Digital Preservation Strategy for the State of Delaware

Digital Preservation Planning Project

*Working Paper – Discussion Draft*

**Deliverable 3**

August 15, 2007
## CONTENTS

Introduction ........................................................................................................................................... 1

1 Long-Term Implications of Failing to Improve the State of Digital Preservation .................... 3
  1.1 Background ................................................................................................................................... 3
  1.2 Long-Term Implications ................................................................................................................. 4

2 Digital Preservation Policy .................................................................................................................. 8
  2.1 Policy Rationale ............................................................................................................................... 8
  2.2 A Digital Preservation Mission Statement ...................................................................................... 8
  2.3 Digital Preservation Policy Background ......................................................................................... 9
  2.4 Digital Preservation Policy ............................................................................................................. 9

3 Digital Preservation Strategy ............................................................................................................. 11
  3.1 Background ................................................................................................................................... 11
  3.2 Digital Preservation Stakeholder Value Strategy ............................................................................. 11
    3.2.1 Discussion ............................................................................................................................... 11
    3.2.2 Recommendation .................................................................................................................... 12
  3.3 Digital Preservation Economic Value ............................................................................................... 12
    3.3.1 Discussion ............................................................................................................................... 12
    3.3.2 Recommendation .................................................................................................................... 12
  3.4 Digital Preservation Infrastructure Strategy .................................................................................... 12
    3.4.1 Roles and Responsibility .......................................................................................................... 12
    3.4.2 Digital Preservation Collaboration Awareness Strategy ........................................................... 14
    3.4.3 Digital Preservation Technical Expertise Strategy .................................................................... 15
  3.5 Digital Preservation Processes Strategy ......................................................................................... 16
    3.5.1 Background ............................................................................................................................... 16
    3.5.2 Open Archives Information Systems (OAIS) Model ................................................................. 16
    3.5.3 Digital Preservation Strategy .................................................................................................... 17
    3.5.4 Technology Neutral Open Standard File Formats ................................................................. 19
    3.5.5 Technology Watch ..................................................................................................................... 20
    3.5.6 Ingest ....................................................................................................................................... 21
    3.5.7 Storage .................................................................................................................................... 22
    3.5.8 Planned Media Renewal ............................................................................................................ 23
    3.5.9 Digital Object Authenticity ....................................................................................................... 24
3.5.10 Security ................................................................................................................. 25
3.5.11 Access .................................................................................................................. 26
3.5.12 Digital Preservation Metrics ................................................................................. 27
4 Digital preservation policy and strategy ....................................................................... 28
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 Deliverable 3 that recommends a digital preservation strategy for the State of Delaware. This digital preservation strategy is a high level planning document that subsequently will be used to drive the phased implementation of a sustainable digital preservation program. A sustainable digital preservation strategy, however, does not have an independent existence. Rather, it is the logical extension of a digital preservation vision and digital preservation policy. Consequently, Deliverable 3 takes into account the need for such a vision and policy.

Deliverable 3 is organized into three sections. Section 1 reviews the current state of digital preservation in the State of Delaware with special attention on long-term implications of failing to implement a sustainable digital preservation program. Section 2 proposes a digital preservation vision and policy for the State of Delaware. Section 3 reviews optional digital preservation strategies and recommends a digital preservation strategy for the State of Delaware.

---

1 A digital preservation policy articulates the high level requirements for achieving a digital preservation vision while a digital preservation strategy delineates at a high level how this vision and policy will be carried out. To put it differently, a digital preservation policy stipulates what will be done at a high level while a digital preservation strategy identifies at a high level how this will be done.
1 LONG-TERM IMPLICATIONS OF FAILING TO IMPROVE THE STATE OF DIGITAL PRESERVATION

1.1 Background
In Deliverable 2 Cohasset developed a digital preservation readiness capability maturity model that included two key components. The first component consisted of five digital preservation infrastructure and seven digital preservation process requirements. The second component included five incremental levels of digital preservation capability - nominal, minimal, intermediate, advanced, and optimal – in which each level incorporated all of the capabilities of previous levels. A numeric value – 0, 1, 2, 3, and 4 – representing each of the five levels was assigned to the 12 digital preservation requirements. The findings from the digital preservation survey (Deliverable 1) were mapped to these 12 digital preservation requirements with the following results:

♦ A statutory mandate and supporting regulations for the long term preservation of information objects exist but there is some uncertainty as to which organization has ultimate authority for the program. (Level 2, numeric value “1”)
♦ A formal written digital preservation policy does not exist. (Level 1, numeric value “0”)
♦ A systematic digital preservation strategy does not exist or if it exists, it is not being implemented. (Level 1, numeric value “0”)
♦ There is little or no awareness of the role of collaboration in a digital preservation project. Digital preservation projects do not take into account the potential for collaboration with other parties having an interest in digital preservation. (Level 1, numeric value “0”)
♦ Little or no professional expertise in digital preservation exists. (Level 1, numeric value “0”)
♦ At least one textual and one digital image technology neutral open standard format for digital preservation has been adopted that protects the content and structure of digital objects. (Level 2, numeric value “1”)
♦ No dedicated logical or physical digital preservation repository exists. (Level 1, numeric value “0”)
♦ No formal media renewal procedure is in place. (Level 1, numeric value “0”)

COHASSET ASSOCIATES, INC. - - - Discussion Draft - - - August 15, 2007
Deliverable 3 Digital Preservation Strategy

- No procedure is in place to validate the integrity of digital objects. (Level 1, numeric value “0”)
- Disaster recovery and backup processes are network based. Physical security is assured through role-based permission access. (Level 3, numeric value “2”)
- Preservation metadata for digital objects of long-term value is collected and maintained on an ad hoc basis. (Level 2, numeric value “1”)
- Some digital objects of long-term value are electronically accessible but only in ASCII text, TIFF images, or PDF. (Level 3, numeric value “2”)

Summing these numeric values produced a digital preservation score of “7,” which is a mid-range nominal score. This score, however, obscures the fact that on 8 of the 12 preservation requirements the current state of digital preservation is “nominal.” In truth, the current state of digital preservation is closer to being “Nominal” rather than “Minimal.”

