Hawai'i Digital Archives Plan

2012

The Hawai'i Digital Archives Plan analyzes Hawai'i's ability to collect and preserve digital records and presents a working blueprint to establish a digital archives capable of preserving and providing access to historical digital records of Hawai'i government. This initiative was made possible through a grant from the National Historical Publications and Records Commission (NHPRC).

Preserving our Cultural and Intellectual Heritage

Table of Contents

Executive Summary

The Hawai'i State Archives has embarked on a mission to preserve and provide access to historical digital records of Hawai'i's government. While there is a clear direction on what records need to be preserved, there is no mechanism within the state to assist agencies in addressing the difficulty of the long-term preservation of digital records. The State Archives is mandated by law to preserve the permanent records of state government; creating a centralized digital archives for the preservation of digital records is a necessary next step. A digital archives will not only ensure the migration of digital records in a cost effective manner but will provide a single access point to all citizens of the state.

Hawai'i State Archives received a grant from the National Historical Publications and Records Commission (NHPRC) to create a plan for the development of a digital archives. Hawai'i State Archives hired a consultant to assist in developing the plan. For the last two years, Hawai'i State Archives has discussed digital archiving across state government, exploring what it entails, why it is necessary, and how it can be developed and sustained; sent out surveys to departments to get a better picture of the digital archiving needs of Hawai'i's government; held training sessions and focus group session to engage and educate government workers regarding the issues of records management; engaged in conversations with agencies across state government and the City and County of Honolulu.

The major issues that were consistently repeated in the responses to the survey include: the need for more frequent and intensive records management training that includes digital records management and maintenance, the need for guidance from the State Archives on how to manage digital records (in the form of draft policies, procedures and guidelines which agencies can use to develop their versions), and the need for a long term repository that will preserve digital records throughout the required retention period. To fulfill these needs, it is the recommendation of the project consultant that a state digital archives be developed and modeled on international standards for the preservation of digital records. The survey demonstrated a clear need for a solution to the rapidly growing problem concerning the maintenance and preservation of digital records. To assist the digital archives initiative, extensive training and best practice documentation must be developed and distributed statewide in order to educate all state employees on how to fulfill their responsibilities in managing digital public records. To continue with business-as-usual places the State of Hawai'i at a high risk of non-compliance with state and federal records laws, rules and regulations. The foundation of democracy is government accountability to the people; without records there is no accountability; without management, there are no records.

State Archives' Role and Responsibility

The responsibilities of state employees with regards to the management of public records are clearly stated in law and policy. All public employees are required by §94-3 HRS to maintain the records created or received in their usual and ordinary course of business for the length of their approved retention schedule. In Comptroller Circular No. 2001-02, it explicitly states that all public officials are responsible for the protection and accessibility of government records under their purview. The Archives has a statutory responsibility, outlined in HRS chapter 94, to play a leading role in the management of all state government records, regardless of format and a requirement to collect and preserve the historically significant records that government creates. §94-1 HRS directs the State Archives to collect all public archives; arrange, classify, and inventory the same; provide for their safekeeping; and compile and furnish information concerning them.

Records created by government agencies in digital format are now legally accepted as government records. (See, Act 177, SLH 2005). As such, they must be available and usable throughout their required retention periods. Unlike paper records, digital records are dependent on computer hardware and software to be usable. There are no current policies and/or laws that specify digital recordkeeping standards that apply to all Hawai'i government agencies. The Archives collects, stores, and makes available the permanent records of Hawai'i's government and oversees the State's overall records management program. The majority of these records are in paper and microform formats. With the exception of a few CDs and DVDs, the Archives is not yet collecting the permanent digital records of state agencies.

As government records are increasingly generated and stored in computer-based information systems, the state faces the challenge of managing and preserving these digital documents. The Archives must work with agencies as new systems are developed or existing ones are updated. Improper maintenance of these records could result in the permanent loss of historical, legal, and vital information. The plan to create a digital archives to manage government digital records will enable the Archives to fulfill its functional duty to properly administer a records management program that includes records created digitally, and to ensure their preservation for future generations.

Digital Archives Survey

Survey Background

As part of the NHPRC funded Hawai'i Digital Archives Initiative, it was determined that the previous digital records survey conducted in December 2005 would need to be expanded and revised in order to gain a better understanding of the current state of digital records management in the State of Hawai'i. A second, more expansive survey instrument was designed and distributed electronically to targeted public employees. In an attempt to increase participation, the survey was announced at the Digital Archives Project Kickoff Meeting conducted in September 2010 by the consultant on the project.

To gain a comprehensive understanding of the state of digital records issues and awareness within the state, the survey was circulated in October 2010 to a cross-functional group of records creators, records managers, and IT professions from all three branches of government, higher education and local government. Invitation to partake in the survey was distributed via several email lists maintained by Information and Communication Services Division (ICSD). In the email, participants were presented with several options for completing the survey: online through SurveyMonkey, completing a PDF version of the survey and emailing it back to the archives, or printing out the survey and completing it manually. For those who chose not use the online survey instrument, the results were manually entered into the online tool by the archives staff. The survey instrument is included in Appendix A. Of the 27 agencies that were identified in the survey, 23 are represented in the results, an 89% agency representation rate. In total, seventy-seven individuals started the survey, with 66 completing the survey – an 86% completion rate.

Methodology

Survey questions built upon the previous digital records survey, with the new survey expanding on several areas of interest that were uncovered by the previous survey. Specifically, the survey instrument was expanded in order to gain a better understanding of the use of Records Management Applications (RMAs), the breadth and depth of digital imaging within the state, and the management of email. The survey instrument was designed with both structured and semi-structured questions. Structured questions required the participant to select fixed responses to questions posed; semi-structured questions posed a question to the participant and provided a fixed length field for the participants to respond. Questions were grouped into four thematic areas: Digital Records, Scanned Records, Electronic Document Management Systems (EDMS), and General Records Issues

The survey was created and hosted through SurveyMonkey, an online-survey services provider. This vendor was selected due to the easy-to-use interface for design and survey completion, as well as the powerful results analysis and reporting tools. During the design, it was decided to allow the participants to complete the survey anonymously. Anonymity allowed participants to submit answers that could be perceived in an unfavorable light, without the responses being traced back to any one individual. At the end of the survey, all participants were asked if they would like to participate in focus group discussions by providing their contact information.

The survey was limited to a maximum of 38 questions with the intention of increasing rate of completion while minimizing the impact on already busy work schedules. Business logic was built into the form design to designate specific questions as mandatory and others as optional. Additional business logic was built into the flow of the survey form to allow participants to by-pass questions that did not apply. For instance, if a participant did not have a digital imaging program (defined as the conversion of paper records into a digital image), they would automatically skip past the remainder of the questions concerning digital imaging and move directly to the next section. The responses to all questions were captured in a spreadsheet for later analysis. Several options for completing the survey were offered to participants: online through SurveyMonkey, completing the survey in Adobe Acrobat and emailing results to the Hawai'i State Archives, or printing the survey and returning the completed survey through regular mail. Eighty-five percent of the respondents chose to participate in the survey online.

Analysis of Survey

The survey collection phase ended on November 13, 2010, at which point the online survey was no longer accessible and no additional responses were added to the spreadsheet. The results were then tabulated and responses analyzed. Of particular interest were the semi-structured questions that allowed the participants to provide feedback on issues or areas of concern. From the individual responses, a larger thematic picture was formed of the current state of digital records management and preservation in the State of Hawai'i. The individual results from each question, grouped by the four thematic areas, and recommendations from the consultant specifically addressing the findings from each question are included as Appendix B. Below are the broad-based recommendations based on the results of the survey. Note: As is to be expected with any survey, the number of total responses decline as the survey progresses due to a number of participants not completing the survey.

Agencies Represented in the Survey

Accounting and General Services
Agriculture
Attorney General
Budget and Finance
Business, Economic Development and Tourism
City and County of Honolulu
Commerce and Consumer Affairs
Defense
Education
Hawai'i County
Hawaiian Home Lands
Health
Human Resources Development
Human Services
Judiciary
Labor and Industrial Relations
Land and Natural Resources
Legislature

Office of the Governor
Public Safety
Taxation
Transportation
University of Hawai'i

Recommendations Based on the Survey Responses

General Recommendations for Digital Records Survey Questions:

Based on the responses received, with regards to Digital Records the following general recommendations are made. It is recommended that:

- The State Archives should obtain copies of existing digital records master plans from those agencies where they exist, conduct research on what other archives are doing for migration and produce guidelines for the creation, maintenance and preservation of digital records that can be used by all public institutions to create agency specific master plans.
- Due to the wide diversity of media types storing digital records within the state, the State Archives should draft a list of media types that are recommended for the long-term storage of digital records, as well as a list of media types that should be avoided for all but short-term storage or transport.
- Due to the volatile nature of the hardware, software and media upon which digital records are stored, efficient, and effective backups are necessary to ensure adequate protection of the records against loss. Therefore, the State Archives should work closely with ICSD to draft guidelines for digital records backup that includes verification and auditing procedures to ensure that the backups are being carried out as planned.
- As email is increasingly being used as a method of choice for conducting business and must be treated as a record, additional e-records training modules that include a strong emphasis on email should be developed and delivered across the state.
- Lastly, as change is inevitable in the digital world, the State Archives should obtain copies of existing migration plans, conduct research on what other archives are doing for migration, and then produce guidelines for digital migration that public institutions can incorporate into agency specific policies and procedures.

General Recommendations for Digital Imaging Survey Questions:

Based on the responses received, with regards to Digital Imaging the following general recommendations are made. It is recommended that:

- The State Archives develop a set of standards for the digital imaging of paper records. This publication should include minimum resolution, acceptable file formats, quality assurance

procedures, and auditing protocols. It is also suggested that legislation be drafted to require State Archives approval of digital imaging systems prior to the destruction of paper records.

- A cost-benefit analysis be created based on the imaging processes currently in use. The study should focus on the primary areas of benefit listed in the survey: access, storage, and management.
- The State Archives seek out and engage in partnerships with those agencies imaging documents of historical value in order to ensure the maintenance and preservation of important government records.
- Further research be conducted into the possibility of centralizing the large scale imaging operations in order to reduce overall costs.
- A model imaging contract be created to detail the necessary controls required to ensure the accuracy and integrity of the images, as well as the state's ownership of the product, both images and indexes.

General Recommendations for Electronic Document Management System (EDMS) Survey Questions:

Based on the responses received, with regards to EDMSs the following general recommendations are made. It is recommended that:

- Agencies with EDMSs be made aware of the dangers of maintaining records in potentially proprietary systems. Guidelines should be developed that recommend data export protocols that will ensure continued access and management of records residing in EDMSs in the event that the system is no longer supported.
- The State Archives be involved in all future EDMS procurements and implementations to lend subject matter expertise on the issues of the maintenance and preservation of public records of long-term value.
- A 'recommended practices' guide be developed for a minimum core set of metadata that has been determined necessary to ensure the maintenance and preservation of public records of long term value.

General Recommendations for General Records Survey Questions:

Based on the responses received, with regards to general records the following general recommendations are made. It is recommended that:

- The State Archives continue and expand its current records management training to include online training for remote locations, electronic records management training that emphasizes email management, and advanced records management training courses.

- Further research be conducted on the recommended practices and guidelines that have been issued by other public sector digital archives programs and adapt these to State of Hawai'i use. Based on the feedback received in the survey, public employees are in need of guidance on the appropriate methods and practices for managing their digital records according to approved retention schedules.
- Interested stakeholders should continue to be engaged as the project moves on, through regular updates, focus group meetings, and solicitation of beta testers.
- A State Digital Archives be developed to allow for the centralized preservation of digital records of enduring fiscal, historical, or legal value. As part of this project, guidelines for the management and maintenance of digital records should be created and incorporated into training modules.

Phased Implementation Plan

Purpose

In order to build a system that works for Hawai'i State Government, it is recommended that the Digital Archives project pursue a phased implementation plan. Based on extensive evaluation of the digital archives software applications currently available, it will be necessary to modify and grow the chosen archiving platform to fit the unique functional requirements determined to be essential for the proper operation of the Hawai'i Digital Archives. The recommended implementation plan is comprised of three phases: a prototype phase to familiarize the digital archives technical staff and prototype partners with the operation and capabilities of the base system; a pilot phase that improves upon the base system by increasing functionality and demonstrating this modified system to a larger group of partners; and a production phase that rolls out the fully featured digital archiving system to Hawai'i state government. A phased implementation plan will allow the staff of the Digital Archives project to:

- develop a strong familiarity with the individual components of the digital archive;
- incrementally build the system to meet the needs of state government agencies and customers;
- provide key milestones to demonstrate needed functionality and solicit feedback from partner agencies and customers; and,
- use the knowledge and experience gained in previous phases to increase the robustness, features, and capacity of the Digital Archives system.

Phase One: Prototype

The Prototype Phase will be comprised of five primary streams: hiring of project technical staff, a comprehensive review of the prototype agencies' record-keeping systems, collection of sample records from partners and other willing agencies, a code review of the chosen digital archives software package, and initial rollout of the prototype system architecture with demonstration of the core functionality.

It is recommended that upon approval of the Digital Archives project, job postings for technical project staff be advertised. It is anticipated that there will be minimum of a three-month lead-time between initiation of the hiring process and the effective start date of the selected personnel. It is critical to the project timeline that the project technical staff be available at the start of the project. Delays in hiring will directly affect the proposed timeline on a one for one basis – that is, every one-month delay in hiring the technical staff will result in a one-month slip of the proposed timeline.

The prototype agency review will consist of conducting a current records inventory for the agency's paper and digital records. Staff from the State Archives will compare the current inventory to the agency's current retention schedule, update the retention schedule as needed, and note any digital records of permanent value that would be eligible for preservation in the Digital Archives. Identified records series will be the primary targets for the prototype phase of the project. Once identified, Digital Archives project members will work with the prototype agencies to document the location, record profile, and possible

transmission methods for the target records. A Memorandum of Understanding (MOU) will be drafted with each prototype agency detailing the targeted records and expected transmission method.

To supplement the targeted records from partner agencies, samples of known archival digital records from other willing agencies will be collected in order to provide as broad a picture as possible of the breadth and depth of records that could potentially be sent to the Digital Archives. Record samples collected from non-partner agencies will be used solely for the purpose of designing a flexible ingestion mechanism that will be capable of processing a wide variety of records from a broad base of agencies. Sample records will not be permanently accessioned into the Digital Archives; rather, the sample records will be used to design and test the Digital Archives functionality and ability to process a wide variety of records from a diverse cross-section of agencies.

The code review will begin with a top-down analysis of the source code for the digital archiving software selected to provide the base foundation for the Hawai'i Digital Archives project. This analysis will map the individual code functions and classes to the functional requirements detailed in the Digital Archives Plan, while inspecting any code documentation that exists. Existing code that does not meet the required standard, in either programming methodology or inline code documentation, will be noted. Logged deficiencies will be prioritized and addressed in the Pilot Phase.

After the analysis of the code base and identification of prototype agency records have been completed, an action plan for addressing any missing core functional requirements detailed in the Digital Archives Plan will be developed. The missing core functionality will be prioritized and assigned to the technical staff for development and integration of the needed modules. The initial prototype system hardware will be installed, tested, and documented. As modules are developed, the collected record samples will be used to test the prototype Digital Archives system, demonstrate its functionality to partner agencies, and solicit feedback from the partners and State Archives staff. Feedback received will be evaluated for implementation and prioritized based on need and feasibility of the responses.

Estimated time frame: 12 Months

Staff resources required:

Full time: Project Consultant, Project Systems Developer, Project Digital Records Acquisition Developer

Part-time: Project Manager (entire phase), Records Analyst (2nd quarter of phase)

Conditions for Successful Completion of Phase:

- Hiring of technical staff
- Up-to-date records inventories for prototype agencies
- Draft MOUs completed with prototype agencies

- Broad representation of sample archival digital records from state agencies
- Top-down code review of base digital archives software
- List code deficiencies
- Prioritization of required functionality remaining to be implemented
- Rollout and documentation of prototype system hardware
- Coding and implementation of core functional requirements
- Demonstration of core functionality to partner agencies

Phase Two: Pilot Phase

The Pilot Phase will consist of five streams: full documentation of Digital Archives' code, rollout of pilot hardware, onboarding of pilot partner agencies, implementation of functional requirements listed as 'necessary' in Archives Plan, a security analysis of the Digital Archives' system, and a comparison of the developed system to the Trustworthy Repository: Audit and Checklist (TRAC) created by the Research Library Group (RLG).

