Prepared for:

Delaware Division of Libraries and the
Delaware Digital Preservation Steering Committee

Current Digital Preservation Capability of the State of Delaware

Digital Preservation Planning Project

Deliverable: 01

July 20, 2007

Contract Number: DD:FY_07-200801CD
Delaware State Project Manager: John Sthele

Prepared by:
Charles Dollar
COHASSET ASSOCIATES, INC.

CONTENTS

INTRODUCTION

Digital information has become an integral part of the governance of society and our cultural heritage. Increasingly, ordinary citizens are confronted with documentation of their legal rights, historical events, and other useful information that is only available in electronic form. The rapid pace of digital technology obsolescence, both hardware and software, poses a major challenge to the long-term availability and use of this digital documentation. Equally as important is the potential loss of irreplaceable digital documentation of long term value as a consequence of a "Katrina-like" natural disaster, especially in light of the fact that the Delaware Department of Technology and Information's backup and business continuity protection of digital information is not equivalent to the accepted standards and best practices of archival preservation.

In December of 2006 the National Association of Secretaries of State (NASS) and National Electronic Commerce Coordinating Council (eC3) published a symposium report, "Digital Archiving – From Fragmentation to Collaboration," that identified three major concerns a sustainable digital preservation program at the state level should address:

1. The integration of digital preservation into a broader context of state information, research, and cultural policy.

2. The collaboration of state organizations whose programs create, collect or disseminate digital information on behalf of stakeholders.

3. Sound technical solutions whose estimated costs and benefits are accurately calculated so they can be presented to state resource allocators.

Early in January 2007 the NASS/eC3 call for integration and collaboration along with concern about the potential loss of digital documentation created, received, or acquired by Delaware state agencies and local governments spawned the ad hoc Delaware Digital Preservation Steering

Group (DDPSG). The goal of the DDPSG is to develop a comprehensive digital preservation planning document for the State of Delaware that establishes the framework within which to develop a two phase comprehensive collaborative digital preservation program. Phase 1 involves creation of a high level digital preservation planning document while Phase 2 is a digital preservation implementation program.

With funding provided by the Delaware Division of Libraries, Cohasset Associates, a records management consulting firm, was engaged to carry out three tasks associated with Phase 1: (1) Collect data about the current state of digital preservation awareness and capability in state agencies; (2) Develop a digital preservation capability maturity model and digital preservation balanced scorecard; and (3) Recommend a digital preservation strategy that can be used to drive the completion of Phase 2.

PURPOSE AND SCOPE

This working paper constitutes the deliverable for Task 1 to establish a base line of the current state of digital preservation in state and local government agencies. Toward this end, Cohasset Associates designed a set of questions to elicit specific information about current digital preservation awareness, programs and activities. The data would be analyzed to establish the current state of digital preservation in state and local government agencies. Selected portions from this assessment of the current state of digital preservation will be used in Deliverable 2, which involves a digital preservation capability maturity model that includes high level metrics for measuring success as the State of Delaware implements a digital preservation program.

This deliverable is a report about the current state of digital preservation in the State of Delaware based upon an analysis of data collected from a web enabled survey that Delaware Information Resource Managers, Records Officers, Librarians, and Archivists were invited to complete. Cohasset Associates has organized this deliverable into three sections that encompass: (1) The Digital Preservation Readiness/Capability Survey; (2) An Analysis of Survey Responses; and (3) Recommendations. Also included in Appendix A are frequency distributions of responses to selected questions in the survey

1 THE DELAWARE DIGITAL PRESERVATION READINESS/CAPABILITY SURVEY

1.1 Background

Cohasset Associates has developed a structured interview methodology that has been very successful in eliciting information from organizations about the life cycle management of digital information. Consequently, Cohasset's work plan for Task 1 included a data collection component that involved the gathering of digital preservation information through interviews with ten (10) Delaware State agencies, departments, and divisions[1] about a number of topics, including:

- Previous digital preservation initiatives
- Current and projected future volume of born digital information
- Current human resources devoted to digital preservation
- Projected future human resources that will be devoted to digital preservation
- Current financial resources devoted to digital preservation
- Awareness and understanding of digital preservation issues
- Projected future financial resources that will be devoted to digital preservation
- Potential digital preservation collaborative initiatives that leverage current digital information technology initiatives.

Cohasset prepared draft questions that addressed these topics for review by the Steering Committee. During a review of the draft questions the Steering Committee learned of a web enabled survey tool that the Department of Technology and Information (DTI) had used successfully in collecting information about the current state of disaster recovery and business continuity in state agencies, departments, and divisions. The Steering Committee decided that

[1] The number of interviews to be conducted was determined by time constraints and available resources.

the data collection should be done through a web enabled survey rather than through face to face interviews.

The survey process usually begins with a kick off meeting attended by survey participants. In most instances, part of the kick off meeting is spent walking participants through the questions and explaining the scope and focus of each question. The remainder of the kick off meeting is devoted to people actually completing the survey form (where possible). The proponents of this methodology believe it offers several advantages:

- It is much more efficient than face to face interviews in which open ended discussions form the basis for data collection and important details may not emerge in the discussion
- More questions can be asked and the questions are presented consistently to all survey participants
- All participants hear the same explanations.
- Participants can complete the survey at their convenience
- There is a basis for personal follow-up if necessary.

Use of a formal, structured survey to collect data requires narrowly focused questions with predetermined selected responses (e.g., "Yes," "No," "Do not know"). The responses may also include branching options for further drill down, depending upon the response (e.g., "If Yes, please answer this question"). Cohasset adapted the original interview questions to a survey structure that resulted in 78 questions.[2] A Strohl Systems consultant under contract with the Department of Technology and Information formatted the 78 questions into a web enabled survey that participants could complete electronically.