1.2 Long-Term Implications

Some observers of the current state of digital preservation in the State of Delaware might argue there is no real impending digital preservation crisis. No digital objects have been reported to have become unusable or irretrievable, nor has there been a legal challenge to the authenticity of digital objects. The Delaware Public Archives has an ongoing program to scan selected historical documents, maps, and photographs and to post these images on its Web site where they are accessible to any one. This program is not likely to cease nor is it likely that the Delaware State Library will discontinue a program to mount scanned images of selected publications on its Web site as part of facilitating citizen access to digital publications. In short, these observers would argue that the “digital sky is not falling in” and there is no urgency to take immediate action. Nothing is broken so nothing needs to be fixed. Maintaining the status quo is sufficient.

This “mind set” is understandable. But it is a mindset that is reminiscent of the awareness (or lack thereof) many people have about the safety condition of bridges. Most bridges are not falling down (the Minneapolis-St. Paul I35W is an exception): traffic continues to move smoothly and ongoing repair and maintenance work seem to be sufficient. Nonetheless, the collapse of one span of the Minneapolis-St. Paul I35W Bridge has raised serious questions about assumptions drivers have about the safety of bridges they traverse. Undoubtedly, further exploration of the cause(s) of failure of the Minneapolis-St. Paul I35W Bridge will highlight numerous problems – flawed construction techniques, unanticipated deterioration of the physical structure, vehicle loads that the bridges were not designed to accommodate, and delayed repairs, among others – that must be corrected in order for bridges to be safe.
Although the current state of digital preservation in Delaware is unlikely to cause death or injury to any citizen, probing beneath the surface of digital preservation has identified problems that if left uncorrected over time are likely to cause serious legal and financial challenges. A case in point is that the current state of digital preservation can give rise to an environment in which there is a significant legal risk exposure because digital objects that are required for ensuring regulatory compliance are not readily identified, accessible or trustworthy. This inability to have access to usable, retrievable, and authentic digital objects for use in litigation could undermine legal efforts to ensure regulatory compliance. Even more compelling is the need for a systematic approach to comply with “e-discovery requirements” laid out in the revised Federal Rules of Civil Procedure that took effect in December 2006. If digital objects scheduled for long term retention fall under the purview of a “discovery order,” the failure of the State to have in place a digital preservation program that ensures reasonable and timely access to digital objects through a sustainable digital preservation may result in adverse legal and financial consequences.

There are several other areas where the consequences of failing to improve the current state of digital preservation are less conspicuous but equally disastrous. The absence of a digital preservation policy gives rise in essence to a dysfunctional infrastructure in which each agency, department, or business unit pursues its own vision (or lack thereof) digital preservation without an overarching commonality of what is best for the State of Delaware and its citizens. Individual digital preservation projects may emerge that “reinvent the digital preservation wheel” and fail to take advantage of lessons learned from unsuccessful digital preservation activities. Without a compelling digital preservation vision driving investment in digital preservation there will be little measurable return on that investment.

Another conspicuous area of digital preservation failure is the absence of a unified digital preservation strategy. A unified digital preservation strategy describes how a digital preservation policy will be executed, what the priorities are, exemplary practices that will be followed, and metrics that will be used to measure success. A unified digital preservation strategy inspires the confidence of stakeholders in the preservation of digital objects that the digital preservation policy is being properly executed. The absence of a unified digital preservation strategy is a strong indicator of fragmentation and chaos in digital preservation. Furthermore, it helps to perpetuate other areas of digital preservation failure, including the absence of digital preservation collaboration.
Collaboration can be characterized as both a mindset and a process. The collaboration mindset sees a particular problem as having ramifications for other parties and seeks opportunities to communicate this perspective. Once the desirability of collaboration has been established, a process with specific procedures is in place to facilitate collaboration that lead to leveraging resources to achieve a good greater than what a single individual or organization can accomplish. This is no less true of digital preservation collaboration. Absent a digital collaboration mindset and a process, any digital preservation effort is likely to be narrowly defined and to operate in a vacuum so that it is difficult to leverage successes and to avoid repeating failures.

Technology obsolescence is the major challenge that a digital preservation program must address. It is inevitable and unless it is addressed within a “technology obsolescence time window” it is irreversible. Newer technology neutral open standard formats that are backwardly compatible with older ones provide a reliable path that mitigates technology obsolescence. Some observers may argue that because there are few legacy digital records in Delaware State agencies, departments, and divisions (Deliverable 1) technology obsolescence is not an issue. However, the clear message from the Rescue and Salvage case studies reported in Deliverable 2 is that digital objects that are forgotten about, ignored, or have a very low digital preservation priority and are not migrated to technology neutral formats in a timely manner eventually come back to “bite” their owners. This “bite” is likely to be in the form of heroic and very costly digital object rescue initiatives or a technology obsolescence so advanced that a “window of opportunity” is tightly closed and the digital objects are beyond the reach of any rescue or salvage undertaking.

Technology neutral open standard formats support the usability of digital objects over time but they do not assure their integrity over time. The underlying bit streams of digital objects are vulnerable to alteration and corruption that may be accidentally introduced during preservation activities with no obvious evidence that they occurred. More importantly, as digital objects are widely disseminated from Web sites there is a potential for malicious alterations, some minor and others major, to occur. The history of fraudulent paper based documents to support ideological, financial, political, or criminal agendas suggests that the problem is likely to be more pervasive in a digital environment. In this digital environment users will want assurance that digital objects are authentic, that is, they are what they purport to be because the integrity of the underlying bit streams can be validated. Absent this integrity validation capability, which is
a key component of a sustainable digital preservation program, future users of digital objects will have little confidence that the digital objects are authentic. Relying upon the good intentions of users to protect the integrity of digital objects rather than integrity validation tools gives rise to an environment in which the authenticity and integrity of all digital objects are suspect.
2 DIGITAL PRESERVATION POLICY

2.1 Policy Rationale
A digital preservation policy is not a panacea that resolves all issues and concerns of digital preservation. However, it is the essential foundation for a sustainable digital preservation program. A digital preservation program without a digital preservation policy will not have clear vision of where it is going so that ultimately going anywhere no matter how wrongheaded will be the measure of success.