The Pilot Phase will begin by addressing those code deficiencies in the base digital archives software package identified in the Prototype Phase. Under the direct supervision of the Project Technical Staff, interns from a local university will rewrite and document any code determined to be deficient by the coding standards established by the Project Team. All third-party tools utilized by the system will be documented to the fullest extent possible. The interns will be tasked with providing a documentation level sufficient for other programmers who are not familiar with the system or digital archiving to be able to support the program without the aid of the Project Technical Staff. This level of extensive documentation is necessary to provide for the long-term, on-going support of the Digital Archives after the completion of the project.

Additional hardware will be deployed and existing hardware will be re-tasked to accommodate the expansion of both the functionality and capacity of the Digital Archives. Storage subsystems will be added to the existing storage pool to accommodate the incoming digital records that will be transferred to the Digital Archives by the pilot agencies in this phase. A three-tier architecture will be deployed to allow for the extensive testing of workflows for transfer, ingestion, and retrieval of records, along with the establishment of the necessary security protocols.

While the interns are documenting code and the Project Technical Staff is rolling out, testing and documenting the hardware, staff from the State Archives will be meeting with Pilot Phase partners to perform records inventories of the agencies' paper and digital records. These inventories will be compared to each agency's current retention schedule. Updates to the retention schedule will be implemented as needed. Those record series identified as archival will be targeted for transfer to the Digital Archives by drafting MOUs with the pilot agencies, noting the location and anticipated transfer protocols for each series.

Once the hardware is implemented, tested, and documented, a review of the current Functional Requirements detailed in the Digital Archives Plan will be conducted by the State Archives staff and partners. Based on feedback received, a priority list for implementing missed or additional 'Core' functionality -- as well as 'Necessary' level functionality -- that can be addressed in this phase will be developed; while those 'Core' and 'Necessary' Functional Requirements that cannot be addressed in this phase will be prioritized for later phases. Developed modules will be extensively tested for robustness, throughput, and efficiency.

This phase will conclude with top-down analysis of both hardware and software as deployed. Once the final Pilot Phase code has been deployed, a comprehensive security audit will be conducted. Any deficiencies will be noted and prioritized for correction in the next phase. Upon completion of the security audit, a TRAC audit will be conducted. Again, any deficiencies will be noted and prioritized for correction in the final phase. The TRAC audit will be used as a litmus test for the trustworthiness of the records contained within the system. Prior to going live, it is strongly recommended that a complete TRAC audit be fully documented and the audit worksheets be completed in order to demonstrate the trustworthiness of the Digital Archives' systems and methodologies.

Estimated time frame: 18 Months

Staff resources required:

Full time: Project Consultant, Project Systems Developer, Project Digital Records Acquisition Developer

Part-time: Project Manager (entire phase), Records Analyst (1st quarter of phase), 2-4 interns (3-6 months each)

Conditions for Successful Completion of Phase:

- Entire code base meets specified coding and documentation standards
- Rollout of three-tier architecture
- Drafting MOUs completed for partner agencies
- Successful ingestion of targeted record series from pilot partner agencies
- Full implementation of all 'Core' functional requirements
- Implementation of 'Necessary' functional requirements scheduled for Pilot Phase
- Transfer, ingestion and retrieval functionality stressed tested and documented
- Top-down comprehensive security audit preformed and deficiencies documented
- Comparison of system to the TRAC requirements, with deficiencies noted

Phase Three: Production Phase

The final phase of the project will consist of seven streams: final development of any functionality determined to be necessary prior to launch of the production system, rollout of production hardware,

stress test of the production system, address any final functionality or revisions necessary, complete documentation required for TRAC audit, knowledge transfer to State Archives staff and public unveiling of the Digital Archives.

Any remaining functional requirements that have not been implemented that are been determined to be necessary for the public launch will be implemented by the project technical staff and project interns. Feedback will be solicited from partners and select customers throughout the final development cycle to ensure a product that is intuitive and useful to the targeted users. Remaining functional requirements not implemented will be prioritized for development after the system is in production.

Prototype hardware will be re-deployed with additional hardware and storage sub-systems into the final production environment. A second security analysis will be conducted to ensure that the hardware is appropriately protected. The code base will be moved over to the production system and tested for correct operation. System redundancy and failover will be tested and any deficiencies will be immediately addressed. System backup policies will be tested with entire system restores performed. The production system will be fully documented and tested for correct configuration. Finally, the disaster recovery manual detailed in the Digital Archives Plan will be created.

The system will be stress tested to determine the upper limits of the systems throughput, in terms of the maximum number of records that can be transmitted, ingested, and retrieved at any given time. Records from the partner agencies, as well as the samples collected from other agencies, will be replicated and used for this test. The impact of large-scale movement of records through the system will be analyzed, with particular attention paid to any affect such movement has on the security of the system. The results will be studied to determine if the current deployment is sufficient to handle the expected volume of records while simultaneously providing the needed level of response to customers. Any shortcomings will be noted, prioritized, and addressed prior to launch.

Outstanding issues will be addressed based on the reviews and analyses performed earlier in this phase. Upon successful resolution of outstanding issues, a final code review will be performed to ensure that the entire existing code base strictly adheres to the coding and documentation standards established for the project. The State Archivist will be presented with the results of the stress test, security review, code review, disaster recovery policy test, and partner and customer feedback in order to make the 'Go-No Go' call to proceed with the public launch of the Digital Archives. Any concerns raised by the State Archivist will be addressed and re-presented for final approval.

Once the 'Go Live' order has been issued by the State Archivist, a complete TRAC audit will be conducted with full documentation collected and codified into the audit. Additionally, the code base will also be duplicated and stored in the offsite security location along with the initial 'Red Box' disaster recovery manual and associated documentation. It is intended that the system will be documented to a sufficient level of detail such that any Information Technology professional can replicate the system based solely on the documentation contained in the 'Red Box' offsite disaster response kit.

Using the documentation created, permanent State Archives technical staff will be thoroughly briefed on the inner workings for both the hardware and software. The design methodologies and philosophies used by the project technical staff will be detailed so that the ongoing technical staff will be well versed in the inter-workings of the various components of the system. As part of the handoff, an extended project plan for the ongoing maintenance of the Digital Archives, as well as the continued growth and expansion of the system, will be developed in conjunction with the State Archives staff. This project plan will include recommended feature sets, a roadmap for hardware expansion, a list of the next round of partner agencies to onboard, and upcoming projects or technologies that may impact the Digital Archives. It is intended that by the end of this phase, the permanent State Archives technical staff will have the knowledge and system experience necessary to seamlessly transition into providing full support of the system.

Finally, upon approval by the State Archivist, the Digital Archives will be unveiled to the public. Project staff will be on hand to assist the State Archives staff with answering questions, addressing the media, providing demonstrations of system functionality, and helping customers. As part of the rollout, user manuals and 'quick start guides' will be developed to assist partner agencies and researchers. This phase will close with a final de-brief of the project and technical staff after the public unveiling. From this point forward, the State Archives technical staff will assume the sole responsibility of maintaining and developing the Digital Archives' system.

Estimated time frame: 6 Months

Staff resources required:

Full time: Project Consultant, Project Systems Developer, Project Digital Records Acquisition Developer

Part-time: Project Manager (entire phase), 2-4 interns (entire phase), Permanent Technical Staff (Second half of phase)

Conditions for Successful Completion of Phase:

- Address all shortcomings identified in security audit
- Implementation of all identified 'Core' and 'Necessary' Functional Requirements
- Implementation of 'Highly Desirable' Functional Requirements planned for roll-out
- Successful configuration of production hardware
- Development and documentation of system backup policies
- Identification and implementation of offsite backup location
- Creation of the disaster recovery 'red box'
- Transition from temporary to permanent State Archives technical staff, if budgeted and approved
- Development of extended project plan to address additional functionality, as well as agency and customer feedback
- Knowledge transfer and handoff of project to State Archives Staff
- Successful rollout of Digital Archives to the public

Digital Archives Initiative Partners

The State Archives has engaged in discussions with various stakeholders from government agencies, higher education, and the public in order to solicit their support and feedback on the digital archives initiative. Because of this outreach, several key partners have stepped forward and expressed a willingness to join the State Archives on the Digital Archives initiative. These partners have been classified into two categories: Contributing Partners and Technology Resource Partners. Contributors are those agencies that will serve as test-beds for the transfer of records from their agencies to the digital archives and will be actively involved throughout the project design and development. Depending on the sophistication of the agencies systems and the time commitment involved, transmission of target record types either will be automated on a predetermined schedule or manually transferred using removable hard drives and/or DVDs. Technology and Resource Partners will be involved in the exchange of information and expertise by being actively involved in the requirements gathering and system design phase by providing input and feedback on the function and operation of the digital archives.

Contributing Partners

The Legislature

The legislative power of the State, under the State Constitution, Article III, Section 1, is vested in a Legislature that consists of two houses, a Senate and a House of Representatives. The Legislature's power extends to all rightful subjects of legislation not inconsistent with the State Constitution or the Constitution of the United States.

Annually, on the third Wednesday in January, the Hawai'i State Legislature convenes in Regular Session to consider all proper subjects for legislation. In addition to its law-making functions, the Legislature performs functions which include fact-finding and similar investigations, receiving and considering requests or petitions from groups and individuals, confirming certain officers appointed by the Governor (a function that is the prerogative of the Senate under Article V, Section 6, of the State Constitution), participating in amending the Constitution, and exercising quasi-judicial authority to punish in cases of certain offenses against the Legislature or its members.

The Senate consists of 25 members elected from 25 senatorial districts for staggered four-year terms. The House of Representatives consist of 51 members selected from 51 representative districts for two-year terms. Each house adopts its own rules, establishes standing committees, maintains its own records, and elects its own officers. The presiding officer of the Senate is the President of the Senate. In the House, the presiding officer is the Speaker of the House of Representatives.

Governor's office

The Governor is the Chief Executive of the State of Hawai'i and is responsible for the faithful execution of the laws of the State and for the effective performance of the executive branch of state government. The Governor establishes the broad goals of the State and determines the priorities for achieving the goals. The Governor provides community, executive, and political leadership for the people of the State, plans for the development of the State and its resources, and provides for the general well-being of the people. The Governor directs the development of the executive budget and the generation of recommendations for revenue producing measures and programs that are submitted to the Legislature for each fiscal biennium and administers the execution of the legislatively approved budget. The Governor's authority includes the powers to call the Legislature into special session, to veto legislation passed by the Legislature, to grant reprieves and pardons, and to arrange interstate compacts. The Governor is also the Commander in Chief of the armed forces of the State. Unless otherwise provided by law or the Constitution, the Governor nominates and, with the advice and consent of the Senate, appoints members of boards and commissions and members of the Cabinet and fills vacancies in the House of Representatives and the Senate. From a list of nominees submitted to the Governor by the Judicial Selection Commission, the Governor fills vacancies in the Supreme, Intermediate Appellate, and Circuit Courts. In addition, the Governor, as required by the State Constitution, Article V, section 5, appoints an Administrative Director, who functions as the Governor's chief of staff and assists the supervision of the executive departments and major state initiatives.

The Office of the Governor is comprised of five functional areas: (1) Executive, (2) Administration and Operations, (3) Communications, (4) Policy, and (5) Collective Bargaining.

Judiciary

The Judicial branch of the State Government of Hawai'i is separate and distinct from, but coequal with, the Executive and Legislative branches. The State Constitution, Article VI, section 1, provides for the Hawai'i Supreme Court, the Intermediate Court of Appeals, Circuit Courts, District Courts, and such other courts as may be established by the Legislature. The Legislature utilized this constitutional authority to create the Land Court and Tax Appeal Court, established under sections 501-1 and 232-11, Hawai'i Revised Statutes, respectively. The Family Courts, pursuant to section 571-1, Hawai'i Revised Statutes, were subsequently added as divisions of the Circuit and District Courts.

Bureau of Conveyances, Department of Land and Natural Resources

The Bureau of Conveyances, as mandated under Chapters 501 and 502, Hawai'i Revised Statutes, maintains accurate records of land title registration and other comprehensive records of documents related to land titles. The Bureau also makes copies of recorded land documents available to various agencies and individuals. Certificates of title on registered land issued by the Bureau are guaranteed by the State against the loss, damage, or deprivation of land, estate, or interest in the land, arising through the fault of the

assistant registrars in the performance of their duties. Hawai'i is the only state in the union with a single statewide recording office.

Hawai'i State Public Library System

The management of the Public Library System is under the Board of Education, pursuant to Section 312-1, Hawai'i Revised Statutes. The State Librarian, under Section 312-2.1, Hawai'i Revised Statutes, is appointed by the Board of Education and is responsible for all public and public-and-school libraries in the State. The System operates 51 public libraries, including 24 on Oahu, 13 on Hawai'i, six on Maui, six on Kauai, and one each on Molokai and Lanai. Following major reorganization in 1995 and 1996, the System includes the Hawai'i State Library and the Library for the Blind and Physically Handicapped; a Public Libraries Branch, which replaced library district offices in overseeing the forty-nine branch libraries; the Administrative Services Branch, which handles administrative matters and Capital Improvement Projects; the Human Resources Branch, which provides personnel services; the Electronic Services Support Section, which manages automation services and the System's network; the Technical Services Support Section, which purchases and catalogs new acquisitions and processes many new titles; the Library Development Services Section, which provides grant writing, program coordination, and promotional support for system libraries; and Logistical Support Services, which provides centralized mail, deliveries, and supplies.

Technology and Resource Partners

Information and Communication Services Division

The Information and Communication Services Division plans, coordinates, and administers the statewide information processing and telecommunications services and programs, and operates an overall program for improving government efficiency and effectiveness through telecommunications and information processing technologies. It provides statewide voice, data, radio, and video communications and networking. The Division operates a central information-processing center, electronic mail, Web page hosting for state departments, Internet support services, and video conferencing system. It is also responsible for developing and maintaining information systems, and evaluates and implements image processing and electronic commerce technologies.

Office of Information Management and Technology

The Office of Information Management and Technology, under the direction of a full-time chief information officer and an Information Technology Steering Committee, is established within the Department of Accounting and General Services to organize, manage, and oversee statewide information technology governance; develop, implement, and manage the state information technology strategic plans; develop and implement statewide technology standards; report annually to the governor and legislature on the status and implementation of the state information technology strategic plans; and perform other necessary or desirable functions to facilitate its mandates, including the supervision and oversight of the Information and Communication Services Division of the Department of Accounting and General Services.

University of Hawai'i

The University of Hawai'i is a land, space, and sea grant institution and the only public institution of higher education in the State of Hawai'i. It began in 1907 as a land-grant college of agriculture and mechanic arts called the College of Hawai'i. In 1919, it was established as the University of Hawai'i (Act 203, Session Laws of Hawai'i 1919; sections 26-11 and 304-1, Hawai'i Revised Statutes). The University of Hawai'i comprises one comprehensive university campus, two baccalaureate-granting institutions, and a system of seven community colleges. The fundamental mission of the University is to provide all qualified people in Hawai'i with an equal opportunity for quality college and university education and training; create knowledge and gain insights through research and scholarship; provide public service through the dissemination of current and new ideas and techniques; preserve and contribute to the artistic and cultural heritage of the community; and respond to the changing needs of the people of Hawai'i. The State Constitution, Article X, sections 5 and 6 establish the University of Hawai'i as a body corporate, governed by an appointed Board of Regents and administered by the board-appointed President of the University. The standard of instruction is equal to that given and required in similar universities in the United States and, upon the successful completion of the prescribed courses, the Board confers a corresponding degree upon all students who are entitled thereto.

Office of Hawaiian Affairs

The Office of Hawaiian Affairs (OHA) was established in 1978 by amendment to the State Constitution, Article XII, Section 5, and Chapter 10, Hawai'i Revised Statutes. Under the direction of nine trustees elected statewide, OHA has functioned operationally as both a government agency with a strong degree of autonomy and as a trust. Its purpose is to provide the opportunity for a better life and future for all Hawaiians.

The mix of trust funds, earmarked for Hawaiians of at least fifty per cent blood quantum, and general funds provided annually by the Legislature, allows OHA to serve Hawaiians of any blood quantum. In addition to running its own programs, OHA provides major funding to many community-based programs and coordinates joint activities with participating organizations. It advocates for Hawaiians in the State Legislature, state and federal courts, the United States Congress and the local media, as well as by supporting community initiatives and interests.