[2] Some of these questions (e.g., "number of employees in the business unit" or "number of years of employment") were intended to be linked to other questions in order to identify more complex issues. For example, does the length of employment or business unit appear to make any difference in how business units or respondents rate their understanding of digital preservation issues?

1.2 Survey Questions

The goal of the survey is to collect information about the digital preservation readiness and capabilities of Delaware agencies, departments, and divisions. Some questions focus on agency or department-wide activities, while other questions emphasize individual activities and aptitude. The 78 questions are divided into eight categories:

1. Organizational and Respondent Information
2. Digital Preservation Knowledge and Skills
3. Digital Preservation Policy
4. Creation, Receipt, or Acquisition of Digital Objects
5. Archival Storage of Digital Objects
6. Access and Dissemination of Digital Objects
7. Human, Technical, and Financial Resources
8. Impediments to Digital Preservation

1.3 Survey Administration

The revised work plan envisioned two working sessions with survey participants in which the survey would be introduced, questions answered, and the survey forms filled out during the working sessions. However, this approach had to be dropped because of a break down in communications in obtaining official clearance to notify the survey participants about the two working sessions. The Steering Committee decided to broaden the scope of the survey to include Information Resource Mangers and Records Officers and to rely exclusively on participants to complete the survey electronically at their work place and at their convenience.

The survey was sent electronically to the Steering Committee,[3] 80 Information Resource Managers,[4] and approximately 250 Records Officers who were requested to complete the survey by June 8, 2007. This deadline was later extended to June 15, 2007. Fifty-three out of 336

[3] The Steering Committee included three representatives of the Department of Technology and Information but only one individual was designated to be the "official" DTI survey participant. Similarly, the Public Archives is represented by two individuals but only one individual was the official survey participant.

[4] This included individuals in school districts whose work is not related to the preservation of digital objects.

individuals invited to participate in the survey completed a survey form, which is nearly a 16 percent response rate.

2 DIGITAL PRESERVATION SURVEY ANALYSIS

2.1 Methodology Review

Cohasset established an analytical framework for the survey analysis in conjunction with the Strohl consultant as part of the survey design. This analytical framework involved the following assumptions:

- Responses accurately reflect the respondents' assessment because the respondents actively selected each response (or Non-response).
- "Non-responses" are equivalent to "Do not know" or "Not sure" responses.
- A frequency distribution (numerical and percentage) summarizes the responses made to each question.
- Cross tabulations of one or more selected questions with other selected questions can serve as a consistency check and to be used to identify underlying patterns or central tendencies.
- Tests of statistical significance are not used with the cross tabulations because a statistically normal distribution of respondents can not be assumed. The Strohl Systems software application used to support the web-enabled survey required "locking down" the identification of questions to be cross-tabulated at the time of survey design. Any subsequent changes in the survey target population or new questions could not be analyzed with the pre-programmed cross tabulations. This limitation came into play when the Steering Committee made the decision to send the survey to Records Officers and Information Resource Managers with the result that it is not feasible to compare

systematically the responses of Records Officers vis-à-vis the responses of Information Resource Managers.

It also should be noted that certain questions were excluded from this analysis because of the linkage issues discussed above or because the question was determined not to be useful (e.g., number of employees as provided by respondents) in eliciting relevant information about the current state of digital preservation. Finally, where several questions asked about closely related topics (e.g., understanding of selected International Standards) and where the responses were virtually identical, only the responses to a single question were used.

As noted earlier, 53 individuals responded to the survey. However, this included two respondents respectively from the Delaware Public Archives and the Historical and Cultural Division that essentially duplicated responses. Consequently, only one survey response from each of these two business units is included in this report, which resulted in a pool of 51 respondents.[5]

2.2 Organizational and Individual Profiles

2.2.1 Background

The goal of the questions in this section was to elicit an organizational and employment context for the overall findings of the digital preservation readiness/capability of various departments, divisions, and programs.

2.2.2 Findings

- The pool of 51 respondents represented 41 departments, divisions, and business units at state and local levels that are spread across the spectrum of state and local government entities. This distribution of representation suggests that survey findings are likely to represent central tendencies for the State of Delaware.

[5] This discrepancy was not discovered until after completion of frequency distributions for responses to questions. In addition, the base for percentage calculations varied with the number of responses to a specific question, which in some instances was 51 and in others it was as low as 25. These two circumstances required manual tabulation of survey responses to produce uniform frequency distributions.

- Forty-two (42) of the respondents have been employed in their current position for four years or more, *which means that they are likely to know a great deal about business/work environment and are in a position to speak reliably about the state of digital preservation.*
- More than one-half of the respondents are in managerial positions. *This indicates that digital preservation issues are being called to the attention of program managers.* Follow-up contact with these managers may lead to opportunities to present digital preservation issues to broader range of senior managers.
- Eight of the respondents identified themselves generally as information technology professionals. *Follow-up contact with these information technology professionals about digital preservation issues may lead to their becoming knowledgeable advocates of a digital preservation program.*
- The preponderance (44) of respondents spend less than 1 hour per week on digital preservation, which suggests that *digital preservation is not a priority*. This could be a function of the level of their digital preservation knowledge and skill, the amount of time available for digital preservation, the importance that their agency, department, or division attaches to digital preservation, or some combination thereof. This is a topic that merits further research.[6] It is noteworthy that in the absence of a digital preservation program for the State, six respondents spend at least five hours per week on digital preservation. *With encouragement, these six respondents could form the core of a digital preservation cadre that could become the foundation of a sustainable digital preservation program.*
- According to survey participants 12 digital preservation projects have been initiated over the past five years.[7] One project was unsuccessful and four other projects were marginally successful. Two projects were rated moderately successful and five were very

[6] The research methodology anticipated that cross tabulations of responses to these questions could identify underlying central tendencies. However, the "linkage limitation" mentioned earlier precluded this level of analysis.