2.2 A Digital Preservation Mission Statement
A clear and compelling mission statement articulates the general direction in which an organization plans to move. Although a mission statement simplifies hundreds, if not thousands, of detailed decisions that must be made, it implicitly defines the core values that drive an organization to accomplish certain goals. This is no less true for a digital preservation mission statement. It should articulate the general direction in which digital preservation will move and what will be accomplished. It should reflect the core values that undergird digital preservation.

A digital preservation mission statement for the State of Delaware communicates the purpose of digital preservation to the executive, judicial, and legislative branches of government as well as to citizens of the state and other interested parties. It answers the question at a high level of why a digital preservation program is important. Deliverable 2 contained a proposed digital preservation mission statement that Cohasset believes is appropriate for the State of Delaware.

  Protect the legal and financial rights of state government, local governments, and citizens and information about people, places, and events in Delaware history that are captured in digital objects through a sustained digital preservation program that ensures their accessibility, authenticity, and usability for current and future users.

This is a dynamic, open-ended mission statement that incorporates the commitment of the State of Delaware to the proposition that digital objects bearing on the legal and financial rights of government and citizens or containing noteworthy information about people, places, and events in Delaware history that are related to continuing self-educational, political, culture, economic,
recreational, and intellectual enrichment will be accessible, usable, and authentic for as long as necessary. In effect, it declares that technology obsolescence will not be a barrier to the accessibility, usability, and authenticity of digital objects for the government and citizens. The Digital Preservation Policy that follows in the next section builds off this digital preservation mission statement.

2.3 Digital Preservation Policy Background
The mandate for this digital preservation policy takes into account requirements in the Delaware Code that relate to the Government Information Center [Title 29, Chapter 87, §8707], the Department of Technology and Information [Title 29, Chapter 90C, §9004C (2), §9006C(d)], the Division of Libraries [Title 29, Chapter 87, §8731], and the Delaware Public Archives [Title 29, Chapter 5, §502, §503]. In addition, this digital preservation policy draws upon four primary guiding documents in digital preservation:

♦ *An Audit Checklist for the Certification of Trusted Digital Repositories* (Research Library Group, 2005)
♦ *ISO 14721, Open archival information System (OAIS) - Reference model* (2003)
♦ *Digital Preservation Readiness Capability Maturity Model and Digital Preservation Readiness Balanced Scorecard* (Deliverable 2, August 10, 2007)

Deliverable 2 is especially important because the Digital Preservation Readiness Balanced Scorecard section implicitly articulates both a policy and a strategy.

2.4 Digital Preservation Policy
Cohasset recommends the following Digital Preservation Policy.

*This Digital Preservation Policy formalizes the commitment of the State of Delaware to the long-term preservation of digital objects that protect the legal and financial rights of state and local governments and citizens or contain information about people, places, and events of Delaware history related to continuing self-educational, political, culture, economic, recreational, and intellectual enrichment. Digital objects may be either “born-digital,” that is, they are created, used, and stored in a digital format, or scanned digital images of paper documents. Furthermore, this digital preservation policy declares that*
digital objects of long-term value are critical assets of the State and their preservation shall be carried out through a sustained digital preservation program supported by stakeholders in the current and future use of accessible, usable, and authentic digital objects.

This Digital Preservation Policy involves four primary interrelated areas of activity and one supporting level of activity:

1. Digital preservation requires promotion of the value of digital objects to stakeholders who include current and future users of the digital objects

2. Digital preservation requires articulation of the economic value a digital preservation program has for current and future users

3. Digital preservation requires implementation of appropriate digital preservation processes that mitigate the effects of technology obsolescence

4. Digital preservation requires a technical infrastructure capable of supporting digital preservation processes

5. This Digital Preservation Policy (and supporting Digital Preservation Strategy shall be communicated to all senior managers and staff of state and local governments who create, receive, or acquire digital objects in the ordinary course of business, to stakeholders in digital preservation, and to the public at large.
3 DIGITAL PRESERVATION STRATEGY

The discussion of a digital preservation strategy for the State of Delaware in this section merges the findings of Deliverable 2 regarding an incremental digital improvement program and measurable performance metrics with the digital preservation policy presented in section 2 of this deliverable.

3.1 Background
Deliverable 2 delineated five levels of digital preservation capability – Nominal, Minimal, Intermediate, Advanced, and Optimal – and identified the current state of digital preservation in the State of Delaware as being midway between Nominal and Minimal. Based on this incremental improvement model, implementation of the preservation strategy presented in this section over the next five years or so could improve the current state of digital preservation to Advanced. The actual formulation of this recommended digital preservation strategy flows out of the four broad areas of digital preservation policy.

3.2 Digital Preservation Stakeholder Value Strategy

3.2.1 Discussion
A sustained digital preservation program requires the engagement and support of digital preservation stakeholders. Unfortunately, many citizens, students, researchers, public interest groups, local and state government agencies, and business organizations may not realize or appreciate how a digital preservation program protects their legal, financial, and informational rights. To correct this problem an outreach program must be put in place that identifies these stakeholders, provides them a forum within which to exchange views, and engage public officials on the value of a robust digital preservation program. Among the tools a digital preservation stakeholder outreach program may use are a newsletter or blog, newspaper and magazine articles, and a Digital Preservation Coalition.
3.2.2 Recommendation
Cohasset recommends the following Digital Preservation Stakeholder Value Strategy:

An incremental Digital Preservation Stakeholder Value Program will be established to identify stakeholders, engage them on the direct value to them of a robust digital preservation program, and promote public support for the program.

