/digital_preservation/digital_preservation_in_the_regions/
- Solum, L.B. (2006). Blogging and the transformation of legal scholarship. *Washington Law Review*, 84, 1071–1088. Retrieved April 13, 2008, from <http://papers.ssrn.com/abstract=898168/>
- Special Committee on Permanent Public Access to Legal Information. (2005). Member's briefing: Preservation. *AALL Spectrum*, 10(3), 1–4.
- University of North Texas Libraries. (n.d.a.). *Congressional Research Service Reports, hosted by UNT Libraries*. Retrieved March 29, 2008, from <http://digital.library.unt.edu/govdocs/crs/>
- University of North Texas Libraries. (n.d.b.). *CyberCemetery*. Retrieved March 29, 2008, from <http://govinfo.library.unt.edu/>
- US Government Printing Office. (2003). *The GPO and National Archives unite in support of permanent online public access*. Retrieved March 29, 2008, from <http://www.gpoaccess.gov/pr/media/2003/03news46.pdf/>
- US Government Printing Office. (n.d.). *FDsys Overview*. Retrieved March 29, 2008, from http://www.gpo.gov/projects/fdsys_overview.htm/
- Web-at-Risk. (n.d.). *Collection plans*. Retrieved April 2, 2008, from <http://web3.unt.edu/webatrisk/cpg.php/>
- Yin, R.K. (2003). *Case study research: Design and methods* (3rd edn). Thousand Oaks, CA: Sage.

APPENDIX

Law Libraries & Digital Preservation: A Survey

1. Welcome

Dear Librarian:

We are two academic law librarians interested in law library activities relating to digital preservation. (Dana Neacsu, at dana.neacsu@law.columbia.edu, and Sarah Rhodes at sjr36@law.georgetown.edu.)

We have developed a survey to gauge the law library community's digital preservation efforts, and we are contacting you today to request your participation in our research.

The survey is divided into the following four short sections:

- (a) Demographic information,
- (b) Digital-preservation activities,
- (c) Perception and attitudes on digital preservation, and
- (d) Copyright and access policies for archived digital materials.

We know how busy you are, so the survey has mostly multiple-choice and scale-based questions, with very few questions requiring more than selecting a checkbox to indicate your answer.

Certainly, we will be happy to share the results with all respondents. Because we are working on a deadline, we would appreciate receiving your responses within two weeks from today. Please do not hesitate to contact either one of us with questions, suggestions, etc.

Sincerely,
Sarah and Dana

Next >>

2. Demographic information

1. I represent a(n)

- Academic Law Library
- State Law Library
- Other

2. If you answered "Other," please describe your library:

3. My position is that of

- Director
- Other

4. If you answered "Other," please describe your position:

<< Prev Next >>

3. Digital preservation activities

1. How many large-scale and/or small-scale digital preservation projects have you planned and/or executed within your library in the last five years?

Please provide approx. number of projects:



2. Are these projects focused on the preservation of digitized (scanned) print items?

- Yes
- No
- Some are, but not all

3. If you answered "yes" or "some," how many of your library's projects have focused on the preservation of digitized (scanned) print items?

Please provide approx. number of projects:



4. Are these projects focused on the preservation of digitally born materials (originally created and disseminated in a digital format, often with no print counterpart)?

- Yes
- No
- Some are, but not all

5. If you answered "yes" or "some," approximately how many of your library's projects have focused on the preservation of items digitally born?

Please provide approx. number of projects:




6. How many of these projects have you successfully completed?

Please provide approx. number of projects:



7. How many of these projects are currently in progress?

Please provide approx. number of projects:



8. What digital repository and content management systems have been used for these projects?

Select all that apply.

- CONTENTdm
- DigiTool
- DSpace
- EPrints
- Fedora
- Greenstone
- Hyperion Digital Media Archive
- Lots of Copies Keep Stuff Safe (LOCKSS)
- MetaSource
- OCLC Digital Archive
- VITAL
- Other

9. If you answered "other," please name/describe your digital repository and content management systems:



10. Does your institution use ...

(Select all that apply.)

- An in-house digital repository
- Repository hosted offsite by vendor
- Other

11. If you answered "other," please describe your repository:



12. If you use a commercial digital repository, who operates your commercial digital repository?

Please provide name of vendor(s):



13. To what extent do you collaborate with other institutions and/or nonprofit/for-profit partners in developing your digital projects?