[7] Email exchanges with several survey respondents suggest that some of the digital preservation projects were in fact projects in which paper records were scanned to digital images.

successful. *These 12 projects should be carefully reviewed to distill lessons learned that can be useful in designing a sustainable digital preservation program.*

2.3 Digital Preservation Knowledge and Skill

2.3.1 Background

A key requirement for a successful Digital Preservation Program is a staff that understands digital preservation issues and has the requisite skills to carry out digital preservation activities. The questions in this section attempt to establish a baseline of individual digital preservation knowledge and skill levels. A threshold question in this regard asked respondents to rate their understanding of digital preservation issues from "Very High" to "None." Following this question was a series of questions about their familiarity with digital preservation standards, awareness of digital preservation best practices, and digital preservation training.

2.3.2 Findings

- Almost 70% of respondents have no or little understanding of digital preservation issues, knowledge or skills. *An immediate top priority for the State of Delaware should be the implementation of an on-going digital preservation training and awareness program that enables employees to grow in their knowledge and understanding of digital preservation issues.*
- The absence of familiarity with key International Standards on digital preservation, especially ISO 17421, and the Research Library Group Audit Check Guidelines for Trusted Digital Repositories collectively denote the low state of employee understanding of critical digital preservation issues and *reinforce the need for a substantial investment in digital preservation training.*
- Another strong indicator of the poor state of digital preservation knowledge and skills is that 38 respondents (75%) do not keep informed about digital preservation issues. The 13 respondents who attempt to keep informed about digital preservation issues do so largely through conferences (8) and web sites (13). Although this is a relatively small number, the very fact that as many as 13 respondents are attempting to learn more about

digital preservation *suggests that they could become the nucleus of a digital preservation training cadre that can be expanded over time*.

- Only 4 out of the 51 survey participants attempted to identify digital preservation best practices and guidelines that could be implemented in a digital preservation program. Interestingly, one respondent cited the State of Delaware Model Guidelines for Electronic Records which has little direct bearing on digital preservation. *This particular finding reinforces the need for mobilizing digital preservation awareness and training programs that over time can correct this great deficiency.* At the very least, such a program should be incorporated into the Stage 2 digital preservation program.

2.4 Digital Preservation Policy

2.4.1 Background

The questions in this section of the questionnaire focused on identifying the scope of a digital preservation readiness/capability policy (agency, department, division, or program) and to elicit an assessment of the effectiveness of the policy. Specifically, the questions asked about the authority of the agency, department, or division to issue a digital preservation policy, whether the policy had been issued as a formal document, and if in fact the digital preservation policy had been implemented. Respondents who indicated that a formal written policy had been implemented were asked to select the features it incorporated from a list of possible features. Supplemental information was obtained in interviews and queries with the Delaware Public Archives, the Department of Technology and Information, and the State Library that further clarified the findings about a digital preservation policy.

2.4.2 Findings

- Eight (16%) of the 51 respondents[8] reported that their agency, department, or division had the statutory or regulatory authority to issue a digital preservation policy. One respondent reported that a formal digital preservation policy had been issued. [9] Five respondents indicated that the actual implementation of the policy was either still in the planning stage or was partially implemented. The accuracy of these responses is questionable.

The Delaware Public Records Act 29 De. Code, §503 assigns to the Delaware Public Archives the authority to establish and enforce "the policies and guidelines for the management and preservation of All public records of the State …." However, Title 29, Delaware Code, § 9004C, states that the general powers and duties of the Department of Technology and Information (DTI) include statewide and interagency technology solutions, policies, and standards.

A good example of how DTI has used this mandate can be found in a Data Classification Policy (In-DataClass-001. There are several interesting aspects of this policy that have a bearing on digital preservation. They include:

- A provision to evaluate protection requirements of computerized data with respect both to data integrity and confidentiality.
- Documenting the retention periods for computerized data in the General Records Retention Schedule and Agency Records Retention Schedules.
- Distinguishing between computerized data and computerized records without an explicit acknowledgement that under the Delaware Public Records law computerized data are defined as public records.

[9] A second respondent reported that a formal digital preservation policy had been issued but a follow up query disclosed that the Social Security Administration "owns" the records and had issued the policy.

There is an additional layer of confusion that is associated with the DPA's issuance several years ago of "Model Guidelines for Electronic Records" and "Guidelines for Maintaining and Preserving Records of Web-Based Activities." The former articulates recommended best electronic records management practices but it does not directly address what digital preservation entails.[10] The latter guidelines essentially inform agencies how to capture web site snapshots and transfer them to the archives. It appears that some respondents may have mistakenly viewed these DPA guidelines as well as DPA guidelines for converting paper records to digital images to conversion as a digital preservation policy. *In fact, the DPA does not have a digital preservation policy.* In this context it should be noted that the DPA "Model Guidelines for Electronic Records" do not reference DTI other than it being a signatory to an example Memorandum of Understanding.

Given this overlapping mandate of statutory authority and confusion about what constitutes a digital preservation policy, *the development and adoption of a collaborative state wide digital preservation policy with DTI should be a top priority for the Delaware Digital Preservation Project.* Without such a collaborative initiative it is unlikely that a sustainable digital preservation program can be grown. Deliverable 3 will contain a recommended digital preservation policy that could form the basis for such collaboration.

2.5 Capture, Receipt, or Acquisition of Digital Objects

2.5.1 Background

The way(s) in which digital objects are created, received, or acquired can have an enormous downstream impact on digital preservation readiness/capability. Digital objects created, received, or acquired in a proprietary file format, an obsolete file format, or a file format that has limited use are likely to pose significant technological and financial burdens over time. These questions, therefore, attempt to identify the scope of these technological and financial burdens through estimates of legacy, potential preservation ready, and preservation ready digital objects.