3.3 Digital Preservation Economic Value

3.3.1 Discussion
Like the cost to preserve paper-based information objects, the cost to preserve digital objects is likely to increase. Given the limited financial resources currently available to fund public programs, competition for these resources is likely to increase, at least in the near future. Success in obtaining the necessary funding for a digital preservation program requires a clear and persuasive articulation of the economic value of a digital preservation program. Two ways to develop this persuasive articulation of economic value that immediately come to mind are (1) making a business case based on future cost avoidance of the rescue and salvage of legacy digital objects and (2) increased revenue streams. Careful study of this issue undoubtedly would identify other ways to present the economic value of digital preservation.

3.3.2 Recommendation
Cohasset recommends adoption of the following Digital Preservation Value Strategy:

Incremental business case studies and other relevant techniques will be used to identify the economic value of digital preservation for the State of Delaware.

3.4 Digital Preservation Infrastructure Strategy

3.4.1 Roles and Responsibility

3.4.1.1 Discussion
One important aspect of today’s digital technologies is an inherent capacity to break down or ignore the traditional boundaries separating disciplines and organizations. Unfortunately,
organizational governance, especially in the public sector, has been slow to adjust to this change as defined roles and responsibilities have not been updated to address this fundamental issue. The result is curious amalgam of organizational fragmentation in terms of defined roles and responsibilities and integration of information technologies.

This is no less true for digital preservation in the State of Delaware. Based upon statutory mandates, there are three key players in digital preservation at the state level - the Delaware Public Archives (DPA), the State Library, and the Department of Technology and Information. The primary responsibility of the Delaware Public Archives is for the preservation of digital objects created, received, or acquired by state and local government agencies. The State Library’s primary responsibility is for the access to and preservation of electronic publications produced by and for state and local government agencies. The primary responsibility of the Department of Technology and Information is for the technical infrastructure supporting state agency programs and ensuring system compliance with approved digital technology standards and practices.

A careful reading of these statutory mandates reveals overlapping roles and responsibilities. For example, a sustained digital preservation program cannot exist without the use of technology neutral standards but the Department of Technology and Information is largely responsible for promulgating technology standards and ensuring agency compliance with them. The same point could be made about metadata standards supporting access to digital objects in the custody of the Delaware Public Archives and the State Library. To improve the governance of digital preservation vis-à-vis digital information technologies, the Delaware Public Archives, the State Library, and the Department of Technology and Information should identify these overlapping roles and responsibilities and develop operational ways to ensure that they are reading from the “same page” and supporting a shared digital preservation agenda.

Identification of these overlapping roles and responsibilities and development of a shared digital preservation agenda can be achieved through a formal agreement between the Delaware Public Archives, the State Library, and the Department of Technology and Information to identify overlapping roles and responsibilities and to develop a shared digital preservation agenda. In some instances it may be possible to work out a modus operandi that achieves desired objectives while in other areas it may necessary to propose revisions to statutory mandates and regulations.
3.4.1.2 Recommendation
Cohasset recommends the following strategy for clarifying the digital preservation roles and responsibilities of the Delaware Public Archives, the State Library, and the Department of Technology and Information:

The Delaware Public Archives, the State Library, and the Department of Technology and Information will put in place a formal agreement to review their respective roles and responsibilities with regard to digital preservation and to develop a shared on-going digital preservation agenda that takes into account any clarification about these roles and responsibilities that may be required over time.

3.4.2 Digital Preservation Collaboration Awareness Strategy

3.4.2.1 Discussion
Collaboration among multiple entities promotes leveraging resources, including technical expertise, and subject matter expertise, that achieves a multiplier effect that is not possible when undertaken by a single entity. Digital preservation issues and concerns frequently cut across the boundaries of organizational structures and technical expertise so a collaborative awareness mindset can play a key role in a successful digital preservation program. Leveraging the subject matter expertise and technical expertise through collaboration can create a digital preservation multiplier effect. The key to collaboration is having a collaborative awareness mindset that continuously scans work programs to identify initiatives that relate directly or indirectly to digital preservation. One useful way to promote a digital preservation readiness collaboration awareness is to include a digital preservation readiness impact statement in the business case template that the Department of Technology and Information requires state agencies to complete as part of the approval process for procurement of new systems and applications. A case in point is the Enterprise Geographic Information System project currently under development. Another instance would be periodic postings to all state employees about the status of digital preservation and soliciting comments and suggestions about opportunities for collaborative digital preservation projects.

3.4.2.2 Recommendation
Cohasset recommends adoption of the following Digital Preservation Collaboration Awareness Strategy.
3.4.3 Digital Preservation Technical Expertise Strategy

3.4.3.1 Discussion
A sound digital preservation program requires digital preservation technical expertise that can address all aspects of the program. Some aspects of digital preservation technology expertise are likely to involve organizing a physical or logical digital preservation repository, including hardware and software specifications. Other aspects are likely to include having skill sets to manage the periodic copying of digital objects to new storage media, conduct annual readability inspections, recommend migrating digital objects from old technology neutral open standard formats to new ones, to work with Agency/Department Information Resource Managers and Records Officers to ensure preservation ready digital objects are created, received, or acquired in the normal course of business, and advise the Delaware Public Archives on specific digital preservation issues that emerge in the future.

A strategy to recruit new recruits who have professional training in digital preservation technologies may take considerable time to succeed because such individuals are in short supply. A far more successful strategy, at least in the term, is to have temporary contracts with information technology consultants who can provide specific services. Supplementing these short term consulting contracts could be digital preservation training for selected employees that include customized courses and digital preservation webinars and workshops. Of course, the long-term and short strategies discussed here cannot be implemented without adequate financial support.