Please mark the appropriate box:

	Never collaborate	Rarely collaborate	Sometimes collaborate	Almost always collaborate	Always collaborate
Collaborate with libraries/institutions	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Collaborate with nonprofit partners	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Collaborate with for-profit partners	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

14. With whom have you collaborated?

Please list institutions, organizations, and commercial collaborators:

15. What was/is the collection scope of your institution's digital preservation projects? (Preservation of which materials takes precedence in your projects and why?)

Please explain:

16. Do you preserve primary legal materials in digital formats?

- No
- Yes
- Sometimes

17. If so, are those... (Please select all that apply.)

- Domestic
- International
- Foreign

18. If preserving domestic primary legal materials, are those... (Please select all that apply.)

- Federal
- State
- County
- Municipal

19. Do you preserve secondary legal sources in digital formats?

- Yes
- No
- Sometimes

20. If so, are those... (Please select all that apply.)

- Domestic
- International
- Foreign

21. If preserving domestic secondary legal materials, are those... (Please select all that apply.)

- Federal
- County
- Municipal
- State

22. Do you preserve digital items harvested from the Web?

- Yes
- No
- Sometimes

23. If so, what Web harvesting software do you utilize?

Please describe:



Section 4: Perceptions and attitudes about digital preservation

1. Please use the scale below to indicate your level of agreement with the statements below, as applicable.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Law libraries should be involved in preventing the loss of law-related information published to the Web.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Law libraries should be involved in the preservation of digitally born government information.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Law libraries should be involved in preventing the loss of information published on the Web and cited within law review articles.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Law libraries should be involved in the long-term preservation of and sustained access to law review articles and other legal materials published digitally within subscription databases (HeinOnline, LexisNexis, Westlaw, etc.).	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. My library's level of involvement in digital preservation activities has been limited by:

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Lack of funding	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Concerns about technology/file format obsolescence	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Staffing shortages	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lack of staff with digital preservation/technological expertise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Digital preservation is not an institutional priority	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lack of partners/opportunities to collaborate with other libraries and organizations in digital preservation activities	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Lack of an organized statewide/nationwide/international digital preservation movement in which to participate	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

3. Are there factors, not listed above, which have limited your library's level of involvement in digital preservation activities?

If so, please describe:

4. The following would encourage greater involvement in digital preservation activities at my library:

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
Increased funding	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Increased staffing	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Recruitment/cultivation of staff with digital preservation/technological expertise	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Increased opportunity to collaborate with other libraries and organizations in digital preservation activities	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Establishment of an organized statewide/nationwide/international digital preservation movement in which to participate	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

5. Are there factors, not listed above, which would encourage greater involvement in digital preservation activities at your library?

If so, please describe:

6. Which materials, in your opinion, deserve more attention when it comes to preservation:

- Print materials
- Digitally born materials

7. Why do you think either a) print or b) digitally born materials are in greater need of preservation? Please explain your choice:

8. Please describe, briefly, the role and responsibility that you envisage law libraries should take in the preservation of Web-published and digitally born legal information:

<< Prev Next >>

Section 5: Copyright and access policies for archived digital materials

1. Is permission from the copyright holder obtained for copyright-protected items that are being digitally archived by your library?

- Yes
- No
- Sometimes

2. Do you preserve copyright-protected materials under a claim of fair use?

- Yes
- No
- Sometimes

3. Do you preserve only materials that are in the public domain?

- Yes
- No
- Sometimes

4. How is copyright being managed for Web-harvested items?

Please explain:

5. How is copyright being managed for digitized (scanned) print items?
Please explain:



6. How is copyright being managed for digitally born items (not Web-harvested)?
Please explain:



7. To what extent are digitally archived items made available for patron use?
- Fully accessible to the public online
 - Accessible only to authenticated patrons online
 - In-library access for library patrons
 - Not accessible to patrons
 - Other

8. If you answered "Other" above, please explain how your digitally archived items are made available for patron use:



9. What limitations on access, if any, are in place?

Please describe:



10. Are digitally archived items accessible via a Web database or portal, or in-house only?

Please describe:



We thank you for your time and consideration in completing this survey.