[10] A Memorandum of Understanding (MOU) between the State Personnel Office, the Department of Technology and Information, and the Division of Accounting references the requirement for documentation that includes "refreshment, migration, and conversion plans."

Legacy digital objects are encoded in obsolete or proprietary software or formats with no export functionality that supports exporting them to newer software and formats. Legacy digital objects are encoded in a proprietary file format that in most instances cannot be read by a text editor and the specifications of the format are not publicly available. Potential preservation ready digital objects typically are encoded in a native proprietary format but tools exist that can transform the digital objects into technology neutral open standard formats (e.g., transformation of Word Perfect documents to Word documents). Preservation ready digital objects are encoded in a technology neutral open standard format with sufficient metadata that they can be moved to a digital repository without any additional processing.

The identification of legacy, potential preservation ready, and preservation ready digital objects that a department, division, or program "owns" through creating, receiving, or acquiring them in the ordinary course of business is the focus of the questions in this section. These questions drilled down to topics such as estimates of the volume of digital records owned (including the volume of digital objects retained for ten years or longer), estimated growth, the file formats used to create, receive, or acquire digital objects with special attention to digital objects in a legacy format, a potential preservation ready format, and a preservation ready format.

2.5.2 Findings

- 72% (38) of respondents identified one or more applications used to create, receive, and acquire digital objects. The largest number of applications was Microsoft (17) followed by Adobe Acrobat (6). None of the applications (see Appendix A, Create, Receive, or Acquire Digital Objects) is now considered obsolete. Of course, this is no guarantee that in the future such applications will not become obsolescent (e.g., File Maker Pro). *A technology obsolescent monitoring watch program that keeps track of applications that are on the verge of becoming obsolescent should be implemented so that migration to current applications can be undertaken while the export functionality window remains open.*
- Only 16% (8) of respondents could provide rough estimates of the volume of digital objects that will be retained for ten years or longer. The remaining 43 (82%) respondents either did not know or did not respond. The combined estimated volume of digital

objects these 8 respondents reported is about 4 Terabytes, which undoubtedly greatly underestimates the volume of these digital objects. *This underscores the urgent need to establish a base line inventory of the volume of digital objects that a digital preservation program should support.*

- 36 (71%) of the 51 respondents did not know or did not respond to the question about the volume of digital by images, text, vector graphics, spreadsheets, or databases. In addition, 68% (35) of respondents either were not sure or did not respond to the question about the estimated growth in the volume of digital objects. *This limited knowledge and understanding about the digital objects that their agency, department, or division creates, receives, or acquires is of grave concern.*
- This level of knowledge appears to be much better with regard to the estimated volume of digital objects in a legacy format. Almost 40% of the respondents reported that the volume of legacy digital objects is less than 1%. Of course, as digital technologies evolve many current formats will become legacy formats and the digital objects in which they are encoded will become unretrievable and unusable unless they are migrated to new formats.
- This "good news" about legacy digital objects should be seen in the context that only 10 (20%) respondents reported that at least 75% of the digital objects created, received, or acquired are in a potential preservation ready format, and only 3 (6%) respondents reported that at least 75% of the digital objects are in a preservation ready format.
- A related issue is the extent to which digital objects are converted to a standard format upon their creation, receipt, or acquisition. Only 4 respondents noted that digital objects are converted to a standard format, which included PDF, TIFF, and JPEG. The identification of TIFF and JPEG may be related to the conversion of paper documents to digital images and therefore may not be relevant for "born digital objects."

On balance, the responses to these questions about digital objects created, receive, or acquired indicate the need for both a digital preservation awareness training program and an accurate inventory of the digital objects that agencies, departments, or divisions create, receive, or acquire during the normal course of business.

2.6 Archival Storage of Digital Objects

2.6.1 Background

Archival storage presumes a repository that stores digital objects scheduled for long term retention. Archival storage is not the same thing as "digital archiving," which typically involves the IT function maintaining one or more copies of digital objects of digital objects (usually on magnetic tape). These copies are maintained for business continuity and disaster recovery in the event that the operational digital objects are lost or corrupted through a natural disaster, an accidental system malfunction, or intentional unauthorized destruction. Digital preservation specialists and many IT specialists alike recognize that "backing up" digital objects for business continuity and disaster recovery does not meet the requirements for long-term retention of trustworthy, usable, and accessible digital objects.

Many of the questions in this section repeated questions from the previous section but with an entirely different emphasis; the focus is on archival storage, not the creation, receipt, or acquisition of digital objects. In other words, archival storage questions deals with the issue of how to ensure the long-term retention of digital objects, not technical circumstances that prevailed at the time of creation, receipt, or acquisition. Clearly, there is a connection between the two sections but there is not duplication of information.