3.4.3.2 Recommendation
Cohasset recommends adoption of the following Digital Preservation Expertise Strategy:

>The Digital Preservation Steering Committee will develop and monitor implementation of plans and scenarios for the promotion of digital preservation collaboration awareness. Such plans and scenarios may include, for example, adding a digital preservation impact statement to the business case procurement template for new systems and applications.
3.5 Digital Preservation Processes Strategy

3.5.1 Background
The specific components of the digital preservation process strategy discussed in this section are a synthesis of the Digital Preservation Readiness Preservation Processes and associated performance metrics in Deliverable 2 and high level requirements of the International Standards Organization 14721 Open Archives Information System Reference Model (2003). Basically, there are three digital preservation strategies on how to mitigate the technology obsolescence of file formats that over time will render digital objects unusable.

3.5.2 Open Archives Information Systems (OAIS) Model

3.5.2.1 Discussion
The overall framework of the Delaware Digital Preservation Program is based on the Open Archives Information System Reference Model which is technology independent and agnostic with regard to a specific digital preservation strategy. It consists of three primary components: Ingest, Storage, and Access. Ingest is the function that prepares digital objects determined to have long-term value for transfer to a digital preservation repository. It is a temporary holding area with no immediate or direct access to the digital preservation repository. Storage essentially consists of a logical or physical repository within which all of the preservation actions and activities necessary to extend the retrievability, usability, and authenticity of digital objects occur. Access is the function that delivers usable and authentic digital objects to a user community. Typically, copies of selected digital objects are copied from those in Storage and posted on an access web site so that no user ever has direct access to digital objects in Storage.

3.5.2.2 Recommendation
Cohasset believes that the Open Archives Information System Reference Model provides a framework for the development and implementation of a robust digital preservation strategy.

Meet the need for sufficient technical expertise to support a fully operational digital preservation program by recruiting one or more new recruits with professional training in digital preservation technologies. In addition, utilize short term consulting contracts with information technology consultants to provide specific digital preservation services and contract with third parties to develop customized digital preservation training courses for selected employees.
Accordingly, Cohasset recommends its adoption and incorporation into the Delaware Digital Preservation Program.

The Delaware Digital Preservation Program will use the Open Archives Information Systems Reference Model to frame its overall digital preservation strategy.

3.5.3 Digital Preservation Strategy

3.5.3.1 Discussion

Generally speaking, there are three digital preservation strategies on how to mitigate the technology obsolescence of storage media and file formats that over time will render digital objects unusable. Below is a discussion of each of the digital preservation strategy along with a listing of advantages and disadvantages.

Digital Archaeology

Digital archaeology involves preserving the original bit streams as legacy digital objects that are in application dependent native formats through planned media renewal and retaining all of the hardware, operating system, and application software along with appropriate documentation required to retrieve and use the digital objects. Digital archaeology enters the picture, so to speak, because over time the operating system, hardware, and application software will no longer work. Once this happens the original bits can be rendered only by writing computer code that replicates the software functionality on a contemporary operating system and hardware. Of course, it is possible that in the intervening years a third party may have written conversion software that would reliably render the original bits.

The digital archaeology strategy supports accessible, usable, and authentic digital objects on demand. In other words, nothing is done to legacy digital objects other following a planned media renewal program until there is a request for them.

Advantages

♦ No limitations on file formats that are acceptable
♦ Low front-end cost because normalization to a standard format is not required
♦ User requests drive the conversion of legacy digital objects to contemporary operating systems, hardware, and software applications
Except for on demand conversion of legacy digital objects, digital preservation essentially consists of extending the life of bit streams through planned media renewal.

**Disadvantages**
- Obtaining the intellectual property right to obsolete software and operating systems may be difficult.
- Creates a huge backlog of legacy digital objects that are unusable until new software applications are written.
- The case studies of rescuing legacy digital objects show that it is very expensive and difficult to do.
- The cost of on-demand conversion is unknown.

**Migration**

Migration is the process that mitigates digital technology obsolescence by converting digital objects in obsolete (or soon to be) to new file formats that can be processed by contemporary computers and software. It relies on technology neutral open standard file formats that are backwardly compatible so no special computer code needs to be written to enable format migration.

**Advantages**
- A proven methodology that has been replicated innumerable times.
- Relatively low cost per GigaByte.
- Technology neutral open standard formats are driven by the market place so backward compatibility between new file formats and old file formats is a standard feature.
- Technology neutral open standard formats with substantial market penetration are likely to persist over time.

**Disadvantages**
- Assumes that new software applications will always be able to open technology neutral open standard file formats.
- Migration is a never ending process because in the foreseeable future file formats will continue to change so the migration cycle must be repeated each time a new format replaces an old format.
- Migration relies upon backward compatibility between new file formats and old file formats.
- Migration only deals with the file formats of digital objects so it is not appropriate for digital objects that rely upon software applications with dynamic features, such as a Geographic Information System.

**Emulation**
In preservation emulation a computer program called an emulator that is written for a contemporary computer replicates the functionality of an obsolete operating system and software so that legacy digital objects can be retrieved, opened, and used without any loss in their original structure and content (i.e., their original appearance). Because preservation migration uses legacy digital objects in a native format, there is no requirement to periodically convert digital objects from an old format to a new one. The only requirement is to protect the bit stream of the digital objects from degradation through periodic media renewal. Emulation can replicate all of the functionality of the original software, which can be important when the software had interactive features that produced logical views of digital objects, as in the case of a Geographic Information System.

**Advantages**

- Ensures preservation of digital objects as viewed by their creators or recipients
- Legacy digital objects require no conversion to new file formats
- An emulator written for one specific operating system and software can be

**Disadvantages**

- Implementations of emulation have been to limited demonstration, research, or rescue and salvage projects so emulation tools are not commercially available
- Multiple emulators must be written over time for new operating systems and hardware to ensure the continued accessibility and usability of legacy digital objects

3.5.3.2 **Recommendation**
Cohasset believes that migration is the preferred preservation strategy that the Delaware Digital Preservation Project should adopt.

> Technology obsolescence shall not impede or otherwise prevent the extending the retrievability, usability, and authenticity of digital objects for current and future users. A migration strategy will be employed to achieve this goal.