Functional Requirements

Over the past two years, the staff from the State Archives has held a series of meetings with key stakeholders, focus group sessions, and public presentations with the goal of educating government employees and public citizens on the importance of preserving records of cultural and intellectual value. During each gathering, the State Archives solicited feedback from attendees on the features and functionality that they would like to see in the proposed Digital Archives. Along with these suggestions, the project consultant reviewed several international standards, leveraged the experiences of other public digital archives, and performed extensive scholarly literature reviews to compile a list of functional requirements to guide the development of the Hawaii Digital Archives. With over 120 functional requirements identified, to make the list more manageable, the requirements are grouped thematically as follows:

- Access and Usage Restrictions -- The system must control who has access to what aspects of the repository as directed by the system administrator;
- Architecture Design -- The design of the system, both physical and logical, must support the long-term access to trustworthy digital records through time and space;
- Digital Objects and Rendering -- As all digital records are comprised of one or many digital objects, it is essential that all of the digital objects necessary to render a digital record are maintained and available for the life of the record;
- Ingestion and Normalization -- In order to manage the contents of the repository over the long-term, it is essential that records are complete when transferred and are stored in a standardized way;
- Integrity of the Repository -- The maintenance of trustworthy records requires that the integrity of the records be established at point of transfer and maintained throughout the life of the records;
- Interface Design and Searching -- The system must provide a web interface that allows researchers to search, select and view non-restricted records contained in the repository;
- Management Tools -- Tools must be provided to the repository archivist that will allow for effective management of the repository;
- Metadata Creation and Handling -- Long-term preservation of trustworthy records requires both descriptive and preservation metadata to be indelibly linked to the record;
- System Security -- Records accepted into the custody of the repository must be protected from unauthorized alteration, addition or deletion; and,
- Purging Records from the System -- Rules, regulations, policies, procedures and legislation change over time, technological and human errors occur, and the need to keep records of historical, legal, and fiscal value is periodically re-appraised necessitating the ability to remove records from the system.

Each functional requirement has been reviewed by the Digital Archives project staff and prioritized according to need, development effort, overall impact, and customer experience as follows:

- **Core Requirement:** A requirement that is essential to the proper collection, processing, maintenance and preservation of trustworthy records based on archives staffing and infrastructure.

- **Necessary Requirement:** A requirement that, while not essential, is important to the proper operation of a trustworthy repository. Omission of necessary requirements will result in significant additional work on the part of the staff to demonstrate the integrity of the digital archive.
- **Highly Desirable Requirement:** A requirement that adds additional functionality to assist either the archives staff, contributing agency or researcher.

Based on the breadth and depth of the functional requirements currently identified, the time and budget required to develop a digital archives from scratch that complies with the list of functional requirements was determined to be unrealistic. As an alternative, the project consultant compared the core functional requirements to several viable digital archiving software packages that are in wide-spread, current use. The project consultant evaluated each digital archiving software package on a number of core functional requirements based on the identified needs of Hawaii government agencies and the public; the programming language and design architecture employed, and whether the source code for the application was available to the State for further modification. The results of the scoring rubric used in the analysis are included in the next section, *Comparison of Digital Archiving Software*.

While extensive, the proposed functional requirements (included as Appendix C) are not comprehensive. As the project continues to progress, key stakeholders become more actively involved in the project, further focus groups are convened, and feedback from public consultations are incorporated, the list of functional requirements will be expanded and refined. The proposed phased implementation plan will use the documented functional requirements in order to develop the project development plan, allocate resources based on the priority of the functional requirement, and validate software rollouts against the current functional requirements set.

Comparison of Digital Archiving Software

Digital Archives Scoring Matrix

Functionality Scoring

	WaDigitalArchives	Alfresco	Archivmatica	DSpace	OCLC Digital Archives
Code is Open Source	Yes[1]	Yes	Yes	Yes	No
Code Base	.Net	Java	Java	Java	???
Operating System Platform	Msft	Linux	Linux	Linux	Web
Virtual Machine Capable	No	No	Yes	Yes	No
Onsite or Vendor Hosted	Onsite	Onsite	Onsite	Onsite	Vendor
Number of Core functional requirements supported	25	9	22	19	9
Level of Support available[2]	Low	Poor[3]	Med[4]	High	High[5]
Install base[6]	<10	Low	<20	>1000	High

[1] Code is kept as 'shared source', limited to a sub-set of participants and not openly available

[2] From worst to best Poor, Low, Med, High

[3] Open Source Version has poor support, higher level of support available with commercial version.

[4] High level of support available at a premium

[5] All support is part of contract, free support not available.

[6] Listed as number of production installations/customers of product

About the Software Compared

A synopsis of the five viable digital archives software platforms has been included as Appendix D.

Proposed Architecture

Overview

The Hawai'i State Archives has begun planning a digital archives for preservation and access of the important digital records of the state. The proposed architecture described within this document is intended to be a representational architecture that will provide sufficient resources to meet the needs of the dual function of both a flexible methodology for taking records from agencies and the ability to provide powerful search tools for researchers. Once a digital archiving system is selected, the proposed architecture will be updated to reflect the changes required by the implementation. Regardless of the system selected, securing and maintaining the trustworthiness of the records accepted into custody of the digital archives will be of primary concern. This is reflected by the selection of three-tier architecture, comprised of a web tier, an application tier, and a data tier.

The web tier will provide access to the digital archives to the public through the internet, allowing access 24 hours a day, seven days a week, 365 days a year– with the exception of planned maintenance cycles scheduled for low utilization periods. Researchers, historians, students, schoolteachers, genealogists, and anyone interested in learning more about the history and current function of the State of Hawai'i will be able to access these records from home, local library or any internet café. Government agencies will be able to send their archival digital records to the State Archives for permanent preservation and access. Protocols will be in place to ensure that those records which are restricted by law from public disclosure due to the sensitive nature of their contents (for example: personally identifiable information such as social security numbers) will be restricted from public view for as long as required by law.

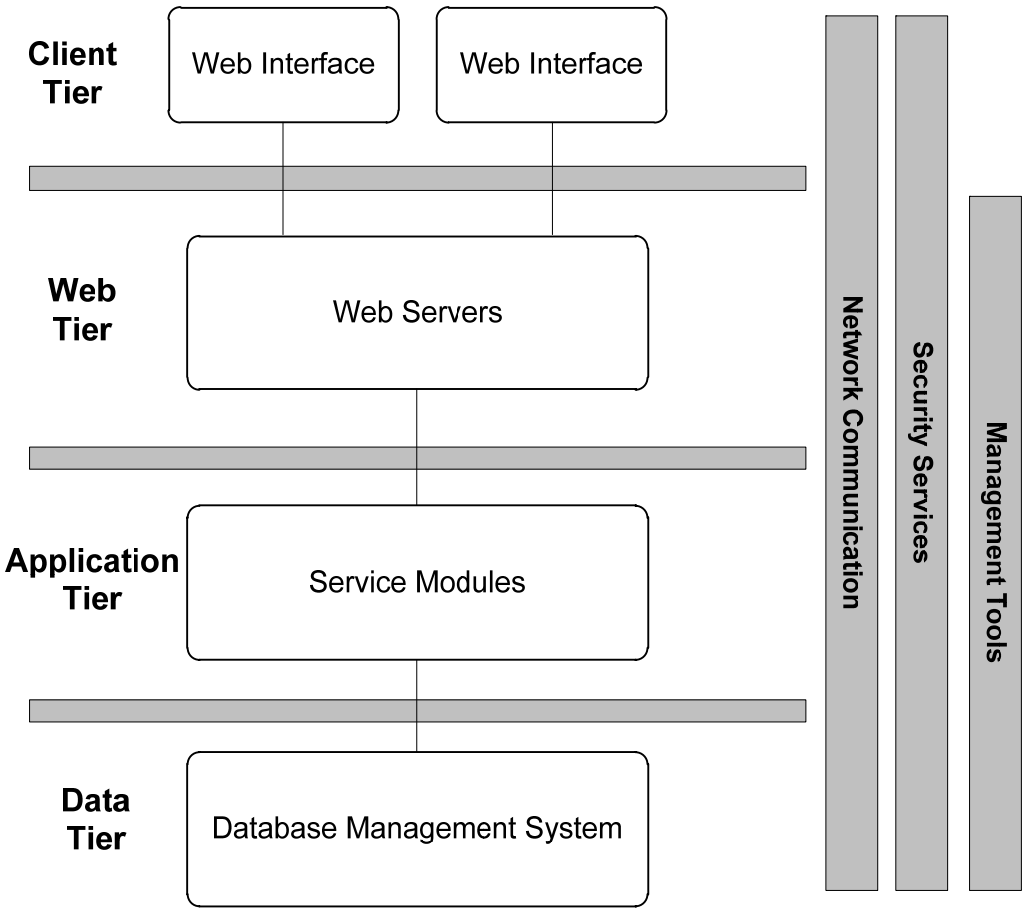
The application tier provides the core of the functionality of the Digital Archives, with a suite of small programs that provide the necessary ingestion, maintenance, retrieval, and migration functions. Incoming records will be processed through a segregated quarantine zone for virus scanning and then verified for accuracy and completeness. Once the ingestion routines begin, data normalization will allow the incoming records to be placed within the database for later searching and retrieval by researchers, while file format conversion will ensure that the records sent today can be viewed by the researchers of tomorrow.

The data tier will store the metadata that describes the records and the digital objects that make up the record. Management of the metadata and the digital objects will be handled by a database management system. Requests for content will be passed by the application tier to the database, which will search the repository and return any results back to the application tier. Any changes to the database will be noted in the transaction log to protect against unexpected interruptions in service. The records themselves will reside on Serial-Attached SCSI (SAS) disk drives contained in an external drive enclosure that provides a high level of data throughput (6 GB/s) and availability with multiple connections, drive redundancy, and dual power supplies. As data storage requirements increase, additional external storage units can be added to the system and populated with hard drives as the need arises. High-end servers, configured with quad-core processors and the ability to add additional memory as needed allows for incremental system growth on a more cost-efficient basis than front-loaded hardware purchases. A tape backup system allows for efficient,

automated backups of the database, file storage, logs and full-text searching catalogs, while also providing the ability to export the contents of the digital archives in an xml format following international standards.

Logical Architecture

The following diagrams depict the conceptual deployment architecture for the Hawai'i Digital Archives. The architecture deployment phases have been designed to be iterative; each phase builds upon the previous, adding to or redeploying resources from previous phases in order to be as cost effective as possible. With the proposed three-tier architecture, each tier is optimized for its specific function; web tier servers are optimized to handle web traffic, which typically involves a higher level of network traffic and disk access but less CPU cycles, the application tier is optimized for processing capabilities to handle the myriad of tasks required to ingest and retrieve records from the repository, and the database tier is optimized for maintaining the database in optimal condition, while inserting and retrieving the actual data from the storage devices. Each tier is designed to be able to expand on demand as the utilization of the existing servers in that tier becomes a bottleneck to the efficient operation of the environment. The production deployment plan is to provide scalability in addition to a high level of availability through the use of load balancing and server clustering.



It is proposed that the system be built in three phases:

- Prototype – Provide sufficient capacity to accept into custody a cross-section of the depth and breadth of records that will be accessioned into the system by collecting samples from key agencies.

- Pilot – Deploy the minimum server set required for a three-tier architecture to allow for extensive testing of the workflows used in the transfer and ingestion of the records, along with the establishment of the required security protocols.

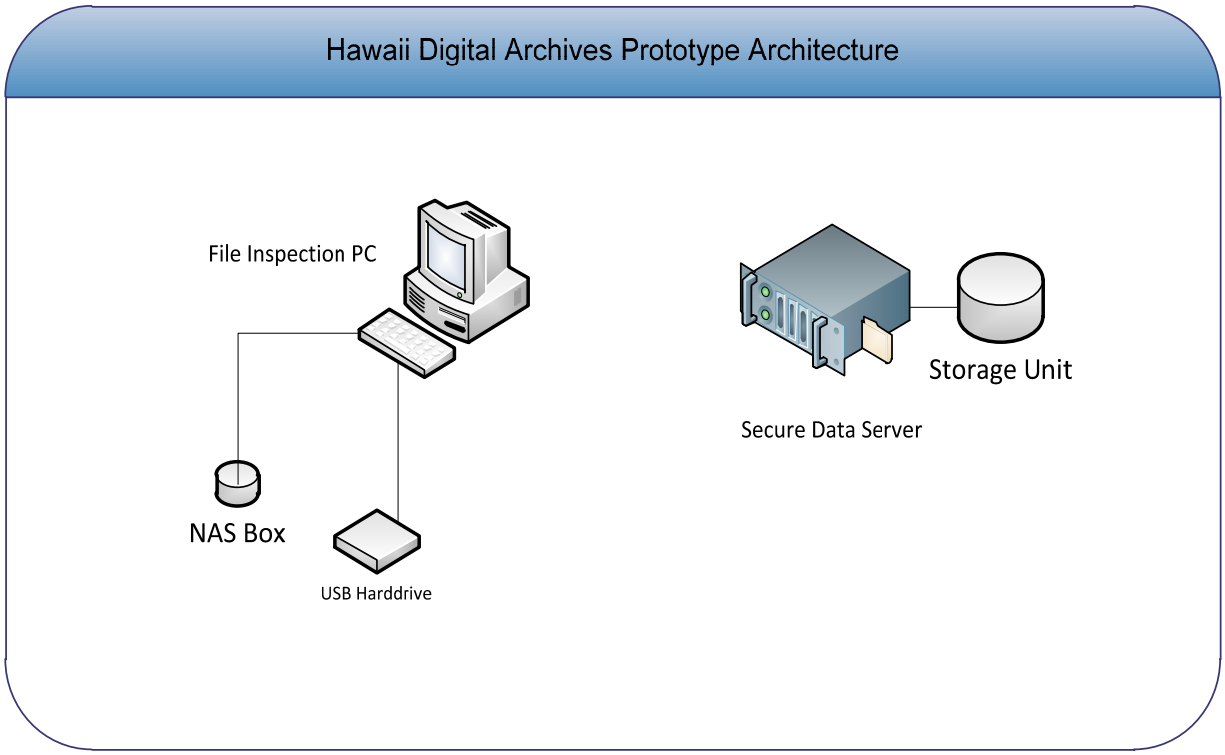
- Production – Add sufficient resources to provide system redundancy and additional throughput to satisfy the performance requirements for producer transfer of records as well as researcher use of the system.