2.6.2 Findings

- Only 25% (13) respondents reported that digital objects were in archival/long term storage. More than one-third (16) reported that no digital objects were in archival/long-term storage. It is unclear if the remaining (22) survey participants did not respond because they did not know or for some other reason. The same observation could be made about the fact that 38 (74%) of the respondents either did not know or did not respond to the question about the volume of stored digital records.
- 13 respondents listed 11 file formats used in archival storage: ASCII, SMP, GIF, JPEG, Lotus 123, PDF, PDF/A, RTF, SQL, TIFF, and XML *Of these 11 formats, only ASCII, JPEG, PDF/A, and XML are considered technology neutral open standard formats.*
- Not surprisingly, the number of non-responses to other questions related to digital objects in a legacy format, a potential preservation ready or preservation ready format ranged

from 24 to 28 which reinforces the likelihood that these respondents simply did not know. *This highlights a central tendency emerging from this survey that at least one-half of the survey respondents know virtually nothing about the digital objects that agencies, departments, and divisions create in the ordinary course of business.* This particular set of questions does not probe the cause for this but it is plausible to speculate that this limited knowledge base about archival/long term storage reflects the absence of digital preservation training. It may also reflect the impact of other pressures and work priorities that do not allow respondents to devote any time to digital preservation. *The good news about archival storage is that almost one-half of those who did respond to the legacy format question reported that legacy formats are not an issue.* So far as potential preservation ready digital objects are concerned, the central tendency is less comfortable: almost one-third (17) of the respondents reported that no digital objects were in a potential preservation ready format. The other side of the potential preservation ready format digital objects is that only two survey participants reported that "most" digital objects (between 96% and 100%) were in a preservation ready format. Unfortunately, these same two respondents also reported that between 96% and 100% of digital were stored in a preservation ready format. This appears to be an inconsistency on their part.

- *A key issue for digital preservation is where and how the archival storage of digital objects is accomplished.* More than one-third (21) of the survey respondents reported they did not know where and how the archival storage of digital objects "owned" by their agency, department, or division is achieved. One-fifth (11) of the survey participants simply did not respond. The remaining responses to this question disclose a great deal about the state of digital preservation in the agencies, departments, and divisions where the survey participants work.

  Twelve survey participants identified a work station hard disk drive or removable storage media (e.g., DVD) as the archival storage location of digital objects. This is not an accepted digital preservation best practice because it does not protect digital objects from natural disasters, theft, and unauthorized destruction. Almost one-third (17) of the respondents indicated that archival storage is accomplished by the use of one or more network drives. Although network drives are backed up daily they do not satisfy the

requirements for archival storage. *A digital preservation policy should specify what constitutes archival storage of digital objects that have long term value.*

- Outsourcing archival storage to a third party can be an attractive archival storage option when a state run digital preservation repository does not exist. Four (4) survey participants noted that their agency, department, or division currently outsource the storage of digital objects of long term value while 21 respondents reported that this is not done. The remaining 26 respondents either did not know or did not respond to this question. In some circumstances, outsourcing archival storage to a third party may be advantageous but this should not relieve the agency, department, or business unit of the responsibility to ensure that outsourced archival storage actually complies with digital preservation requirements. *A digital preservation policy should address this issue by requiring agencies, departments, and divisions to incorporate digital preservation requirements into contracts with third parties.*

2.7 Access and Dissemination of Digital Objects

2.7.1 Background

The ultimate rationalization for the preservation of digital objects of long-term value is to ensure that they are accessible as far into the future as necessary. Consequently, this section of the survey focuses on the access and dissemination of digital objects. Several questions from previous sections are repeated here but strictly within the context of access and dissemination. The purpose of these questions was to identify the current state of public access to and dissemination of digital objects. Nonetheless, it is important to keep in mind that future access to digital objects is unpredictable because we do not have a reliable "crystal ball" of future technologies or the expectations of future users.

2.7.2 Findings

- Only 14% (7) survey respondents reported that their agency, department, or division supported public access to and dissemination of digital objects. They included the Vocational Rehabilitation Division, the Department of Justice, Public Archives Division, State Library, Public Health, Historical and Cultural Affairs Division, and Office of

Auditor of Records. It is worth noting that there were 11 (22%) survey participants who noted that their agency, department, or division did not support public access to and dissemination of digital objects. This leaves 33 respondents who did not know or did not respond.

- Respondents from four (4) agencies (Archives, Historical and Cultural Affairs, Justice, and Library) identified the software application used to support public access to and dissemination of digital objects. The Public Archives reported the use of Microsoft Office. Historical and Cultural Affairs reported that it used Past Perfect Museum, a historical collections management software tool. Hyperion, an interactive library search and retrieval system, is the State Library tool of choice for supporting public access to and dissemination of digital objects.[11] The Department of Justice reported that it uses Alchemy, a document management system, to support public access to and dissemination of digital objects. Of these four software applications, Hyperion appears to be the most flexible and user friendly tool to support public access to and dissemination of digital objects.
- Only the State Library and the Public Archives reported the percentage of digital objects by type of record that can be publicly accessed. The preponderance of digital objects publicly available through the State Library is images, closely followed by text and databases. The digital objects the Public Archives makes available is equally distributed across text, images, databases, and spreadsheets. The Public Archives and State Library share in common providing public access to and dissemination of text and digital images. *This suggests that one component of a digital preservation program could be a shared public access service platform for text and images.* The Stage 2 Digital Implementation Plan should take this possibility into account.

[11] Hyperion is still in an implementation stage so it is not fully functional.

2.8 Human, Financial and Technical Resources

2.8.1 Background

A Delaware digital preservation program that flourishes and meets expectations must have sufficient on-going human, financial, and technical resources. Of course, what constitutes "sufficient human, financial, and technical resources" for a sustainable digital preservation program is open to interpretation and depends in part on the volume, type, and state of preservation readiness (legacy, potential, and ready) of the digital objects. Nonetheless, the questions dealing with human, financial, and technical resources were framed with the expectation of extrapolating indicators of the resources currently available to the agencies, departments, and divisions represented in the survey participants. For example, an annual budget could be linked to an estimated increase in program funding necessary to support a sustainable digital preservation readiness program. Estimates of future resources required to support a sustainable digital preservation program clearly are educated guesses but they are all that we have until a far more detailed analysis is undertaken.