3.5.4 **Technology Neutral Open Standard File Formats**

3.5.4.1 **Discussion**
A file format is set of specifications associated with the bit stream underlying a digital object that tell a computer how to interpret the bit stream and render it as an object understandable to humans. A technology neutral open standard file format essentially is a non-proprietary
consensus based standard issued by organizations such as the International Standards Organization (ISO) or the American National Standards Institute (ANSI). Technology neutral open standard file formats share several characteristics:

- Run on multiple platforms
- Requires no special software
- Transparent (publish specifications)
- Self-contained
- Widely implemented/adopted

The leading technology neutral open standard file formats in use today include XML, ODF, PDF, PDF/A, TIFF, JPEG, MPEG, and PNG. The migration preservation strategy discussed previously normalizes digital objects to one of these technology neutral open standard file formats and any future ones. Many software applications use a native file format (e.g., Microsoft Word) but also support the conversion of digital objects to a technology neutral open standard format such as PDF and PDF/A, which can greatly simplify normalization.

Cohasset believes that the use of technology neutral file formats should be phased in incrementally over five years or so, beginning with TIFF and JPEG for digital images and PDF for textual documents. During this five year transition period the owners of digital objects should be strongly encouraged to use technology open standard file formats at the point of creation, receipt, or acquisition because this reduces the workload of normalizing digital objects to a technology neutral open standard format.

3.5.4.2 Recommendation
Cohasset believes that technology neutral open standard formats are critical to a sustained digital preservation program and recommends their adoption.

Incrementally adopt technology neutral open standard formats for use in the storage of digital objects, beginning with TIFF, JPEG, and PDF.

3.5.5 Technology Watch

3.5.5.1 Discussion
Operating systems, software applications, file formats, and storage media widely used today eventually will succumb to technology obsolescence. Digital technology obsolescence may
occur quickly because of the discovery of a totally new technology that has little connection with current technologies or it may occur in an evolutionary mode because backward compatibility tools and procedures are in place. In either case, a sound digital preservation program will monitor trends in storage media, file formats, operating systems, hardware, and software to ensure that appropriate actions are taken well in advance of technology obsolescence. This can be accomplished through a Technology Watch Program that collects information about trends in digital technology obsolescence and alerts the program when corrective action is necessary. A Delaware Digital Preservation Technology Watch Program can obtain up to date information from programs such as the UK Digital Preservation Coalition Technology Watch Program and reports issued by the National Archives of the United States.

3.5.5.2 Recommendation
Cohasset believes that a Technology Watch Program that monitors trends in digital technology obsolescence is an important component in the Delaware Digital Preservation Program. Accordingly, Cohasset recommends adoption of a Technology Watch Program as part of a Digital Preservation Strategy.

A Technology Watch Program will be used to monitor trends in technology obsolescence and alert the Delaware Digital Preservation Program when action must be taken to avert technology obsolescence.

3.5.6 Ingest

3.5.6.1 Discussion
From an OAIS framework perspective Ingest is a function that receives digital objects and prepares them for storage in the repository. The physical delivery of digital objects can be accomplished in a variety of ways, including FTP and removable media such as Cds, DVDs, and magnetic tape. Determination of the physical form of transfer is likely to depend upon the volume of the digital objects and the data transfer rate. The Ingest function should confirm that the digital objects received from their owner in fact are complete and they are what they purport to be, check the digital objects for a virus, and quarantine them until the virus is fixed. In addition, the Ingest function should identify digital objects that are encrypted or protected by a password and place a hold on them until these impediments to future file format migration are
removed. Finally, the Ingest function should normalize digital objects to the appropriate technology neutral open standard file format(s).

3.5.6.2 Recommendation
Cohasset believes that selected OAIS Ingest functions play an important role in a sound digital preservation program and recommends their incorporation into the digital preservation strategy.

Receive complete and authentic digital objects from their owners and prepare them for transfer to the storage repository. Included among the preparation techniques to be used are placing a hold on encrypted or password protected digital objects until they are decrypted and passwords removed, checking digital objects for virus infections and quarantining them until fixed, and normalizing digital objects to an appropriate technology neutral open standard format.

3.5.7 Storage
3.5.7.1 Discussion
The content of the digital repository should reside on a physical (as opposed to logical) storage system that is exclusively dedicated to digital preservation, segregated from an operational environment, and protected from external intrusion by a robust firewall. The storage system should support ingesting of digital objects and exporting copies of digital objects to an external access application. The storage system itself may consist of direct access devices or a robotic library of removable storage media, depending upon the volume of digital objects, the frequency of ingesting digital objects or exporting copies of digital objects to an access application, and the data transfer rate required. In addition, the dedicated digital preservation repository system should support planned media renewal (discussed in greater detail later) and migration of digital objects from old formats to new formats in order to extend their retrievability and usability. The digital repository should be backed up daily for disaster recovery purposes. In addition, a second set of backup tapes shall be stored in an environment controlled off site storage location. A key question is who is the “owner” of this digital repository. Given its statutory mandate, the Delaware Public Archives is the preferred “owner.”
3.5.7.2 Recommendation
Cohasset believes that the Delaware Digital Preservation System requires a robust storage system and recommends adoption of the storage system strategy that follows.

The storage of digital objects utilizes a physical digital repository that is hosted on a storage system exclusively dedicated to digital preservation, segregated from an operational environment, and protected from external intrusion by a robust firewall. This storage supports ingesting of digital objects, exporting digital objects to an external access system, periodic media renewal, migration from old formats to new formats, and backup copies for disaster recovery purposes. The actual storage system itself may consist of direct access devices or a robotic tape library with removable storage media, depending upon the volume of digital objects and the frequency of imports and exports of digital objects. The preferred “owner” of this storage repository should be the Delaware Public Archives.

3.5.8 Planned Media Renewal

3.5.8.1 Discussion
All known digital storage media today are vulnerable to media degradation that causes the loss of bits comprising digital objects. In addition, all digital storage media eventually will succumb to technology obsolescence. To mitigate the effects of media degradation and technology obsolescence, a planned media renewal program should be in place that transfers digital objects to new storage media at specified intervals, depending upon the project life expectancy of the storage device and/or associated storage media. Planned media renewal also should include a periodic readability check of digital objects to determine if a loss of data is imminent.