Physical Architecture

Prototype Environment

Assumptions

- Will not utilize three-tier architecture.
- Will not have system redundancy.
- Is designed only for the accumulation and inspection of representative data types for further development of functional requirements and metadata schema



Representative Configuration of Hardware in Prototype Environment Based on HP models

Servers:

Model	CPU	Memory	Hard Drives	Function	Layer
DL160	1	4GB	(4) 250GB	Secure data storage of data samples	N/A
Desktop	1	2 GB	500GB	Virus scan incoming data	N/A

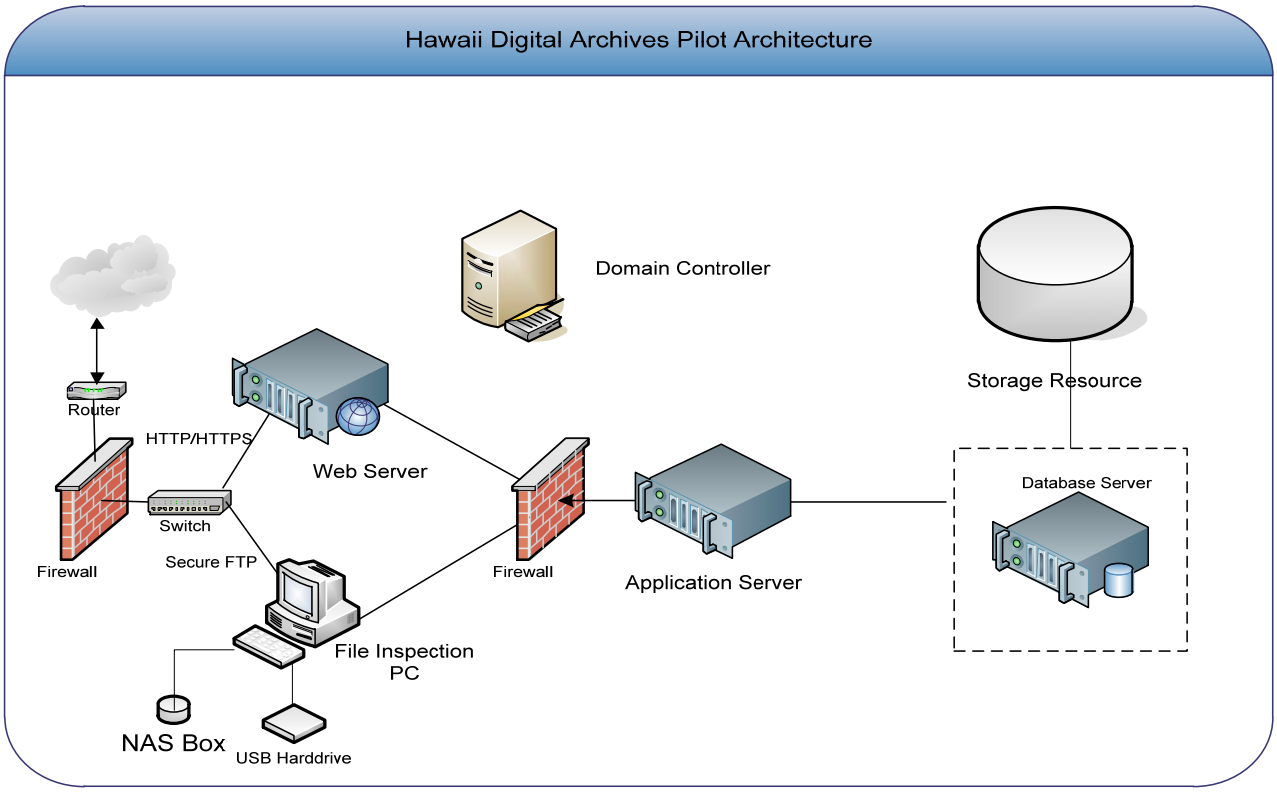
Storage:

Model	Hard Drives	Function
External USB	(1) 2TB SATA	Transportation of data from producer's environment
Network Attached Storage	(1) 2TB SATA	Transportation of data from producer's environment
D2600 Storage Array	(12) 2TB SAS	Data storage for repository

Pilot Environment

Assumptions

- Implements a three-tier architecture.
- Does not have system redundancy.
- Existing firewall infrastructure is properly secured.



Representative Configuration of Hardware in Pilot Environment Based on HP models

Servers:

Model	CPU	Memory	Hard Drives	Function	Layer
DL160	1	4GB	(4) 250GB	Web Server	Web
DL160	1	4GB	(4) 250GB	Applications Server	Application

Model	CPU	Memory	Hard Drives	Function	Layer
DL160	1	4GB	(4) 250GB	Database Server	Data
505B	1	2 GB	500GB	Virus scan incoming data	Application
505B	1	2 GB	250GB	Domain Controller	N/A

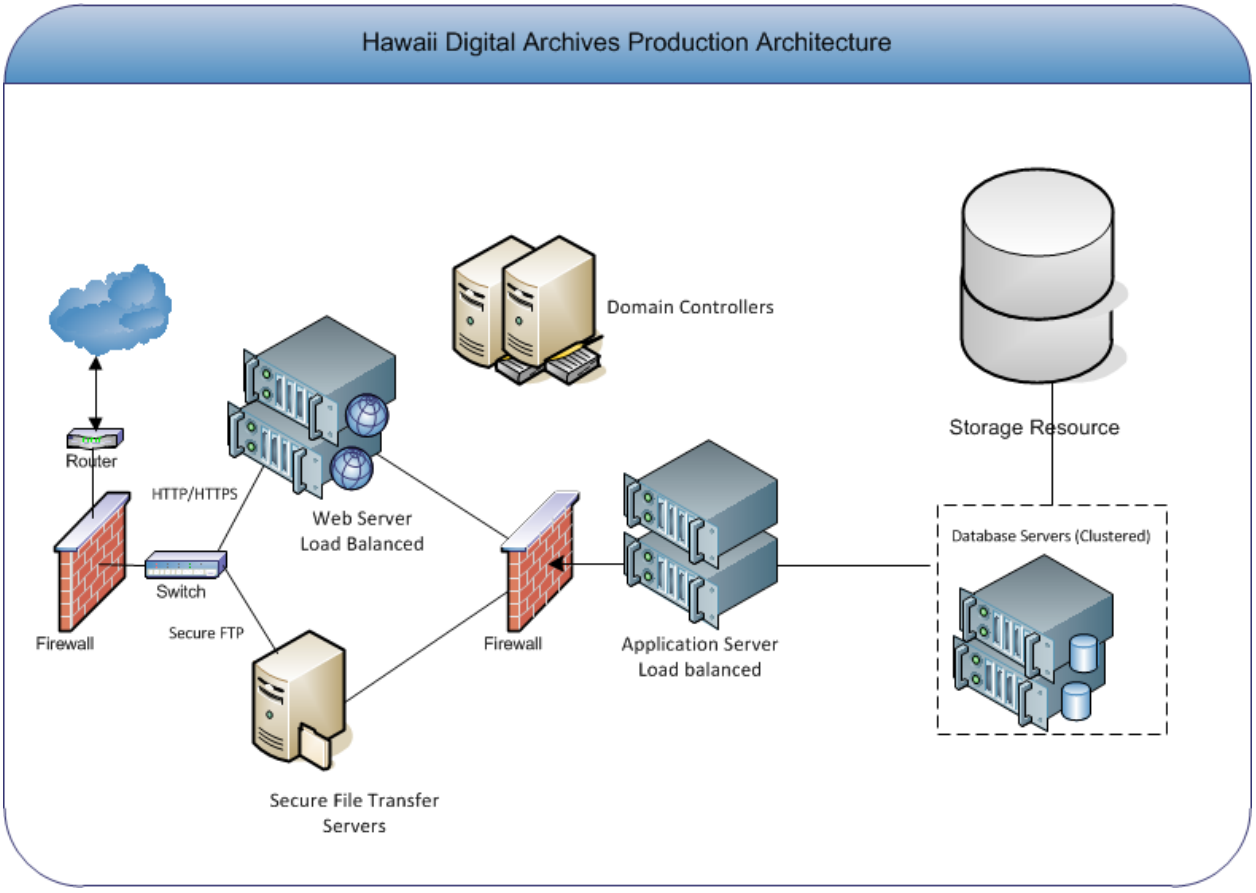
Storage:

Model	Hard Drives	Function
External USB	(1) 2TB SATA	Transportation of data from producer's environment
Network Attached Storage	(1) 2TB SATA	Transportation of data from producer's environment
D2600 Storage Array	(12) 2TB SAS	Data storage for repository

Production Environment

Assumptions

- Sufficient internet access is available.
- Redundant power is onsite.
- HTTPS/SSL is enabled
- Full security with intrusion preservation is in operation



Representative Configuration of Hardware in Production Environment Based on HP models

Servers:

Model	CPU	Memory	Hard Drives	Function	Layer
DL160	1	4GB	(2) 120GB	Web Server	Web
DL160	1	4GB	(2) 120GB	Web Server	Web
DL160	2	4GB	(2) 120GB	Applications Server	Application
DL160	2	4GB	(2) 120GB	Applications Server	Application
DL160	2	8GB	(2) 120GB	Database Server	Data
DL160	2	8GB	(2) 120GB	Database Server	Data
ML110	1	2 GB	500GB	Virus Scan Incoming Data	Application
ML110	1	2 GB	250GB	Domain Controller 1	Domain
ML110	1	2 GB	250GB	Domain Controller 2	Domain

Storage:

Model	Hard Drives	Function
External USB	(2) 2TB SATA	Transportation of data from producer's environment
Network Attached Storage	(1) 3TB SATA	Transportation of data from producer's environment
D2600 Storage Array	2 Units each of (12) 2TB SAS	Data storage for repository

*Note: Proposed architecture is for cost planning purposes only. Final configuration is highly dependent on the chosen digital archiving environment and may vary greatly from what is shown above.

Legislative Review and Recommendations

For a digital archives to be successful in the state, it must be backed by strong laws and policy that address four core requirements: establishing the State Archives' authority to create and enforce records management rules, including the final disposition of records; a clear and concise definition of 'record' that includes those that are created in, or converted to, a digital format; clearly worded rules and regulations that require state employees to manage digital records; and strong laws governing the public's right to access government records. As part of the due diligence for the Hawai'i State Digital Archives project, the project consultant reviewed the current Hawai'i Statutes governing the creation, management, and preservation of digital records based on the four categories listed above.

Establishment of Authority:

Clearly defined authority for both records management and preservation is essential for a digital archives to be successful. Without this authority, the ability of the State Archives to collect and manage records is compromised. The primary public official responsible for the proper management and preservation of records within the State of Hawai'i is the state Comptroller, head of the Department of Accounting and General Services (DAGS). The State Archives is within DAGS and has been delegated by the Comptroller as the permanent repository for state records of enduring legal, historical or fiscal value.

The authority of Comptroller to determine the disposition of state government records, other than judiciary or legislative branch, is established in §94-3 HRS, which states:

> The comptroller shall determine the disposition of the records; stating whether the records should be retained by the office, department, or bureau; be transferred to the public archives, the University of Hawai'i, the Hawaiian Historical Society, or other agency; or be destroyed. The comptroller shall have full power of disposal of all records submitted for this purpose.

Section 94-1 HRS authorizes DAGS to "collect all public archives; arrange, classify, and inventory the same; provide for their safekeeping; and compile and furnish information concerning them." Section 26-6 HRS locates major records responsibility with DAGS by investing in the department the authority to "manage the preservation and disposal of all records of the State" and to "administer the state risk management program."

Section 94-4 HRS states the archivist is "authorized and empowered to certify, as true and correct, copies or reproductions of any of the books, documents, papers, writings, or other records, or excerpts therefrom in their custody."

Recommendations

The Hawai'i Revised Statutes provide sufficient authority to the Comptroller to create and enforce the necessary policies and procedures to collect and manage digital records within a common defined

framework. Based on the experience with other state archives digital archive programs, the consultant has strongly recommended that the State of Hawai'i explicitly detail the duties and authority of the State Archivist in a Comptroller's memorandum or through modification of §94-1 HRS in order to provide the State Archivist with documented authorization to actively collect and preserve records of long-term legal, historical and fiscal value.

Definition of Record

A clear, statewide understanding of what is a record and what is not a record is important in determining what to keep and what to dispose of in accordance with the retention schedules. The Revised Statutes define records in several locations: a record is "information maintained by an agency in written, auditory, visual, electronic, or other physical form" [§ 92F, HRS]; public records are "any writing, memorandum, entry, print, representation, report book or paper, map or plan, or combination thereof, that is in the custody of any department or agency of government" [§ 94-3, HRS]; and personal records are "any item, collection, or grouping of information about an individual that is maintained by an agency. It includes, but is not limited to, the individual's education, financial, medical, or employment history, or items that contain or make reference to the individual's name, identifying particular assigned to the individual, such as a finger or voice print or a photograph" [§ 626, HRS].

The definition of 'record' must also encompass all formats that a record may be created in – including digital. Act 177, Session Laws of Hawai'i, Regular Session of 2005 (HB 515) expressly allows the use of government records in a digital format. The act also updated Hawaii Revised Statutes, Chapters 46, 92 and 94 to broaden the language surrounding records to include digital formats. The acceptance of digital records as having evidentiary value is clearly articulated in §92-30 HRS which states that an "electronic copy of a government record shall be deemed to be an original record for all purposes, including introduction in evidence…" This is reiterated in the Hawai'i Rules of Evidence, Chapter 626, HRS: "A "public record" means any writing, memorandum, entry, print, representation, report, book or paper, map or plan, or combination thereof, that is in the custody of any department or agency of government." And finally in §489E-2 HRS, Uniform Electronic Transaction Act, "'record' means information that is inscribed on a tangible medium or that is stored in an electronic or other medium and is retrievable in perceivable form."

Recommendations

The definition of records within the statutes meets the clear and concise requirement. The wording of "written, auditory, visual, electronic, or other physical form" allows for digital records to be included in the term 'record' and is sufficiently open to account for records created in or stored on future technologies. Section 92-30 HRS continues that definition by explicitly stating that digital copies of records are also accepted as records in their own right. It is recommended that the definition of record be identical throughout all the statutes, with sufficiently broad language to be inclusive of digital formats.

Management of Digital Records

The ability to maintain and preserve digital records is directly related to how the records are created, stored, and disposed of by the record's creators and custodians -- necessitating strong rules and regulations governing the management of digital records. Section 94-1 HRS empowers DAGS with ability to adopt, amend, or revise from time to time such rules and regulations as it may consider expedient for the conduct of its business. Section 94-3 HRS establishes the Comptroller's authority to determine the disposition of public records by stating "whether the records should be retained by the office, department, or bureau; be transferred to the public archives, the University of Hawai'i, the Hawaiian Historical Society, or other agency; or be destroyed." The combination of the two statutes provides the Comptroller with the authority necessary to meet the required "clearly worded rules and regulations that include the management of electronic records."

Recommendations

The statutes provide the necessary strong language required to allow the State Archives to enact requirements for the creation, maintenance, and handling of digital records by the agencies that are essential for digital archives to be successful.

Retention of Records in Digital Format

Section 489E-12 HRS states that any records retention requirements may be satisfied by maintaining a digital version of the record provided that it accurately reflects the "information set forth in the record after it was first generated in its final form" [§489E-12 (1)] and that it remains accessible for later reference. Section 489E-7 specifically allows for the maintenance of digital records even when a law requires that the record be in writing, and likewise allows for the use of an electronic signature in instances where a signature is required.

Recommendations

Section 489E-12 allows agencies to maintain digital versions of records provided that the integrity and accessibility of the record remains intact, but does not provide sufficient clarification on how agencies can satisfy these requirements. Clarification can be provided by DAGS as section 489-12 (G) states that it does not "preclude a governmental agency of this State from specifying additional requirements for the retention of a record subject to the agency's jurisdiction," and as already demonstrated, all records within state government are within DAGS jurisdiction. It is therefore recommended that DAGS, under the authority granted to it by §26-6 HRS, §94-1 HRS and §489E-12 (G), produce additional requirements that clearly delineate how agencies can satisfy the dual requirements of maintaining the integrity and accessibility of digital records of long-term value. It is also recommended that one of the requirements set forth should entail State Archives approval of any process that migrates, coverts or reformats records to ensure that the requirements stipulated in §489E-12 are being met. With regards to electronic signatures, it is recommended that the State Archives develop guidelines on appropriate methods and technologies

that will allow for the long-term preservation of electronic signatures -- particularly those based on a Public Key Infrastructure (PKI) or Certificate Authority (CA) based technology.

Preservation Media

The statutes concerning digital records make no mention of media formats that can or cannot be used. As a result, a wide variety of formats are being utilized both within and between agencies, as evidenced in the digital records survey recently conducted by the State Archives. Due to the underlying technology and manufacturing techniques, certain media types are better suited to long term preservation and the delay of obsolescence than other formats. Without any guidance, agencies are unknowingly utilizing media formats and technologies that are placing vital digital records at an increased risk of loss, alteration or technological obsolescence, as well as an inability to access records of evidentiary value at a future date.

Recommendations

It is recommended that the State Archives publish a list of recommended media formats, as well as guidelines for the proper care, storage and refresh frequencies of each media type. It is also recommended that the State Archives draft minimum requirements for the migration and/or refresh of magnetic media, as well as the processes that must be followed to ensure the integrity and accessibility of the records.

Microfilm and Imaging Systems

Digital Imaging systems are heavily utilized in the state to convert paper records to digital formats. While some agencies are utilizing outside vendors for the conversion process, a majority of the state agencies are conducting the imaging in-house. Section 92-29 HRS allows for any public officer to reproduce digitally or on microfilm any record in their custody. While this section provides the authority, no standards for the reproduction are provided. There is a check on this authority: section 92-31 HRS requires state agencies to obtain written permission from the Comptroller prior to the destruction of such records.