2.8.2 Findings

- Almost two-thirds of respondents failed to provide meaningful responses (e.g., identify how many FTE are required to support a digital preservation program) to questions dealing with human, financial, or technical resources that support on-going program activities. It is unclear what this high non-response rate means.
- A cross tabulation of the responses to the question about the staff required to successfully implement a digital preservation program with the self-assessment responses about knowledge of digital preservation issues *suggests that there is a linkage between the level of knowledge about digital preservation and identifying the need for staff to run a digital preservation program.*[12] Raising the level of knowledge about

[12] Grouping the self-assessment ratings into two categories for the 45 survey participants who responded to both questions yields the following contingency table.

digital preservation could help create digital preservation synergy that spills over into a number of related activities.

- *This particular finding reinforces the need for a digital preservation awareness training program for people who are likely to be involved in programs that involve digital recordkeeping requirements.*

2.9 Impediments to a Digital Preservation Program

2.9.1 Background

Up to this point, the analysis of the digital preservation readiness/capability survey has focused on responses to highly structured narrow topics. Questions in this section focused on respondent assessments of the impediments to and opportunities for digital preservation readiness for the Delaware State Government. The goal of these open ended comment questions was to allow survey participants to highlight issues and concerns that questions in previous sections may not have raised. Specifically, these questions asked respondents to list top digital preservation priorities and impediments that a successful digital preservation program would have to overcome.

2.9.2 Findings

- Six of the 23 survey participants responded to the question about top issues and impediments to a digital preservation with the observation "Uncertain" or "I have no idea."

Digital Preservation Knowledge Level	Identified staff requirements for digital preservation	Did not identify staff requirements for digital preservation
None	4	10
Low to Very High	16	15

- Responses from the remaining 17 survey participants can be organized into the following categories:
    - EDiscovery and legal (1 response)
    - Ease of retrieval (1 response)
    - Technology incompatibilities (1 response)
    - Cost of digital preservation (2 responses)
    - Coordination of digital preservation programs across multiple agencies (3 responses)
    - Training and promoting digital awareness (4 responses)
    - Funding and staff (8 responses)
- *Almost half of the respondents view funding and staff as the top impediment to a digital preservation program for the State of Delaware*. This is an important finding that confirms a major infusion of financial support for digital preservation is required to reverse the persistent shortfall in digital preservation funding. *However, the fact of the matter is that a major infusion of funding for a digital preservation program without establishing a digital preservation infrastructure is not likely to result in a sustainable digital preservation program for the State of Delaware.* The coordination of digital preservation programs across multiple agencies along with training and promoting digital awareness are far more likely to enable the establishment and maintenance of a sustainable digital preservation program. A case in point is the creation of a collaborative state wide digital preservation policy that involves all of the stakeholders in digital preservation. This should be a top priority for the Digital Steering Committee over the next year.

3 SUMMARY FINDINGS AND RECOMMENDATIONS

This section presents five high level recommendations for immediate action that have been extracted from the findings analysis in Section 2. Several other findings will be incorporated into Deliverables 2 and 3.

The State of Delaware does not have a digital preservation policy or digital preservation program. The Delaware Public Archives has issued electronic records management guidelines and guidelines for digital imaging projects that some agencies, departments, and divisions appear to have interpreted as digital preservation policy.

In the absence of a digital preservation policy and a digital preservation program with a dedicated digital repository, agencies, departments, and divisions mistakenly rely upon DTI Backup/Business Recovery for digital preservation. The goal of DTI backup/business recovery program is speedy restoration of mission critical applications and digital objects, not long term digital preservation. This reliance upon the DTI Backup/Business Recovery system for digital preservation is not a viable alternative to a digital preservation program with a dedicated digital repository.

There is an inadequate knowledge base of the volume of digital objects that agencies, departments, and divisions create, receive, and acquire, which gives rise to an enormous impediment to systematic planning for digital preservation. This inadequate knowledge base also makes it difficult to match available resources to the greatest needs. Not all digital objects are equal because some programs are more important than others and therefore there is a greater potential risk exposure to an erosion (if not loss) in their integrity, irretrievability, and usability. Absent the baseline knowledge of what agencies, departments, and divisions are creating and the range of risk exposure to which they are subject, it is difficult to build a business case for a sustainable digital preservation program

Few, if any, agencies, departments, or divisions appear to be using applications and/or formats that can be characterized as legacy. Most agencies, departments, or divisions are using applications and/or formats that are potentially preservation ready, that is, tools exist that currently can export digital objects encoded in preservation applications and formats to newer applications and/or formats. With few exceptions, there is little reliance on the use of technology neutral open standards formats when digital objects are created, receipted, or acquired or set aside for long-term preservation.

Cohasset recommends that the Digital Preservation Readiness Steering Group consider implementing a short-term digital preservation improvement program that can quickly begin to address some of the problems described above while a comprehensive digital preservation program is being designed and implemented. Cohasset believes that quick action to implement the recommendations below can promote a digital preservation synergism to help establish a strong foundation for a sustainable digital preservation program for the State of Delaware. The recommendations are listed in a priority order that takes into account the concerns raised in the NASS/eC3 symposium report "Digital Archiving – From Fragmentation to Collaboration" and "Foundations for a Successful Digital Preservation Program: Discussions from *Preservation in State Good Practices Exchange 2006*."[13]

Recommendation 1. Adopt the recommended Delaware Digital Preservation Policy that is forthcoming in Deliverable 3.

Recommendation 2. Review the current statutory and regulatory requirements for preservation of records to clarify the role and responsibility of the Delaware Public Archives vis-à-vis those of the Department of Technology and Information and promote a collaborative framework that supports a state-wide digital preservation program.

Recommendation 3. Create a digital preservation specialist position in the Delaware Public Archives and recruit an individual to lead the Archives digital preservation program and to serve as the digital preservation resource person for state and local government agencies.