3.5.8.2 Recommendation
Cohasset believes that a planned media renewal program is a prudent exercise of risk management in digital preservation and recommends adoption of the planned media renewal strategy that follows.

Implement a planned media renewal program for digital objects based upon projected life expectancies of storage devices and/or storage media and the potential risk of bit streams through media degradation.
3.5.9 Digital Object Authenticity

3.5.9.1 Discussion
Authentic digital objects are trustworthy because over time the integrity of their underlying bit streams has been protected against accidental changes and malicious alterations. The bit streams underlying digital objects can be accidentally changed during media renewal and in fact will be changed during format migration. Even more compelling is the prospect that some individuals are likely to introduce malicious changes in digital objects that are posted on web sites or disseminated on the Internet. When this happens it will be a challenge to demonstrate which version is the authentic one. What are required, therefore, are controls and mechanisms that can validate the authenticity of digital objects.

Two methodologies can support controls and mechanisms to validate the authenticity of digital objects. The first methodology essentially involves validating the integrity of bit streams by computing a hash digest, which is a unique digital fingerprint of a digital object, before transfer or media renewal and generating a new one after completion of the transfer or media renewal and comparing the two. The alteration of a single bit in the transferred or renewed digital object will produce a different hash value.

There have been successful attempts to crack older hash digests such as MD-5 (128 bits) and Standard Hash Algorithm 1 or SHA-1 (160 bits), which suggest that more advanced hash algorithms such as SHA-2 (212 bits) should be used.

Hash digests in and of themselves are of no value in validating the underlying integrity of bit streams of digital objects that are migrated to new formats because the new formats are likely to introduce changes in the bit stream. One way to establish confidence that the content and structure of digital objects has not been altered is to retain a copy of the pre-migrated digital objects for a period of time and conduct visual comparisons with the migrated digital objects.

This leads to the second methodology to validate the authenticity of digital objects over time. It relies upon the collection of preservation metadata that describes what preservation action was taken, why it was taken, who was responsible for its execution, and what the results were. In the case of media renewal the preservation metadata would include the pre and post media renewal hash algorithms. In the case of format migration, the preservation metadata would include identification of the old format, the new format, the tool that was used to execute format migration, and any observed anomalies introduced during the format migration.
Digital preservation specialists agree that preservation metadata should be treated with the level of care as that given to digital objects but there is disagreement about where to retain the metadata. One school of thought calls for encapsulating all preservation data into the digital objects so that it is always a part of them. Another school of thought calls for the retention of digital preservation metadata as an organized body of interrelated in a different location from the digital objects. There are persuasive arguments in support of both approaches but Cohasset believes that the second approach should be implemented.

**3.5.9.2 Recommendation**

Cohasset believes that a sound Delaware Digital Preservation Program requires a robust methodology to validate the authenticity and integrity of digital objects in its custody. Toward this end Cohasset recommends the following digital authenticity/integrity strategy.

**3.5.10 Security**

**3.5.10.1 Discussion**

Protecting the physical security of a digital preservation repository is the first line of defense against individuals who want to introduce malicious changes in digital documents that compromise their authenticity/integrity or to corrupt digital documents so they no longer are accessible or usable. Physical security is ensured by restricting access to the location of the digital preservation repository only to those who have authorized access and can demonstrate this through a coded badge or a password. Physical security also is ensured through a robust fire wall that blocks direct external access to the digital objects in the digital preservation repository. Finally, the physical security of digital objects is assured by requiring rigorous control and documentation of all digital preservation activities such as media renewal and format migration.
that are executed within the repository. This documentation would be in the digital preservation metadata discussed earlier.

3.5.10.2 Recommendation

Protect the physical security of the digital preservation repository through the use of restricting access to authorized individuals, blocking external access to the digital preservation repository, and capturing documentation of all digital preservation actions undertaken.

3.5.11 Access

3.5.11.1 Discussion

A recurring theme in the three deliverables produced under the Phase 1 Digital Planning Study is that the preservation of digital objects and ensuring access to them are mutually interdependent. Digital preservation without ensuring access is a waste of resources. Attempting to provide access to digital objects without their preservation guarantees short term access. Doubtless, there are few who would mount an argument against the interdependency of digital preservation and access to digital objects.

The earlier discussion of the Open Archives Information Reference Model postulated that the Delaware Digital Preservation Program access functionality should not be an integral part of the digital preservation repository. Rather, copies of digital objects that are of current interest to users could be made and transferred to a web site that could be accessed by any standard web browser. In some instances, the format used for public access could be different from the one used in the digital preservation repository so some conversion might be necessary. Copies of digital objects transferred to the web enabled access application would be considered temporary records and could be disposed of when they no longer serve a useful purpose. The official digital objects remain within the Digital Preservation Repository, which should be the primary responsibility of the Delaware Public Archives.

The fundamental issue is how access will be supported in the Delaware Digital Preservation Program and who will have the primary responsibility for managing it. The mandate of the Delaware State Library is to support public access to electronic publications and documents produced by and for state and local government agencies. The Delaware State Library supports
an on-line public access catalog to library holdings in the state and has links to other on-line public access catalogs. In addition, the State Library has a demonstration project underway to test a module of the Hyperion Library Software that supports direct access to a variety of digital objects, including book and non-book images. The long history of the Delaware State Library role of supporting public access to digital sources of library publications and the resources it can leverage to support expanding the scope of material that can be accessed through a standard web browser suggests that it should have the primary responsibility for managing the access repository for digital objects of long term value.

3.5.11.2 Recommendation

The Delaware State Library has the primary responsibility for managing the public access function that makes available to the public copies of digital objects stored in the Digital Preservation Repository.