Recommendations

As the agencies must exclusively rely on the digital image after the destruction of the paper original, the processes and equipment used to create and maintain these digital records needs to be regulated and operated within an approved framework for all state agencies. It is recommended that the State Archives be given oversight on the purchase of equipment by state agencies for the purpose of converting paper into digital or microform, as well as any new contracts with outside vendors for imaging services. As illustrated in the digital records survey conducted by the State Archives, many agencies engage in digital imaging in order to trim expenses by destroying the original paper records. It is also recommended that clearly articulated imaging guidelines be drafted setting forth the minimum accepted requirements for digital imaging (to include file formats, scanning resolution, color depth, etc.) and necessary system and process documentation that agencies must produce. Additionally, it is recommended that §92-31 HRS be amended to authorize the State Archives to take into its custody a digital copy of any records so created under this

section, in addition to the pre-existing provision allowing for the transfer of the version proposed to be destroyed. Lastly, it is recommended that as part of the imaging process authorized under §92-29 HRS, agencies be required to delineate the integrity and document the conversion process to the satisfaction of the State Archives prior to destruction of the paper records.

Records Management Guidelines

Under the authority vested in the Comptroller by §94-1 HRS, the State Archives publishes a number of guidelines for use by state employees: Records Management Manual, Disposal of Government Records, Vital Records Protection Policy and Guidelines, and Storage of Inactive Paper Records. These guidelines provide a general overview of the records management processes recommended by the Archives for managing both paper and digital records within the state records management framework. These guidelines are articulated as recommendations with no stated requirements for compliance.

Recommendations

The guidelines have been widely distributed for several years and are easily accessible on the home page of the State Archives website. While these documents provide a good overview of the records management process, they do not provide sufficient information for more advanced records producers or any guidance on the selection or implementation of records systems. As technology changes rapidly, it is important that digital records be created, maintained, and preserved in records systems that are sufficiently open to allow for migration to future generations of software. Without advanced planning and guidelines, the state runs a risk of either permanently losing records of long-term value or incurring significant costs to migrate legacy data into a current, more useable format. If the Digital Archives is to be successful, efficient and effective, proactive management of digital records at the agency level will be essential. Requiring digital records to be managed following a common policy will allow for a seamless transfer of records from the agency of creation to the Archives. Therefore, it is recommended that guidelines be published providing direction to agencies on how best to manage and maintain digital records in a consistent methodology that is within the framework of the state records management policies. Guidelines should be drafted that address the minimum requirements necessary to protect digital records from loss, alteration, deterioration, and technological obsolescence. The guidelines should include recommendations on file formats, format migration, disaster recovery provisions, document security on protected networks, and media use and refresh. After a set period of time for the agencies to use the guidelines in development or modification of their agency records management policies, it is recommended that the guidelines become mandatory policies that all agencies must follow.

Public Access to Records

Chapter 92F Part II Freedom of Information delineates the public's right to open access of government records:

> In a democracy, the people are vested with the ultimate decision-making power. Government agencies exist to aid the people in the formation and conduct of

> public policy. Opening up the government processes to public scrutiny and participation is the only viable and reasonable method of protecting the public's interest. Therefore the legislature declares that it is the policy of this State that the formation and conduct of public policy--the discussions, deliberations, decisions, and action of government agencies--shall be conducted as openly as possible. [§92F-2]

Section 92F-11 states that "All government records are open to public inspection unless access is restricted or closed by law." This provision "applies prospectively, requiring disclosure of records maintained by state agencies regardless of when the records came into existence." Each agency is required to issue instruction and guidelines necessary to ensure the public's access to government records in §92F-18, including compiling a public report "describing the records it routinely uses or maintains" [§92F-18(2b)]. Providing the public access to records is covered in section 26-6 which requires DAGS to "provide a long-term means for public access to public information" [26-6 (10(1))] and to "adopt rules as may be necessary or desirable for…the operation and implementation of a program to provide a means for public access to the State's…public information." [§26-6 (10(4D))].

Recommendations

The public access to government records is clearly articulated in the Statutes; meeting the fourth requirement for 'strong laws governing the public's right to access government records.' As can be evidenced in the digital records survey, few agencies have the policies, procedures, and technical infrastructure required to comply with requirements set forth in Chapter 92F. Technology changes every three to five years, necessitating a migration of the records onto newer media and/or software formats. Failure to keep pace with technology will result in the loss, alteration deterioration or technological obsolescence of those very records of government that are "the only viable and reasonable method of protecting the public's interest" [§92F-2]. The creation and implementation of a digital archives by DAGS, under which the State Archives falls, is legislatively authorized by §26-6 in order to "provide a means for public access to the State's…public information." [§26-6 (10(4D))]. To ensure long term access to digital records of enduring legal, historical and fiscal value, it is strongly recommended that a copy of those digital records with retention periods exceeding ten years, or those that have been determined to be of vital importance to the operation of government, be sent to a centralized digital archives for permanent preservation. By centralizing the records, and placing them in the care of trained, dedicated, professional staff, pro-active management and migration of the digital records can be conducted in a cost effective manner in accordance with international standards. An important secondary benefit of such a consolidation will be the ability to provide a single access point to all citizens of the state, regardless of their location and ability to travel, 24 hours a day, seven days a week.

Staffing Recommendations

Based on the experiences of other state digital archives, the staffing required for the Hawai'i Digital Archives project to move forward is comprised of three major groups: executive management, technical staff and professional staff.

Executive Management:

Success of the Digital Archives project will require strong executive management support and sponsorship. The State Archivist manages the State Archives and is tasked by HRS § 94-1 to "collect all public archives; arrange, classify, and inventory the same; provide for their safekeeping; and compile and furnish information concerning them." Over the past four years, the State Archivist has spearheaded the support of this project and recognizes that the digital archives is a needed component for the State Archives if it is to fulfill its mandate of preserving the public records of Hawai'i. In addition to the State Archivist, the State Comptroller as head of the Department of Accounting and General Services (DAGS) will play an important role in providing support for the project. As the initiative moves forward, legislative changes to the records laws and funding streams will need the continued backing and leadership of executive management in order to be successful.

Technical Staff:

The technical staff is responsible for the development and management of the computer based systems that will be central to the digital archives. For this initiative, the technical staff will be responsible for:

- providing technical expertise as functional and business requirements develop;
- providing system specifications for necessary hardware and software based on budget and need;
- evaluating technical solutions offered by vendors and used in other states;
- installing, or overseeing the installation of, the computer system infrastructure as necessary;
- creating, or modifying, the digital archives software; and
- developing transfer mechanisms that allow agencies to send records to the Digital Archives.

The State Archives currently does not have any information technology staff within the unit. Any required information technology related services are provided by the Information and Communication Services Division (ICSD) and the Systems and Procedures Office (SPO) of DAGS. While ICSD, SPO and the State Archives are in the same department and have strong working relationships, full-time technical support of State Archives projects are not in the scope of ICSD's or SPO's responsibilities. To adequately support the digital archives project, dedicated technical staff will need to be hired.

Professional staff:

Administrative staff will provide the necessary project management, budgeting oversight, and subject matter expertise for the creation, management, and preservation of digital records. While collectively the professional staff at the State Archives possesses a high level of skill in and knowledge of archival and records management administration, the current staff does not have time available in their current workloads or any practical experience with the maintenance or preservation of digital records. To operate a digital archives, the State Archives will need to increase the collective technical skill level of its staff, through training and new hires.

Recommendations:

Cross Functional Project Team

In order to develop a system that meets the needs of all of its stakeholders and mitigate the issues that inevitably arise, it is recommended that a cross-functional team be assembled for the duration of the project. In addition to the executive, technical and administrative staff listed above, it is recommended that representatives from IT, Human Resources, and Accounting Services be part of the project team to provide input throughout the process. Likewise, it is recommended that the archives staff members listed above also work on the project team to share their expertise in the development of the policies and procedures that will be used by the digital archives. The cross-functional representation of the project team follows the best practice methodology as defined in the Joint Application Development process and has been used to great success on other government digital preservation projects. In particular, ICSD possess extensive technical skills and knowledge of State and agency specific hardware, software, and network configurations, while Human Resources and Accounting Services bring both an end user perspective to the team and subject matter expertise in staffing and budgeting that will make important contributions to the project.

Technical Staff

The current staff of the State Archives is highly skilled at archives administration and records management, but collectively possesses a deficit of computer and technology related experience. To successfully operate a Digital Archives, dedicated staff with knowledge and experience in computer-based systems is essential. To address this need, it is recommended that additional staff be hired and specifically assigned to the Digital Archives project. To provide the minimum necessary skill set to operate a digital archives, a minimum of two additional full-time positions is required. The recommended positions are a Digital Records Acquisition Specialist, to manage the flow of incoming digital records, and a Systems Developer, to manage the system infrastructure and develop additional software functionality. As the project continues to grow, and more agencies start sending records to the Digital Archives, additional staff resources will

need to be added to match the increased workload generated by these agencies. Additional staff increases can be mitigated to some extent through the use of interns and contract staffing.

The Project Consultant will have the primary responsibility of oversight on the project to ensure that, first and foremost, the system is designed and implemented in accordance with international standards and current best practices for the long term, trustworthy preservation of digital records, while also serving as the subject matter expert on issues relating to digital preservation systems. Additionally, the Project Consultant will interface with the archives staff, agency project partners, State IT staff and the public to ensure that the functional requirements and satisfaction criteria are documented and achieved. Recommended core skills for the Project Consultant include:

- Project management experience in leading a software development team
- Experience managing a digital archives
- Experience managing projects in a government environment
- Demonstrable knowledge of and experience in archival science and computer science

The Digital Records Acquisition Developer will have the primary responsibility of identifying both current and legacy digital records of enduring legal, historical, or fiscal value and arranging the transfer of these records to the Digital Archives. Recommended core skills for Digital Records Acquisition Specialist include:

- Computer Science background with 2-3 years of Java Programming
- Experience creating and manipulating XML and Metadata
- Previous experience with the transfer, identification and maintenance of digital records
- Strong customer service background

The Systems Developer will be responsible for developing and maintaining the Digital Archives systems, creating the public search interfaces and increasing the system functionality to meet stakeholder needs. Recommended core skills for the application developer include:

- Strong Computer Science background with intermediate level (3-5 years) java programming
- Experience maintaining MySQL or PostgreSQL databases
- Experience creating functional requirements documentation
- Experience creating, documenting and executing software test plans

In addition to the two staff positions above, it is strongly recommended that the current project manager continue to oversee the next phase of the project, including the creation of project requirements, timelines, budgets, staff recruitment, etc. To succeed in the position, the project manager must be given the necessary authority to drive the project forward to ensure its successful completion. Building upon a successful project kickoff meeting and subsequent digital records training, there currently exists a large amount of 'good will' among legislators, stakeholder groups, agencies and customers that must be built upon to support the submission of any legislative and funding changes. It is recommended that Executive

Management work in concert with the project manager to focus on these targeted groups so that allies and interested stakeholders can be located.

It is also recommended that the State Archives enter into a partnership with the Computer Science Departments' of the University of Hawai'i and Hawai'i Pacific University in order to form an internship program. Ideally, such a partnership will allow the best and the brightest students from the University to work on the digital archives project; thereby allowing student workers to increase system functionality while decreasing the workload on the Archives technical staff. By collaborating with the University of Hawai'i, the State Archives will be able to recruit, on a temporary basis, bright, capable technical staff to work on focused projects. The University will be able to place students into high visibility, high impact projects in a real world setting. The students will gain valuable work experience and build their portfolio. Such a partnership would create a win-win situation for all parties involved.

Review and Recommendations of Potential Funding Sources

The State Archives has insufficient funding to staff and operate a digital archives. With an annual operating budget of approximately $754,000 for FY2012 and 16 full time staff, the staffing levels and funding are currently inadequate to provide the appropriate level of operation, preservation, and access to the 63,000 cubic feet of paper and analog holdings in the State Archives and State Records Center; let alone taking on the task of providing preservation and access to the permanent digital records of the State. The requirement to preserve records of permanent value is clearly stated in §94-1 HRS, but to fulfill this mandate the State Archives will need the budget, technical staff and the computer infrastructure to fulfill this requirement.

Alternatives Considered:

Based on the research conducted for this study, five potential streams that have successfully been used by State Archives to fund similar initiatives have been identified: fees on recorded documents, agency charge back, Special Fund allocation, increase in General Fund allocation, and grants.

- Dedicated fee on recorded documents -- The concept behind the fee assessed on recorded documents is that those documents that are publically recorded (such as land, marriage, power of attorney, court, etc.) are done so in order to establish and protect a citizen's rights. As these records are of permanent value, they are eligible to be preserved within the State Archives. Specific categories of records that are known at the time of creation to be of permanent value should, therefore, fund their own preservation. Typically a fee of between one and five dollars is added to the other fees already collected at the time of the recording, resulting in a marginal increase of approximately 2-8% in the overall cost of publically recording a document. The small increase over the many records that are recorded in a typical year should provide sufficient funding for the digital archives.

- Agency charge back -- Mirroring the charge-back schema already utilized at the State Records Center for special fund agencies, agencies sending records to the Digital Archives could be charged a pro-rated share of the overall operations costs based on their agencies proportion of the overall records stored within the archives. The concept behind agency charge back is that the agencies producing records of permanent value have a need to continue to have access to these records over both the near term, as well as infrequent use over the long term. Where agencies do not have the expertise or resources to store their digital records over the long term in a trustworthy state, these records should be placed in the custody of trained professionals. By transmitting the records to the State Archives, agency storage requirements, server utilization, and back-up widows would all decrease -- resulting in an overall decrease in the agency's IT budget. These cost savings could then be shared with the State Archives to fund the Digital Archives initiative. A potential drawback to this funding stream is that a charge back could be seen as a disincentive for agencies to transmit records of permanent value to the State Archives.

- Dedicated Special Fund -- A special fund pursuant to section 36-27 could be created similar to the shared services fund created for the State CIO (see S.B. 2548, S.D. 2, H.D. 1, C.D.1 , http://www.capitol.hawaii.gov/session2010/Bills/SB2548_CD1_.HTM). Such a dedicated fund could be dedicated specifically to the preservation of the records of permanent legal, historical, or fiscal value. The creation of such a fund would reinforce the importance of archival records to the citizens and employees of the state in supporting and advancing their shared understanding of the workings of state government, the reasoning behind the decisions that were made and policies created, and culture of its native and immigrant peoples.

- Increase in State Archives General Fund allotment -- The State Archives was established in 1905, when its primary responsibility was the preservation of the paper records of state government. Since that time, there have been only incremental increases in the agency's budget to account for the rising costs of operations, to support the increase in records acquired, and to provide state employees' and citizens' access to the records currently held within the State Archives; and several large staffing and budget cuts mirroring those of other state agencies during times of budget shortfalls. The State Archives budget was established at a time when digital records did not exist. As such, the State Archives has never been appropriately funded to acquire the infrastructure and skill set necessary to preserve digital records. In recognition of the massive expansion of the domain of information that the State Archives is now responsible for -- if it is to execute its mandate of preserving the records of the State of Hawai'i that are of permanent legal, historical or fiscal value – the State Archives general operating budget could be increased proportionately to the increased resources necessary to address this new workload.

- Federal Grants -- Grants rely upon locating and obtaining external funds, typically with a matching funds requirement. While obtaining grants will cut the State's financial burden for a digital archives in half, the State will still need to fund the other half (through the possibility of one of the two funding sources above). Additionally, grant funding is always of a short duration, typically two to five years, providing only a short term financial benefit; and typically have a lead time of 6-18 months between application and receipt (if the proposal was in the very small percentage of projects selected by the grant oversight committee) of the project funds. This delay can create a level of disruption in the program if work on the project must be stopped halfway through while waiting for additional grants funds to be received.

Recommendations:

Creation of a dedicated funding stream, for the equipment and staffing of the digital repository is strongly recommended. A dedicated funding source will allow the digital archives to recruit the necessary technical

staff and purchase the hardware and software needed to manage and preserve digital records of long-term value. Without sufficient resources to preserve the records in their digital format, the state either risks losing access to the records or being forced to print them out onto paper. Current industry estimates that greater than 95% of all records are produced today in a digital format. If all the digital records that are required to be retained for more than ten years are printed onto paper, the increased volume of records needing to be stored at the State Archives will rapidly outpace the amount of storage space available. This situation would create a cascading effect of requiring that additional storage capacity be added to the existing building, or suitable facilities be located and leased at a higher annual cost than operating a digital archives. As it is unrealistic in the current economic environment to expect that the State Archives budget simply be increased by the necessary amount, alternative funding sources have been explored.