[13] *RLG DigiNews*, June 15, 2006, Volume 10. Available at www.rlg..org/en/page.php?Page_ID=20952

Recommendation 4. Increase the awareness and understanding of digital preservation issues by conducting digital preservation workshops and seminars in collaboration with agencies and departments that create and maintain digital objects based upon the potential risk exposure and importance of the digital objects.

Recommendation 5. Initiate an inventory of the volume of digital objects that require long-term digital preservation based upon the potential risk exposure and importance of the programs that create and maintain digital objects. This will enable the digital preservation program to focus on "high value" digital objects and thereby ensure that state resources are wisely invested.

APPENDIX A SURVEY REPONSE FREQUENCY DISTRIBUTIONS

The frequency distributions in this appendix are extracted from statistical reports generated by Strohl Systems software and relate only to the questions and responses that were selected for used in the analysis.

Employment and Organization Profile

Number of years in the same position:

9 (18%) fewer than 3 years
42 (83%) three years or longer

Based upon job titles the 53 respondents included:

27 (51%) Managers
12 (25%) Administrative Support Staff
7 (23%) Specialists
6 (11%) Management Analysts
8 (15%) respondents also reported they were information technology professionals

Digital preservation project begun over the last five years:

12 (24%) Yes
29 (56%) No
10 (20%) Not sure

Rating of 12 digital preservation projects:

1 Unsuccessful
4 Minimally successful
2 Moderately successful
5 Very successful

Number of hours spent on preservation per week:

31 (59%)	None
10 (20%)	1 hour
1 (2%)	2 - 4 hours
3 (6%)	5 – 10 hours
1 (2%)	11 – 20 hours
2 (4%)	More than 20 hours

Digital Preservation Knowledge and Skills

Self assessment of general knowledge and understanding of digital preservation issues:

20 (37%)	rated None
16 (32%)	rated Low
11 (22%)	rated Moderate
3 (6%)	rated High
1 (2%)	rated Very High

Level of familiarity with International Standards that bear on digital preservation

There was only one response regarding ISO 15489, ISO 14721, or ISO 18492 and it was a question if DPA Model Electronics Guidelines are similar to the ISO standards.

Level of familiarity with the Research Library Group Audit Check Guidelines on Trusted Digital Repositories

46 (89 %)	reported No familiarity[1]
4 8%)	reported Low familiarity
1 (2%)	reported Moderate familiarity

Awareness of digital preservation best practices and guidelines that could be implemented in a digital preservation program

Only four respondents answered this question

One respondent cited the State of Delaware Model Guidelines for Electronic Records

One respondent cited digital preservation programs underway in Arizona and Washington

[1] This number includes the respondent whose self-assessment on understanding of digital preservation issues was "very high."

Two respondents cited discussions with vendors that specialize in digital archives

Digital preservation training by attending a conference over the past two years:

6 (12%) reported Yes
45 (88%) reported No

Primary ways of keeping informed about digital preservation issues:

5 (10%) reported conferences
5 (10%) reported workshops and seminars
12 (24%) reported web sites
6 (12%) reported journal articles
19 (36%) did not respond

Digital Preservation Training Provided by Agency, Department, or Division:

3 (6%) reported Yes
23 (46%) reported No
25 (46%) did not respond

Policy

Agency, department, or division authority to issue a digital preservation policy:

8 (16%) reported Yes
5 (10) reported No
18 (36%) reported Do not know
20 (40%) did not respond

Agency, department, or division issuance of a formal digital preservation policy:

2 (4%) reported Yes
23 (44%) reported No
13 (25%) reported Do not know
13 (28%) did not respond

Implementation of a digital preservation policy:

5 (9%) reported Partial, or planning implementation
35 (66%) reported No implementation
11 (21%) did not respond

Create, Receive, or Acquire Digital Records

Applications used to create, receive, and acquire file formats

Microsoft Office Suite (17 responses)
Adobe Acrobat (6 responses)
Mobius (2 responses)
DFMS (2 responses)
PHRST (2 responses)
Word Perfect (1 response)
Hyperion (1 response)
Fortis (1 response)
Document Management (1 response)
Lotus Notes (1 response)
Imaging Work Flow (1 response)
File Maker Pro (1 response)

Estimated volume of digital objects created, received, or acquired that will be retained for ten years are longer:

5 (10%)	reported between 5 to 400 GB
1 (2%)	reported 900 GB
2 (4%)	reported 1 TB or more
43 (82%)	No response

Percentage of digital objects by image, text, vector graphics, spreadsheet, and database formats[2]:

14 (26%)	Do not know
7 (14%)	Greater than 50% text
2 (4%)	Greater than 50% image
2 (4%)	Greater than 50% database
3 (6%)	Greater than 50% spreadsheet
0 (0%)	Greater than 50% vector graphics
23 (43%)	No response

[2] These categories are not mutually exclusive so multiple responses occur.

Estimated growth:

2 (4%)	Greater than 50%
4 (8%)	26% to 50%
2 (4%)	11% t0 25%
4 (8%)	5% to 10%
2 (4%)	Less than 5%
2 (0%)	None
25 (49%)	Not sure
10 (19%)	No response

Conversion of digital objects to a standard format:

4 (8%)	Yes
19 (36%)	No
25 (47%)	Not sure
3 (6%)	No response

Percentage of digital objects in a legacy format:

20 (38%)	None
3 (6%)	Few
2 (4%)	Some
1 (%)	Many
1 (2%)	Most
1 (2%)	All
23 (44%)	No response

Percentage of digital objects in a potential preservation ready format:

19 (39%)	None or less than 1%
5 (10%)	Few (1% to 14%)
2 (4%)	Some (15% to 39%)
2 (4%)	Many (40% to 74%)
4 (8%)	Most (75% to 95%)
6 (12%)	All (96% to 100%)
10 (19%)	No response