3.5.12 Digital Preservation Metrics

3.5.12.1 Discussion

A sustained digital preservation program must employ widely publicized empirical performance metrics to identify areas where digital preservation is performing as expected as well as where its performance is not meeting expectations. The use of empirical performance indicators to identify deficiencies and weakness is a key component of an incremental digital preservation improvement program. Deliverable 2 contains a set of digital preservation readiness capability performance metrics and objectives of a Balanced Digital Preservation Scorecard that in fact are empirical digital preservation performance metrics.

3.5.12.2 Recommendation

Cohasset recommends adoption of the digital preservation readiness capability performance metrics and the objectives of the digital preservation balanced scorecard (Deliverable 2) as empirical performance metrics for the Delaware Digital Preservation Program.

Implement the digital preservation readiness capability performance metrics and the objectives of the digital preservation balanced scorecard (Deliverable 2) as empirical performance metrics for the Delaware Digital Preservation Program.
4 DIGITAL PRESERVATION POLICY AND STRATEGY

This section displays the digital preservation policy and strategy that Cohasset recommends to be adopted for the Delaware Digital Preservation Program.

Delaware Digital Preservation Policy

This Digital Preservation Policy formalizes the commitment of the State of Delaware to the long-term preservation of digital objects that protect the legal and financial rights of state and local governments and citizens or contain information about people, places, and events of Delaware history related to continuing self-educational, political, culture, economic, recreation, and intellectual enrichment. Digital objects may be either “born-digital,” that is, they are created, used, and stored in a digital format, or scanned digital images of paper documents. Furthermore, this digital preservation policy declares that digital objects of long-term value are critical assets of the State and their preservation shall be carried out through a sustained digital preservation program supported by stakeholders in the current and future use of accessible, usable, and authentic digital objects.

This Digital Preservation Policy involves four primary interrelated areas of activity and one supporting activity.

1. Digital preservation requires promotion of the value of digital objects to stakeholders who include current and future users of the digital objects

2. Digital preservation requires articulation of the economic value a digital preservation program has for current and future users

3. Digital preservation requires a technical infrastructure capable of supporting digital preservation processes

4. Digital preservation requires implementation of appropriate digital preservation processes that mitigate the effects of technology obsolescence
5. This Digital Preservation Policy (and supporting Digital Preservation Strategy shall be communicated to all senior managers and staff of state and local governments who create, receive, or acquire digital objects in the ordinary course of business, to stakeholders in digital preservation, and to the public at large.

**Digital Preservation Strategy**

**Stakeholder Value.**
An incremental Digital Preservation Stakeholder Value Program will be established to identify stakeholders, engage them on the direct value to them of a robust digital preservation program, and promote public support for the program.

**Economic Value.**
Incremental business case studies and other relevant techniques will be used to identify the economic value of digital preservation for the State of Delaware.

**Infrastructure.**
1. The Delaware Public Archives, the State Library, and the Department of Technology and Information will put in place a formal agreement to review their respective roles and responsibilities with regard to digital preservation and to develop a shared on-going digital preservation agenda that takes into account any clarification about these roles and responsibilities that may be required over time.

2. The Digital Preservation Steering Committee will develop and monitor the implementation of plans and scenarios for the promotion of digital preservation collaboration awareness. Such plans and scenarios may include, for example, adding a digital preservation readiness impact statement check to the business case procurement template for new systems and applications.

3. The need for sufficient technical expertise to support a fully operational digital preservation program will be met by recruiting one or more new recruits with professional training in digital preservation technologies. In addition, short term consulting contracts with information technology consultants to provide specific digital preservation services and contract with third parties to develop customized digital preservation training courses for selected employees can be used.
Processes.
1. The Delaware Digital Preservation Program will use the Open Archives Information Systems Reference Model to frame its overall digital preservation strategy.

2. An incremental migration strategy will be employed to ensure that technology obsolescence does not impede or otherwise prevent the extending the retrievability, usability, and authenticity of digital objects for current and future users.

3. Incrementally adopt technology neutral open standard formats for use in the storage of digital objects, beginning with TIFF, JPEG, and PDF.

4. A Technology Watch Program will be used to monitor trends in technology obsolescence and alert the Delaware Digital Preservation Program when action must be taken to avert technology obsolescence.

5. Receive complete and authentic digital objects from their owners and prepare them for transfer to the storage repository. Included among the preparation techniques to be used are placing a hold on encrypted or password protected digital objects until they are decrypted and passwords removed, checking digital objects for virus infections and quarantining them until fixed, and normalizing digital objects to an appropriate technology neutral open standard format.

6. Utilize a physical digital repository that is hosted on a storage system exclusively dedicated to digital preservation and segregated from an operational environment or external intrusion by a robust firewall. This storage system supports ingesting of digital objects, exporting digital objects to an external access system, periodic media renewal, migration from old formats to new formats, and backup copies for disaster recovery purposes. The actual storage system itself may consist of direct access devices or a robotic tape library with removable storage media, depending upon the volume of digital objects and the frequency of imports and exports of digital objects. The preferred “owner” of this digital repository is the Delaware Public Archives.

7. Implement a planned media renewal program for digital objects based upon projected life expectancies of storage devices and/or storage media and the potential risk of bit streams through media degradation.

8. Implement the use of robust hash digest algorithms to validate the integrity of digital objects through comparison of pre and post digital objects transferred from old media to new media.
Validate the integrity of digital objects that undergo format migration by retaining a copy of the bit streams of digital objects in the old format that can be rendered and compared with the rendered digital objects in the new format. Retain all metadata associated with preservation actions as an organized documentation that current and future users can use to confirm the authenticity and integrity of digital objects in the digital preservation storage repository.

9. Protect the physical security of the digital preservation repository through the use of restricting access to authorized individuals, blocking external access to the digital preservation repository, and capturing documentation of all digital preservation actions undertaken.

10. The Delaware State Library has the primary responsibility for managing the public access function that makes available to the public copies of digital objects stored in the Digital Preservation Repository.

11. Implement the digital preservation readiness capability performance metrics and the objectives of the digital preservation balanced scorecard (Deliverable 2) as empirical performance metrics for the Delaware Digital Preservation Program.