Of the five fee structures investigated, it is strongly recommended that the State Archives pursue the strategy of a dedicated fee on recorded documents along with applying for any grants that are made available by OCLC, NHPRC, Library of Congress or any other historic records source. The states of New Jersey and Washington have both successfully used a recording fee to fund their respective initiatives and to create an in-state grant fund to push money back to the local government agencies collecting the fees. Collecting fees in advance on those records most in need of long-term preservation allows for an overall greater level public access to government records by allowing free and open access to everyone. Given Hawai'i's geography, allowing citizens on neighboring islands open access to all public records without having to travel between islands fulfills the spirit and intent of public disclosure requirements of Chapter 92F HRS.

Recommendations for Disaster Recovery

In the event that the digital archives network is severely damaged or destroyed and the technical staff is inaccessible, the disaster recovery documentation outlined within is intended to provide all the necessary resources to rebuild/restore the digital archives to the identical level of functionality and content that existed when the "snapshot in time" for the Red Box was built.

Definition of a Red Box:

Everything intellectual that is needed to rebuild the digital archives to a specific state – with sufficient documentation to purchase replacement physical items and reconstruct the physical system as it originally existed, such as: building, Power, HVAC, and hardware (racks, servers, tape library, switches, PIX, cables, etc.). The Red Box provides all the necessary 'tools' for any competent IT professional to rebuild the digital archives by including:

- Instructions
- Complete system documentation
- Software disks
- Disk Images of all servers and
- Data backups

The Red Box derived its name from the Red Tape used to seal the box, the red color drawn from the allusion to danger and disaster.

Scope

It is intended that the Red Box be an all-encompassing, self-contained 'digital archives in a box'. As such, it should ONLY contain information on the digital archive system configuration; from the Intrusion Prevention System all the way back to the storage devices. The network infrastructure and hardware leading up to the 'front door' of the digital archives is outside the scope of this process. The only resources needed to rebuild and operate the digital archives outside of the Red Box are the 3 Ps: ping, power and POP (network, power and a place to put it)!

Creation and Use

The remainder of this document details the contents of the Red Box, divided into six sections each with a specific focus. It is recommended that each section be stored separately in its own binder so that multiple individuals can concurrently work on system restoration in separate areas. As the Red Box contains the complete network infrastructure diagrams and account credentials, the Red Box *MUST* be kept in a secure, protected location at all times. The recommended process is to place the contents in an appropriately sized container that is sealed with security tape and signed by the department manager and network administrator. As it is intended that the Red Box will only be used in the event of a large-scale disaster, it is

vital that staff outside of the Archives know of its location and be provided access. It is recommended that the State Comptroller and the ICSD Administrator be made aware of the location of the Red Box, along with the State Archivist in the event of a statewide disaster. Due to the sensitive and confidential nature of the materials contained within the Red Box, it is important that its location be restricted to a "need to know" basis only.

The recommended elements of the disaster recovery manual are included in Appendix E.

Appendix A: Digital Archives Survey

As mentioned previous, a cross-functional group of records users, records creators, and IT professionals from all branches of government, higher education and all the counties were invited to take the survey. The primary method of survey completion was online through Survey Monkey, with paper based options presented as well and the results hand entered into the online survey by the archives staff. The original survey instrument, attached below, is followed by the tabulated results of all the respondents. Based on these results, the project consultant provided general recommendations to address specific issues either raised directly by the survey participants or from the accumulated results for the various questions and sections.

[Survey Instrument starts on next page]

Section 1. Hawai'i State Digital Archives Survey

The Hawai'i State Archives greatly appreciates your willingness to participate in this survey. This survey is expected to take 5-15 minutes to complete. Responses will remain anonymous unless you choose to provide your contact information. Your feedback is an important part of creating the look, feel, function, and focus of the Hawai'i State Digital Archives. More information on the Digital Archives Initiative can be found at http://Hawaii.gov/dags/archives. If you have any questions about the project or this survey, please contact Gina Vergara-Bautista at gina.s.vergara-bautista@Hawaii.gov or 808.586.0329.

NOTE: Items with an * are required!

Section 2. Digital Records: Master Plan

***Q1: Has your agency developed a master plan for handling/managing digital records (creation, storage, access, privacy, and imaging)?**

☐ Yes

☐ No

Section 3. Digital Records: Master Digital Records Plan Contact

Q2: Is the master plan for handling/managing digital records available, either internally or on the web?

☐ Yes

☐ No

Q3: Who is the project lead for the master plan? (Name, phone and email please)

Name:

Email Address:

Phone Number:

Section 4. Digital Records: Media Formats For Storage

***Q4: What medium(s) do you currently store your digital records on? (Check all that apply)**

- ☐ CD/DVD
- ☐ Diskette
- ☐ USB
- ☐ Magnetic Tape
- ☐ Zip Drive
- ☐ Network Drive
- ☐ Desktop Hard Drive
- ☐ Removable Hard Drive

Other (please specify)

Section 5. Digital Records: Agency Backups

***Q5: Are your digital records backed up?**

- ☐ Yes
- ☐ No
- ☐ Don't Know

Section 6. Digital Records: Agency Backup Formats

Q6: Who creates the backups?

- ☐ I do

- [ ] The IT staff at my agency does
- [ ] We both do

Q7: What media format(s) the backups stored on? (Check all that apply)

- [ ] CD/DVD
- [ ] Diskette
- [ ] USB
- [ ] Magnetic Tape
- [ ] Zip Drive
- [ ] Network Drive
- [ ] Desktop Hard Drive
- [ ] Removable Hard Drive

Other (please specify)

Section 7. Digital Records: File Formats Produced In The Agency

***Q8: What format(s) of digital records do you produce? (Check all that apply)**

- [ ] ASCII/TXT/RTF
- [ ] Image (BMP, JPG, PNG, PSD, etc.)
- [ ] HTML
- [ ] PDF

- [ ] MSFT Word (DOC, DOCX)
- [ ] MSFT Excel (XLS, XLSX)
- [ ] Open Office (ODF, ODS, etc.)
- [ ] XML
- [ ] Database

Other (please specify)

Section 8. Digital Records: Email Management

***Q9: Do you currently manage your email (such as: separating record from non-record, disposing of emails when retention reached, etc.)?**

- [ ] Yes
- [ ]

Section 9. Digital Records: Retention Schedule

***Q10: Does your agency have a records management schedule that includes digital records?**

- [ ] Yes
- [ ] No
- [ ]

Section 10. Digital Records: File Migration

***Q11: Does your agency have a plan in place to migrate your digital records as new versions of your current software or file formats become available?**

- [ ] Yes

☐ No

☐ Do

Section 11. Digital Records - Migration Plan

Q12: Is the migration plan for digital records available, either internally or on the web?

☐ Yes

☐ No

Q13: Who maintains the migration plan? (Name, phone and email please)

Name:

Email Address:

Phone Number:

Section 12. Scanned Documents: Agency Scanning

*Q14: Does your agency convert paper documents to electronic images?

☐ Yes

☐ No

☐ Don't Know

Section 13. Scanned Documents: In-House or Vendor

Q15: Is the document scanning done in-house or by a vendor?

- ☐ In-house
- ☐ Vendor
- ☐ Both

Section 14. Scanned Documents: Volume Currently Scanned

Q16: What is the approximate volume of records already scanned?
Example: 1000 pages, 10 cubic foot boxes, 8 rolls of microfilm

Section 15. Scanned Documents: Volume Waiting to be Scanned

Q17: What is the approximate volume of records waiting to be scanned?
Example: 1000 pages, 10 cubic foot boxes, 8 rolls of microfilm

Section 16. Scanned Records: Date Range

Q18: What is the approximate date range of scanned records maintained at your agency?
Example: 1992-94, 1996-2010

Section 17. Scanned Records: Business Purpose

Q19: Why did your agency decide to scan/image your records? (Check all that apply)

- ☐ Save Space
- ☐ Faster Retrieval

- [ ] Better Management
- [ ] Multiple Access
- [ ] Security Backup

Other (please specify)

Section 18. Scanned Records: Ongoing Process

Q20: Is the imaging:

- [ ] A One-time Process
- [ ] An Ongoing Operation
- [ ] Periodic

Other (please specify)

Section 19. Scanned Records: Records Disposition

Q21: After scanning, what happens to the original paper records? (Check all that apply)

- [ ] Maintained In-house
- [ ] Sent Off-site
- [ ] Microfilmed
- [ ] Destroyed

Other (please specify)

Section 20. EDMS: Agency Usage

***Q22: Does your agency use an electronic document management system (EDMS)?**
An EDMS is any system that centralizes and manages electronic documents.

- [ ] Yes
- [ ] No

Section 21. EDMS: System Identification

Q23: What is the name of your EDMS system?

Q24: Who is the Vendor of the system?

Section 22. EDMS: Born Digital Records

Q25: Does your agency EDMS store born digital records?
Born digital records are those digital records that never existed in a paper or analog format.

- [ ] Yes
- [ ] No

Section 23. EDMS: File Types

Q26: What types of digital records are stored in your EDMS system? (Check all that apply)

- [ ] Email
- [ ] Office Documents (word processing, spreadsheets, etc.)

- [ ] Images
- [ ] Database outputs/reports
- [ ] Scanned Images

Other (please specify)

Section 24. EDMS: Metadata

Q27: Does your agency collect any additional metadata to help identify and manage the records in the EDMS?

- [ ] Yes
- [ ] No

Section 25. EDMS: Metadata Types

Q28: What additional metadata is collected? (Check all that apply)

- [ ] Agency Name
- [ ] Title of Record
- [ ] Disposition of Record
- [ ] Destruction Date of Record
- [ ] Access Restrictions
- [ ] Case Number
- [ ] Document Number

Other (please specify)

Section 26. General: Records Management Training

Q29: Have you ever attended records management training of any kind?

- ☐ Yes
- ☐ No

Section 27. General: RM Training Date

Q30: When did you last attend records management training?

- ☐ Last Six Months
- ☐ Last Year
- ☐ Last 1-3 years
- ☐ More Than Three Years Ago

Section 28. General: Definition of a Record

*Q31: Do you understand the difference between what qualifies as a business record and what does not?

- ☐ Yes
- ☐ No
- ☐ Not Really

Section 29. General: Records Pain Points

Q32: What problems, if any, are you encountering in the creation, storage, retrieval, privacy, destruction, maintenance, or management of digital records?

Section 30. General: Records Guidelines

***Q33: Would guidelines and/or standards for the management of digital records be helpful?**

- ☐ Yes
- ☐ No

Section 31. General: Training Needs

***Q34: Would you attend training on managing digital records if it was offered?**

- ☐ Yes
- ☐ No

Section 32. General: Focus Group Participation

***Q35: Would you be willing to share your needs and experiences with a focus group developing the Hawai'i State Digital Archives requirements?**

☐ Yes

☐ No

Section 33. General: Focus Group Contact Info

Q36: Please provide your contact information so we may send you details about our focus groups.

Name:

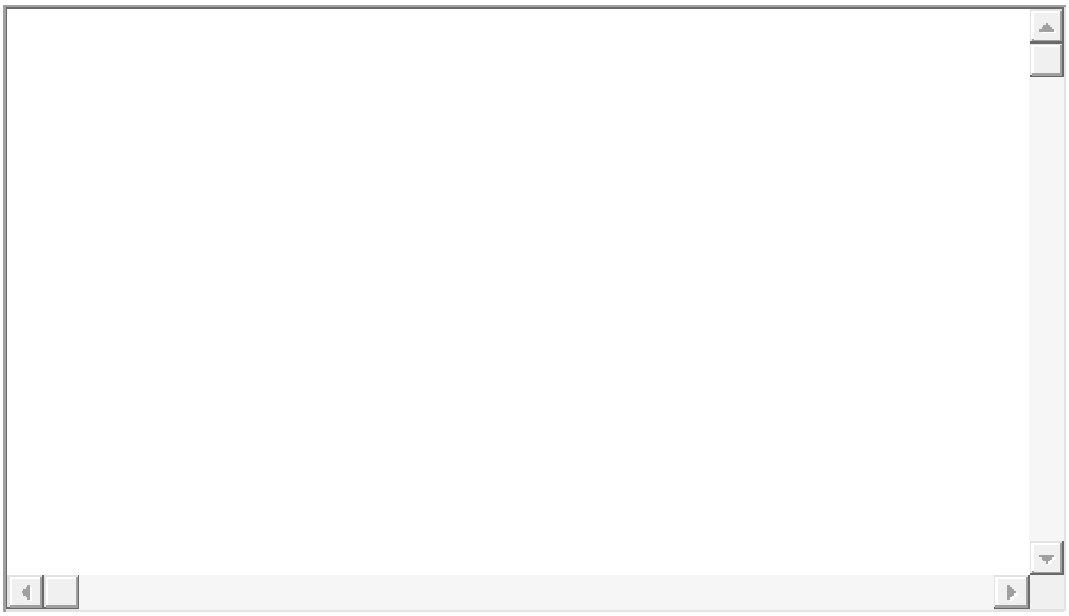
Agency:

Email Address:

Phone Number:

Section 34. General: Last Thoughts

Q37: Please share with us any other comments you have regarding records and their management.



End of Survey

Appendix B: Analysis of Survey Results

Digital Records:

1. Has your agency developed a master plan for handling/managing digital records (creation, storage, access, privacy, and imaging)?		
Answer Options	**Response Percent**	**Response Count**
Yes	14.3%	11
No	85.7%	66

Findings and Recommendations: 86% of participants are without clear agency direction on the appropriate procedures for creating, maintaining and preserving digital records. The lack of a master plan for digital records places agencies at risk due to an ability to produce the digital records required in a public disclosure request pursuant to §92F-12 HRS and of non-compliance with records retention requirements. It is recommended that all agencies be required to have at least a rudimentary plan in place on the expectation of how agency employees with manage their digital records.

2. Is the master plan for handling/managing digital records available, either internally or on the web?		
Answer Options	**Response Percent**	**Response Count**
Yes	41.7%	5
No	58.3%	7

Findings and Recommendations: Of those agencies that do have a plan for managing digital records, 58% of those plans are not easily accessible by the employees of that agency. Without easy access to master plans, consistent enforcement of the retention requirements is difficult. It is recommended that agency master plans be published in easily accessible locations, such as the agency intranet site, and periodically sent to all employees as a reminder of such policies.

3. Who is the project lead for the master plan? (Name, phone and email please)		
Answer Options	**Response Percent**	**Response Count**
Name:	100.0%	6
Email Address:	66.7%	4
Phone Number:	83.3%	5

Findings and Recommendations: Half of the participants who responded that their agency had a master plan provided the contact information for the individual who manages that plan. Further investigation into the extent of the agency master plan for digital records is recommended. It is also recommended that a model master records plan be developed for other agencies use from the existing master plans and archival best practices.

4. What medium(s) do you currently store your digital records on? (Check all that apply)		
Answer Options	**Response Percent**	**Response Count**
CD/DVD	61.0%	47
Diskette	13.0%	10
USB	27.3%	21
Magnetic Tape	33.8%	26
Zip Drive	14.3%	11
Network Drive	83.1%	64
Desktop Hard Drive	59.7%	46
Removable Hard Drive	31.2%	24
Other (please specify)		7

Findings and Recommendations: As expected, the agencies employ a wide variety of media for the storage of digital records. Of particular concern is the 61% that use CDs/DVDs to store digital records --research has shown that CDs have a high media failure rate due to a phenomena called CD rot -- and the 14% that use Zip Drives, a technology at risk of becoming technologically obsolete. Of the 'other' responses, three were networked file shares, two were microfilm and one was an Oracle database. It is recommended that the State Archives issue advisories on recommended storage media and educate state employees of the dangers of unstable and obsolete media formats.

5. Are your digital records backed up?		
Answer Options	**Response Percent**	**Response Count**
Yes	77.9%	60
No	2.6%	2
Don't Know	19.5%	15

Findings and Recommendations: 77% of digital records are known to be backed up and at least 2.5% are not backed up; meaning that as high as 23% of the records may not have any back up. When combined with the myriad of media types (from question 4) that are storing records, those that are not backed up are at a high risk of loss, alteration, deterioration or technological obsolescence.

6. Who creates the backups?		
Answer Options	**Response Percent**	**Response Count**
I do	11.7%	7
The IT staff at my agency does	53.3%	32
We both do	38.3%	23

Findings and Recommendations: 53% of the backups of digital records are created by the IT staff, with another 38% of the backups being performed in partnership with the records custodians. Further research on those performing backups of digital records is recommended in order to determine if the backups follow a consistent, auditable procedure and whether or not the backups are routinely checked for accuracy and completeness.