Percentage of digital objects in a preservation ready format:

13 (26%) None (or less than 1%)
8 (15%) Few (1% to 14%)
4 (8%) Some (15% to 39%)
3 (6%) Many (40% to 74%)
1 (2%) Most 75% to 95%
2 (4%) All (96% to 100%)
22 (38%) No response

Archival Storage

Long term storage of digital objects:

13 (25%) Store digital objects scheduled for long term retention
16 (30%) Do not store digital objectives scheduled for long term retention
22 (42%) Did not respond

Identification of file formats currently used to store digital objects that are scheduled for long term retention:

13 (24%) reported one or more applications or formats used in the long term storage of digital objects that include:

Alchemy (1 response)
File Maker Pro (1 response)
Hyperion (1 response)
Microsoft Office (6 responses)
Adobe Acrobat (1 response)
PHR (1 response)
Web Extender (2 responses)

16 (30%) are not custodians of digital objects in long term storage
22 (45%) did not respond

File formats used in the long term storage of digital objects:

ASCII
SMP
GIF
JPEG
Lotus 123
PDF

PDF/A
RTF
SQL
TIFF
XML

Projected growth in estimated volume of digital objects in long term retention:

6	(11%)	projected No growth
9	(17%)	projected a growth of less than 10%
2	(4%)	projected a growth of between 11% to 25%
7	(13%)	estimated a growth rate of between 2% to 50%
2	(4%)	estimated a growth of greater than 50%
25	(48%)	No response

Percentage of digital objects scheduled for long term retention that are in a legacy format:

11	(22%)	None
5	(10%)	between 1% and 14%
5	(10%)	between 15% and 39%
0	(0%)	between 40% and 74%
0	(0%)	between 75% and 95%
2	(4%)	between 96% and 100%
28	(53%)	No response

Percentage of digital objects in long term storage that are in a potential preservation ready format:

17	(32%)	None
2	(4%)	between 1% and 14%
4	(8%)	between 15% and 39%
1	(1%)	between 40% and 74%
0	(0%)	between 75% and 95%
2	(4%)	between 96% and 100%
24	(45%)	No response

Percentage of digital objects in term storage that are in a preservation ready format:

12	(32%)	None

4 (8%) between 1% and 14%
2 (4%) between 15% and 39%
1 (1%) between 40% and 74%
0 (0%) between 75% and 95%
2 (4%) between 96% and 100%[3]
29 (55%) No response

Storage location of digital objects[4]:

5 (10%) reported work station hard disk
7 (14%) reported removable storage media
17 (32%) reported network shared drive
4 (8%) reported outsourcing the storage of digital objects
20 (38%) reported "Do not know"
11 (21%) No response

Outsourcing storage of digital objects scheduled for long-term retention:

4 (8%) Yes
21 (40 %) No
22 (42%) Not sure
4 (8%) No response

Estimated volume of stored records by format:

38 (74%) No response
13 (26%) Response
Text
0 less than 5%
1 between 5 and 10%
2 11 – 25%
1 26 – 50%
4 5 0% or more
Vector Graphics

[3] These two respondents also indicated that between 96% and 100% of their digital objects were in a preservation ready format. Responses to this question and the previous one were intended to be mutually exclusive but some respondents may have missed this distinction.

[4] These categories are not mutually exclusive so multiple responses occur.

2 less than 5%
2 between 5 and 10%
0 11 – 25%
1 26 – 50%
0 50% or greater

Database

4 less than 5%
1 between 5 and 10%
1 11 – 25%
3 26 – 50%
0 50% or greater

Image

0 less than 5%
3 between 5 and 10%
2 11 – 25%
5 26 – 50%
1 50% or greater

Spreadsheet

2 less than 5%
1 between 5 and 10%
3 11 – 25%
2 26 – 50%
0 50% or greater

Access and Dissemination

Public access to and dissemination of digital objects:

7 (14 %) Yes
11 (22%) No
33 (64%) No response

Percentage of digital objects by image, text, vector graphics, spreadsheet, and database formats[5]:

Greater than 50%

0 Text
1 Image

[5] These categories are not mutually exclusive so multiple responses occur.

0 Vector graphics
0 Spreadsheet
0 Database

Between 26 % to 50%

0 Text
1 Image
0 Vector graphics
0 Spreadsheet
1 Database

Between 11% to 25%

1 Text
1 Image
1 Vector graphics
1 Spreadsheet
1 Database

Applications used to support public access to and dissemination of digital objects:

Alchemy (Document Management Software)
Hyperion (Library software for web enabled search and retrieval)
Past Perfect Museum (Historical collection management software)
Microsoft Office

Outsourcing of access to and dissemination of digital objects:

0 (0%) Yes
7 (14%) No
11 (22%) Do not support public access to or dissemination of digital objects
33 (64%) No response

Human, Financial Resources

Annual budget for department or division:

9 (18%) Not sure
42 (82%) No response

Percentage of annual budget that supports information technology services:

6 (12%) Less than 5%
6 (12%) Between 5% and 10%
0 (0%) Between 11% and 25%
2 (4%) Between 26% to 50%

2 (4%) Greater than 50%
1 (2%) None
34 (66%) No response

Number of Full Time Equivalent (FTE) staff involved in the protection of vital records and disaster recovery/business continuity:

4 (8%) .50 FTE
5 (10%) .51 to 1.5 FTE
4 (8%) 1.51 to 4.0 FTE
2 (4%) 4.01 to 8 FTE
2 (4%) Greater than 8.0 FTE
11 (21%) Not sure
23 (45%) No response

Number of FTE required to successfully implement a digital preservation program:

6 (12%) .5
7 (14%) 1.0
5 (10%) 2.0
1 (2%) 3.0
32 (62%) No response