7. What media format(s) the backups stored on? (Check all that apply)		
Answer Options	**Response Percent**	**Response Count**
CD/DVD	38.2%	21
Diskette	9.1%	5
USB	14.5%	8
Magnetic Tape	63.6%	35
Zip Drive	10.9%	6
Network Drive	54.5%	30
Desktop Hard Drive	29.1%	16
Removable Hard Drive	25.5%	14
Other (please specify)		10

Findings and Recommendations: As the number of media type responses is greater than the total number of participants (145 selections from 55 participants), backups are being performed on an average of approximately three different types of media per participant. As noted in the findings of question four, a noticeable percentage of backups are on to 'at-risk' media types. Further research into the decision making process on selecting media types for backups is recommended. Based on the research findings, it is recommended that the State Archives issue advisories on the recommended backup media and educate state employees of the dangers of using unstable and obsolete media formats.

8. What format(s) of digital records do you produce? (Check all that apply)

Answer Options	Response Percent	Response Count
ASCII/TXT/RTF	29.9%	23
Image (BMP, JPG, PNG, PSD, etc.)	70.1%	54
HTML	42.9%	33
PDF	85.7%	66
MSFT Word (DOC, DOCX)	63.6%	49
MSFT Excel (XLS, XLSX)	51.9%	40
Open Office (ODF, ODS, etc.)	5.2%	4
XML	14.3%	11
Database	48.1%	37
Other (please specify)		9

Findings and Recommendations: A large percentage of records produced are text-based documents (85% of participants produce PDF, 64% word files, 30% text files), while 48% of participants produce digital records in databases, and 70% produce images files. Of particular concern is the 43% of participants producing HTML files, as web records are highly complex in their interrelationships with other web pages and the multiple components necessary to reconstruct the webpage. Two responses to the 'other' category also are of note: Geographic Information System (GIS) -- another file format that requires a strong interrelationship between files and components to render accurate information -- and audio and video files. It is recommended that further research be conducted into the types of digital records being produced, and the findings of this research be used to develop a phased-implementation plan for file type functionality in the digital archives.

9. Do you currently manage your email (such as: separating record from non-record, disposing of emails when retention reached, etc.)?		
Answer Options	**Response Percent**	**Response Count**
Yes	40.3%	31
No	59.7%	46

Findings and Recommendations: As email is a form of correspondence, in many instances email is legally recognized as a record of value and must be maintained and preserved according to a retention schedule. With 60% of participants not managing their email, these agencies are at risk of non-compliance with Hawai'i state law and an inability to produce records in a public disclosure request as required by HRS §92F-12. It is recommended that each agency conduct an 'email awareness' training session, in conjunction with the State Archives records management branch, to stress the importance of proper email management.

10. Does your agency have a records management schedule that includes digital records?		
Answer Options	**Response Percent**	**Response Count**
Yes	32.9%	25
No	38.2%	29
Don't Know	28.9%	22

Findings and Recommendations: A high percentage of participants either do not have (38%) or don't know of (29%) a retention schedule for their digital records. Without access to an approved retention schedule that includes digital records, public employees are at risk of either disposing of records too soon (before the approved retention allows for destruction) or of keeping the records longer than necessary (leading to higher storage costs and increased handling). It is recommended that every agency maintain an agency specific retention schedule and have it published in a readily accessible location.

11. Does your agency have a plan in place to migrate your digital records as new versions of your current software or file formats become available?		
Answer Options	**Response Percent**	**Response Count**
Yes	7.9%	6
No	52.6%	40
Don't Know	39.5%	30

Findings and Recommendations: Only 8% of participants know of a migration plan for their current digital records. As changes in hardware and software are inevitable and unavoidable in a digital world, without a clear migration strategy for digital records, the other 92% of records are at risk of technological obsolescence. It is recommended that all agencies create and publish a migration plan to avoid the loss of digital records due to technological obsolescence.

12. Is the migration plan for digital records available, either internally or on the web?		
Answer Options	**Response Percent**	**Response Count**
Yes	33.3%	2
No	66.7%	4

Findings and Recommendations: Of the six participants who have a migration plan, only one-third have access to the plan. It is recommended that those agencies that have migration plans make them available to employees of the agency and that these plans be verified for accuracy and audited on a periodic basis.

13. Who maintains the migration plan? (Name, phone and email please)		
Answer Options	**Response Percent**	**Response Count**
Name:	100.0%	4
Email Address:	100.0%	4
Phone Number:	100.0%	4

Findings and Recommendations: Two-thirds of the participants with a migration plan know who to contact in regards to the migration strategy. It is recommended that these individuals be contacted and copies of their migration plans be obtained by the State Archives for further analysis and possible integration into a model migration plan for use by other agencies.

Digital Imaging:

14. Does your agency convert paper documents to electronic images?		
Answer Options	**Response Percent**	**Response Count**
Yes	72.4%	55
No	22.4%	17
Don't Know	5.3%	4

Findings and recommendations: A higher than expected percentage (72%) of agencies are converting paper records into a digital format. It is recommended that further research be conducted into the tools, methods and procedures that the agencies are using in order to determine the accuracy and integrity of the imaging process.

15. Is the document scanning done in-house or by a vendor?		
Answer Options	**Response Percent**	**Response Count**
In-house	76.4%	42
Vendor	5.5%	3
Both	18.2%	10

Findings and Recommendations: A relatively low percentage (6%) of imaging is being conducted outside of the agency. It is recommended that the State Archives examines the imaging contracts of the agencies utilizing outside vendors to develop a model contract that provides controls to ensure the accuracy and integrity of the images as well as the state's ownership of the product, both images and indexes.

16. What is the approximate volume of records already scanned? Example: 1000 pages, 10 cubic foot boxes, 8 rolls of microfilm	
Answer Options	**Response Count**
	44

Findings and Recommendations: Volumes produced annually range from hundreds of pages to millions of pages. Five participants are producing more than a million pages per year. It is recommended that further research be conducted into the large scale imaging programs to ensure the accuracy and integrity of the images being created. It is further recommended that all agencies engaged in digital imaging have written procedure manuals and that the process be routinely audited.

17. What is the approximate volume of records waiting to be scanned? Example: 1000 pages, 10 cubic foot boxes, 8 rolls of microfilm	
Answer Options	**Response Count**
	40

Findings and Recommendations: With the large amount of identified materials waiting to be imaged, it is recommended that the State Archives investigate the possibility of centralizing the imaging operations for the purposes of reduction in per scan cost and/or an increase in accuracy and integrity through economy of scale.

18. What is the approximate date range of scanned records maintained at your agency? Example: 1992-94, 1996-2010	
Answer Options	**Response Count**
	45

Findings and Recommendations: Many of the digital images being produced by state agencies are from documents dating to the territorial period. Due to the potentially historical nature of the documents being imaged, it is recommended that the State Archives investigate these imaging operations to explore the possibility of a partnership for the maintenance and preservation of the images and/or historical material.

19. Why did your agency decide to scan/image your records? (Check all that apply)		
Answer Options	**Response Percent**	**Response Count**
Save Space	80.8%	42
Faster Retrieval	88.5%	46
Better Management	75.0%	39
Multiple Access	78.8%	41
Security Backup	55.8%	29
Other (please specify)		13

Findings and Recommendations: As four of the reasons offered for implementing a digital imaging project were selected by participants greater than 75% of the time, it can be surmised that many agencies are

expecting a multi-benefit return from their imaging investment. Accessibility is one of the major reasons given; in addition to multiple access and faster retrieval, six of the thirteen 'other' responses centered on increasing access to the records as well. It is recommended that the State Archives conduct further research into the actual versus perceived benefits from implementing a digital imaging system. In the event the realized benefits can be proven and quantified, this research can then be used by other projects to justify a digital imaging project for their paper records. Further research is also recommended in the area of security backup of paper records through digital imaging to ensure that, in the event the backup is needed, the records are accessible, accurate and have maintained their integrity.

20. Is the imaging:		
Answer Options	**Response Percent**	**Response Count**
A One-time Process	3.8%	2
An Ongoing Operation	80.8%	42
Periodic	15.4%	8
Other (please specify)		4

Findings and Recommendations: As a vast majority of the digital imaging projects (81%) are ongoing, it is recommended that the State Archives develop a 'digital imaging guidebook' that establishes the minimum standards for resolution, file format, auditing protocols and quality control procedures to be used in the imaging of long-term government records. Establishing benchmarks for imaging conducted in the state will allow agencies to create a litmus test for their own projects and help to ensure the accuracy and integrity of the digital images created.

21. After scanning, what happens to the original paper records? (Check all that apply)		
Answer Options	**Response Percent**	**Response Count**
Maintained In-house	85.7%	42
Sent Off-site	22.4%	11
Microfilmed	2.0%	1
Destroyed	40.8%	20
Other (please specify)		2

Findings and Recommendations: Eighty-one percent of the participants listed saving space as a reason for digitally imaging paper records; yet 86% still maintain the paper records and 41% destroy the paper records after imaging -- leading to the conclusion that some agencies are doing both. It is recommended that the

State Archives explore drafting legislation that requires agencies to adhere to a minimum level of standards when imaging paper records, and only those agencies that can demonstrate the accuracy and integrity of the processes used in their digital imaging process be allowed to destroy the paper records.

EDMS:

22. Does your agency use an electronic document management system (EDMS)? An EDMS is any system that centralizes and manages electronic documents.		
Answer Options	**Response Percent**	**Response Count**
Yes	34.2%	26
No	65.8%	50

Findings and Recommendations: Over one-third of participants (34%) use some type of EDMS. As such, these agencies are recommended as potential pilot partnership agencies with the digital archives. Automated EDMS integration will allow for seamless flow of records from creator to preserver.

23. What is the name of your EDMS system?	
Answer Options	**Response Count**
	23

Findings and Recommendations: Several of the larger EDMS vendors are represented in the list of installed systems (FileNet, Stellant, LaserFiche) as well as some that are unfamiliar to the digital archives team. It is recommended that the State Archives investigate what EDMS systems are in use and whether or not they are DoD 5015.2 compliant. It is further recommended that the State Archives draft legislation that would require Archives input into the procurement or development of EDMS systems storing public records of long term value. Such legislation will ensure appropriate protocols are in place to preserve the authenticity and integrity of the records stored in the EDMS system.

24. Who is the Vendor of the system?	
Answer Options	**Response Count**
	19

Findings and Recommendations: The major EDMS vendors are represented, along with a number of smaller vendors of unknown provenance. It is recommended that further research be conducted to determine which, if any, of the EDMS systems installed in state government are proprietary systems supported by a single entity. These systems should be labeled 'at-risk' and additional precautions taken to ensure that the records are not locked into a systems that may be unsupported if a vendor goes out of business.

25. Does your agency EDMS store born digital records? Born digital records are those digital records that never existed in a paper or analog format.		
Answer Options	**Response Percent**	**Response Count**
Yes	52.2%	12
No	47.8%	11

Findings and Recommendations: The close to half of the total number of participants (48%) that responded they stored non-digital records in the EDMS is in line with the large number of digital imaging systems in the state. Additionally, with 52% of the records in the EDMSs being born digital, it is recommended that disaster recovery guidelines be developed to ensure that the records in an EDMS are maintained and preserved in an authentic manner to maintain their evidential value.

26. What types of digital records are stored in your EDMS system? (Check all that apply)		
Answer Options	**Response Percent**	**Response Count**
Email	24.0%	6
Office Documents (word processing, spreadsheets, etc.)	56.0%	14
Images	44.0%	11
Database outputs/reports	40.0%	10
Scanned Images	88.0%	22
Other (please specify)		3

Findings and Recommendations: It is encouraging that nearly one-quarter (24%) of participants with EDMSs manage their emails through the system. With 88% of the EDMSs managing scanned images, this again speaks to the large number of digital images projects within the state. Further investigation is recommended into the methods and practices of those agencies managing database outputs/reports in their EDMS and whether these outputs adequately capture the records stored in the database system.

27. Does your agency collect any additional metadata to help identify and manage the records in the EDMS?		
Answer Options	**Response Percent**	**Response Count**
Yes	63.6%	14
No	36.4%	8

Findings and Recommendations: Over one-third (36%) of agencies add no additional metadata to the records within their EDMS. It is recommended that further research be conducted into how agencies manage and locate records without additional metadata elements. The metadata schemes used in all agencies must ensure that records in the EDMS can be maintained and preserved as long as required by the approved retention schedules.

28. What additional metadata is collected? (Check all that apply)		
Answer Options	**Response Percent**	**Response Count**
Agency Name	50.0%	6
Title of Record	75.0%	9
Disposition of Record	33.3%	4
Destruction Date of Record	0.0%	0
Access Restrictions	33.3%	4
Case Number	58.3%	7
Document Number	66.7%	8
Other (please specify)		6

Findings and Recommendations: Only one-third of participants add the disposition code to the EDMS, and none add the disposition date. Given that, it is recommended that further research be conducted into how the records are disposed of in the EDMSs, or if they are kept in the systems indefinitely.

General Records:

29. Have you ever attended records management training of any kind?		
Answer Options	**Response Percent**	**Response Count**
Yes	31.6%	24
No	68.4%	52

Findings and Recommendations: Only one-third of participants have attended records management training of any kind. It is recommended that Hawaii State Archives offer more records management training and education through online or virtual means in order to reach the public employees spread all across the islands.

30. When did you last attend records management training?		
Answer Options	**Response Percent**	**Response Count**
Last Six Months	25.0%	6
Last Year	8.3%	2
Last 1-3 years	16.7%	4
More Than Three Years Ago	50.0%	12

Findings and Recommendations: Half the participants have not attended any records management training in the past three years, while one-quarter has had training in the past six months -- pointing to the need to continue and expand the current records management training of public employees. It is recommended that the State Archives provide more training and education through online or virtual means in order reach the public employees spread all across the islands.

31. Do you understand the difference between what qualifies as a business record and what does not?		
Answer Options	**Response Percent**	**Response Count**
Yes	34.2%	26
No	11.8%	9
Not Really	53.9%	41

Findings and Recommendations: Almost two-thirds (65%) of participants are unclear of the legal definition of a business record. Without a firm understanding of what qualifies as a record and what does not, records custodians do not have the necessary core tool set to determine which retention schedule is appropriate for the records in their care. It is recommended that a basic training curriculum be introduced to ALL state employees upon hiring, and refreshed on a periodic basis.

32. What problems, if any, are you encountering in the creation, storage, retrieval, privacy, destruction, maintenance or management of digital records?	
Answer Options	**Response Count**
	49

Findings and Recommendations: Nearly two-thirds of the participants (64%) responded to this question by sharing problems they are aware of or encountering on a daily basis. An overwhelming theme is the lack of knowledge – specifically, they are concerned about what to do with records currently in their custody and what are they going to do with them in the future to ensure long term maintenance and retrieval.

"A majority of our documents are confidential (by law). Storage and future retrieval will be a problem."

"each employees archives own email without any guidance as how to clean up and keep only business related emails"

"not sure how to retrieve information if digitized"

"No Cataloging Standards"

"No Guidelines (or no commonly understood guidelines) for what should be stored"

"Pushback from IT - using too much storage space on servers"

"Because we are not specifically managing them yet, we don't know what problems we have."

"We are concerned about the long-term and future unintended negative consequences of moving into a digital archive with digital born documents without any paper back-up documents."

"Because the legislature is primarily an information-based entity, from which our primary product is state law, our documents are relied upon by many local and national entities far beyond the initial development of the law."

"We want to ensure that our records accurately reflect the actions and intent of the legislature at the various points in the legislative process."

"Sufficient space and poor cooperation from ICSD"

"With changing technology, state department cannot keep up."

"We need a way to backup and archive our electronic records in a secure manner so that we can discard the physical records."

"A central storage device which enables retrieval and backup would is preferred."

"Designation of resources to manage and maintain digital records."

"identification of records, personnel knowledge of recordkeeping and retention of what should or should not be imaged, and written procedures."

" Searchability of stored records"

"Destruction: a high potential for such of documents stored on desk PC and no back-up"

"NOT AWARE OF AGENCY POLICY, SOFTWARE, GUIDANCE, OR PLAN."

"Records are scanned without consideration for identification when retrieval is necessary. There are quality control issues. I'm not aware of any departmental procedures."

"We have information on floppy disks, 5 1/4 disks, and zip disks and most current computer do not even have zip drives. I've brought in a 5 1/4 drive from home (my husband's) to try to run disks."

Based on the feedback received, it is recommended that the digital archives project make it a priority to continue to educate public employees of their obligations to manage digital records for the life of the appropriate retention schedule. Additionally, guidelines and recommendations for the care and handling of digital records need to be developed and widely circulated. The development of a digital archives is essential to ensure that those records of enduring legal, fiscal or historical value are kept for as long as they are needed.

33. Would guidelines and/or standards for the management of digital records be helpful?		
Answer Options	**Response Percent**	**Response Count**
Yes	96.0%	72
No	4.0%	3

Findings and Recommendations: It is recommended that the State Archives continues to provide guidance and leadership in the creation, maintenance and preservation of digital records, and to draft model guidelines for the management of digital records.

34. Would you attend training on managing digital records if it was offered?		
Answer Options	**Response Percent**	**Response Count**
Yes	90.7%	68
No	9.3%	7

Findings and Recommendations: There is a strong desire (90%) for records management training among the participants. It is recommended that the State Archives continues its records management training and expand the content to include a stronger emphasis on digital records issues.

35. Would you be willing to share your needs and experiences with a focus group developing the Hawai'i State Digital Archives requirements?		
Answer Options	**Response Percent**	**Response Count**
Yes	56.0%	42
No	44.0%	33

Findings and Recommendations: Over half the participants (56%) are willing to participate in focus group sessions. It is recommended that the State Archives continue to involve interested stakeholders in the development of the digital archives project.

36. Please provide your contact information so we may send you details about our focus groups.		
Answer Options	**Response Percent**	**Response Count**
Name:	100.0%	39
Agency:	97.4%	38
Email Address:	97.4%	38
Phone Number:	97.4%	38

Findings and Recommendations: Of the forty-two participants who expressed interest in the focus groups, thirty-eight provided contact information. It is recommended that this contact information be kept on file to notify interested stakeholders of future developments, as well as to solicit their feedback on requirements and design.

37. Please share with us any other comments you have regarding records and their management.	
Answer Options	**Response Count**
	24

Findings and Recommendations: Below is a sample of the responses received. As can be seen by the feedback provided, the overall consensus is that the digital archives project is timely and needed to provide direction and solutions to digital records issues they are encountering. It is recommended that the

stakeholders continue to be regularly engaged and updated on the progress of the project, and their feedback be solicited and integrated into the functional requirements.

"Thank you for providing the archives seminar. It was very enlightening and thought provoking."
"Each ASO should be leading this initiative within the departments, but they are singularly uninterested."

"Possibly IT personnel, decision makers, etc. should be involved in these focus groups."

"Mahalo for including the Hawai'i State Senate. We are hopeful that we can work together on this very important project."

"I work in the State Procurement Office and over the past few years we have moved significantly towards the paperless environment. We are therefore appreciative that the Archives Office is focusing on the management and storage of electronic records."

"Suggest designated repository(ies) of state and state-related records that has knowledge of where these state and state-related records are maintained and are accountable of them."

"I realize this is an important subject but our institution has no awareness of this. I cannot justify spending a lot of time on this when my priorities are dictated elsewhere."

"The Hawai'i Digital Archive technology should also be shared as a sub-system for managing the active records within the Department. As a State, we do not have an enterprise license for Document Management System. This sub-system would help defray cost and bring standards across the State."

"I attended the presentation but am not "the" designated representative of this entire agency. I recognized the need for digital records management policies because I have worked on large cases involving digital records management."

"I'm really glad the State Archives is taking the lead on electronic records."

Appendix C: Functional Requirements

Definitions

AuERMSspec: *Functional Specifications for electronic Records Management Systems Software released by the National Archives of Australia, ISBN 1-92080734-9, February 2006.*

Core Requirement: A requirement that is essential to the proper collection, processing, maintenance and preservation of trustworthy records based on archives staffing and infrastructure.

Digital Component (or component): A digital object that is part of one or more digital documents, and the metadata necessary to order, structure or manifest its content and form, requiring a given preservation action.

Digital Object (or object): A discrete aggregation of one or more bit streams and the metadata about the properties of the object.

DoD: Department of Defense Electronic Records Management Software Applications Design Criteria Standard 5015.02

Highly Desirable Requirement: A requirement that adds additional functionality to either the archives staff, contributing agency or researcher.

Ingestion: The process of transferring records from a producer, using services and functions to prepare the transferred records for storage, verifying the existence and accuracy of the appropriate and accepting custody of the records by insertion into the digital archives.

Metadata: Information that characterizes another information resource, especially for purposes of documenting, describing, preserving or managing that resource.

Necessary Requirement: A requirement that, while not essential, is important to the proper operation of a trustworthy repository. Omission of necessary requirements will result in significant additional work on the part of the staff to demonstrate the integrity of the digital archive.

OAIS: Reference Model for an Open Archival Information System (OAIS) Standard CCSDS 650.0-B-1

Producer: The role played by those persons or client systems that provide the information to be preserved.

Record: A collection of digital components (i.e. files) and descriptive information (i.e. metadata) that when combined together provide a complete object that conveys meaning and context to the action for which it was created.

Render: To represent a digital object in a human-interpretable way.

Repository: *The whole of the hardware and software that is designed for permanent preservation of digital records*

TRAC: *Trusted Repositories: Audit and Checklist by OCLC and NARA, February 2007.*

UPitt: *University of Pittsburg project on the preservation of electronic records, known as the Pitt Project, "Functional Requirements for Evidence in Recordkeeping"*

Use Copy: *Versions of digital components that are viewed and/or downloaded by researchers.*

WADA: *Washington State Digital Archives Functional Requirements*

Web-Friendly: *Human readable versions of digital components that are typically smaller in file size and of open file formats created in order to allow for ease of download and access by researchers.*

Functional Requirements

A. Access and Usage Restrictions

The system must control who has access to what aspects of the repository as directed by the system administrator; therefore, the system will:

Ref #	Recommendation	Function/Feature
A.1	Core	Indicate on the record if an access restriction exists
A.2	Necessary	Provide sufficient detail on any access restrictions so that restrictions be interpreted by security protocols to limit access to only those individuals authorized
A.3	Necessary	Track for each restriction: reason for restriction, date restriction takes effect, date restriction ends, organization requesting restriction (DoD C4.1.1)
A.4	Highly Desirable	Provide a comprehensive audit trail of any changes made to access restrictions
A.5	Necessary	Indicate which metadata/indexing fields/attachments are restricted from public view
A.6	Highly Desirable	For each indication of field/attachment restriction, provide the governing policy, rule, regulation or law that authorizes the restriction
A.7	Highly Desirable	Provide in a prominent way to the researcher a detailed description of any use restrictions that exist on the record

B. Architecture Design

The design of the system, both physical and logical, must support the long-term access to *trustworthy* digital records through time and space; therefore, the system will:

Ref #	Recommendation	**Function/Feature**
B.1	Core	Utilize three-tier architecture
B.2	Core	Be built predominantly with open-source software
B.3	Highly Desirable	Rigorously utilize in-line code documentation
B.4	Highly Desirable	Comply with coding and database standards as established by the State of Hawai'i
B.5	Necessary	Will be able to scale on demand, for both storage and server utilization
B.6	Necessary	Modularize functionality and tools used in order to allow for replacement, addition or subtraction of individual modules or tools as needed without the need to rewrite an entire tier (i.e. microServices)
B.7	Highly Desirable	Support server clustering
B.8	Highly Desirable	Support 100 concurrent, interactive users, with multiple processes in a networked environment without a slowdown of more than 10% over what is experienced with 10 concurrent users (WADA)
B.9	Core	Utilize a web interface that allows researchers to search for, select, view and print records
B.10	Highly Desirable	Use a webserver that is separate (either physically or virtually) from the database server
B.11	Highly Desirable	Have HTTPS, SSL or equivalent support for web applications (WADA)
B.12	Necessary	Operate from inside a firewall and support a web server(s) that reside between firewalls (WADA)
B.13	Necessary	Allow the web interface to comply with branding standards established by the State of Hawai'i
B.14	Highly Desirable	Utilize a web browser that is fully functional with no browser 'plugins' required (WADA)
B.15	Highly Desirable	Maintain a data dictionary for the repository that conforms to the Data Entity Dictionary Abstract Standard ISO 22643:2002 or similar [OAIS 3.2.2.1.2 F-10]
B.16	Core	Store years in a 4 digit year. Leap year calculations shall be accommodated (DoD C2.1.2)
B.17	Highly Desirable	be designed with redundant components through that system allowing for 'a-channel'--'b-channel' architecture, allowing either channel to be brought off-line for maintenance/upgrades without the need to take down the entire system.
B.18	Highly Desirable	Support standard e-Commerce operations (add, view, remove, change quantities) that would allow researchers to purchase (certified) reproductions of records (WADA)

B.19	Highly Desirable	Allow for caching of popular searches/results/records to the depth and/or number specified by the repository archivist (WADA)
B.20	Highly Desirable	Utilize virtual machines at all three layers of the architecture
B.21	Highly Desirable	Provide a dashboard that allows the system administrator to monitor the overall health of the system, as well as the health of individual servers from a single screen

C. Digital Objects and Rendering

As all digital records are comprised of one or many digital objects, it is essential that all of the digital objects necessary to render a digital record are maintained and available for the life of the record; therefore, the system will:

Ref #		Function/Feature
C.1	Core	Provide the capability to support multiple versions of a record (DoD C2.2.3.19, WADA)
C.2	Core	Link each version of a record to its predecessor (WADA)
C.3	Necessary	Generate web-friendly versions of digital components for all records in the system where such formats exist
C.4	Highly Desirable	Store the attachments of emails as separate records and provide an indelible link between the email and the attachments.
C.5	Core	Be able to convert file formats ingested into the repository into non-proprietary formats according to a file-type crosswalk maintained within the system. (WADA)
C.6	Highly Desirable	Identify the file format and file version of each digital object within the system
C.7	Highly Desirable	Provide an indication of the hardware and software dependencies required to render the digital objects
C.8	Core	Provide sufficient workflows and tools to normalize data into standard formats
C.9	Highly Desirable	Have the ability to place watermarks on use copies of all documents, photos, videos, etc.
C.10	Highly Desirable	Provide the ability to redact text from records within the repository without altering the original record

D. Ingestion and Normalization

In order to manage the contents of the repository over the long term, it is essential that records are completed when transferred and are stored in a standardized way; therefore, the system will:

Ref #		Function/Feature
D.1	Core	Track an incoming group of records as an accession with a unique identifier that is associated with each individual component of the accession
D.2	Highly Desirable	Provide a full report of records transferred, the accession to which they have been assigned, and confirmation of which items have been formally accepted into custody (TRAC B1.7)
D.3	Core	Accept records in both structured and unstructured formats
D.4	Highly Desirable	Compare the hash of the incoming digital components against those already existing in the repository, note any identical objects, notify repository archivist and move potential duplicate records into holding
D.5	Highly Desirable	Check the structure of the incoming records for conformity to expected format agreed to with the producer; when discrepancies are detected, log those instances of non-conformance, send detailed report to repository archivist and move offenders to holding area (OAIS 3.2.2.4 F-21)
D.6	Highly Desirable	Identify if any components of the record transferred to the archives are encrypted prior to ingestion into the repository, notify originator and repository archivist, and move the affected record into a holding area without ingesting into the system
D.7	Necessary	Provide sufficient workflows and tools to normalize textual data into standard structure, formats, and semantics
D.8	Highly Desirable	Provide the ability to spider designated web-sites at a given frequency
D.9	Highly Desirable	Provide the capability to directly file audit data as a record (DoD C2.2.8.4)
D.10	Highly Desirable	Note the language of the record
D.10	Core	Ensure that digital objects can be ingested in their native format, regardless of format and technical characteristics, in order to be described and stored as records within the system (AuERMSspec A.2.1)
D.10	Core	For records comprised of more than one component, track all incoming components as a single record and maintain the association between the various components (AuERMSspec A.2.2)

D.11	Core	Support the bulk ingestion of disparate records from agencies through use of batch files, customize rules, validation and queries (AuERMSspec A.2.16)
D.12	Necessary	Support the transfer of records through a secure webpage
D.13	Highly Desirable	System supports registration of all incoming records to track existence, provide auditing and evidence of presence after destruction (AuERMSspec A.2.25)
D.14	Necessary	System will track authentication means used to authorize the ingestion of records from a trusted source (AuERMSspec C.1.13)

E. Integrity of the Repository

The maintenance of *trustworthy* records requires that the integrity of the records be established at point of transfer and maintained throughout the life of the records; therefore, the system will:

Ref #		Function/Feature
E.1	Necessary	Quarantine and then virus scan on a customizable schedule, all digital components transferred to the repository prior to full ingestion
E.2	Core	Provide a means of verifying the integrity of the transfer by hashing files on the originator's end prior to transfer, verifying hash on the receiver's end
E.3	Core	Enforce data integrity, referential integrity and relational integrity of the database (DoD C2.2.3.23)
E.4	Highly Desirable	Report to both originator and receiver the results of the Hash comparison. Any discrepancies will be moved into a holding area. (WADA)
E.5	Highly Desirable	Track the version of the ingestion routine used to normalize the records into the repository.
E.6	Core	Maintain a hash for each digital component ingested into the system
E.7	Highly Desirable	Run periodic hash checks to insure the integrity of the system and provide report of the results of the hash check, specifically noting any discrepancies
E.8	Highly Desirable	Indicate on each record which version of migration software was used for each version of the record.

E.9	Necessary	Provide sufficient auditing to facilitate reconstruction, review, and examination of events surrounding or leading to mishandling of records, possible compromise of sensitive information or denial of service. (DoD C2.2.8.3.2)
E.10	Highly Desirable	Trap errors encountered in the operation of the repository and maintain them in an error log
E.11	Necessary	Ensure that all digital objects referenced in the records as part of the accession are transferred – including external references within document objects (such as images, fonts, etc.) (TRAC B1.5)
E.12	Highly Desirable	Maintain a change log of all software changes to the repository (TRAC A3.6)
E.13	Highly Desirable	Accept into the repository the source code of each version of software used to ingest the records into the repository
E.14	Highly Desirable	Make a copy of all records transferred to the repository, as transferred by the producer, for storage outside the system as protection against system malfunction, corruption or loss

F. Interface Design and Searching

The system must provide a web interface that allows researchers to search, select and view non-restricted records contained in the repository; therefore the system will:

Ref #		Function/Feature
F.1	Highly Desirable	By default, sort returned search results in alphabetical order descending, but allow the researcher to select other metadata fields to be returned with the search results (WADA)
F.2	Highly Desirable	Allow the repository archivist to specify which metadata fields will be returned with the various types of canned, simple and advanced searches available to the researchers
F.3	Core	Adhere to ADA requirements for web accessibility
F.4	Highly Desirable	Provide a thumbnail of the digital components of the record along with all allowable XML data concerning the record when a researcher selects a specific record to view.
F.5	Highly Desirable	Maintain a full keyword index of all the metadata fields and text-based components stored in the repository for fast, efficient searching
F.6	Highly Desirable	Allow for both simple, broad based searches and focused, advanced searches on parameters established by the repository archivist

F.7	Highly Desirable	Allow the repository archivist to create searches with specified values and save those searches for global use ('canned searches')
F.8	Necessary	Utilize a webpage master template with drop down navigation bars controlled by a style sheet(s)
F.9	Core	Protect against SQL injection on all search field through validation of user response fields prior to execution of searches
F.10	Highly Desirable	Notify the researcher if zero results meet the criteria and provide a means of backtracking to revise the search criteria
F.11	Highly Desirable	Provide the ability to use soundex for name searches (WADA)
F.12	Highly Desirable	Provide a breadcrumb trail for researchers navigating on the website (WADA)
F.13	Highly Desirable	Have all the content, buttons, navigations user selectable for either English or Hawaiian
F.14	Highly Desirable	Support 'live chat' capabilities with a reference archivist during normal business hours
F.15	Highly Desirable	Contain a site map of the webpage to assist in researchers' navigation of the site
F.16	Highly Desirable	Maintain a list of common terms used within the project and classify them by domain field
F.17	Necessary	Contain a FAQ and/or a help button to provide direction, instruction, and tutorials on how to use the repository
F.18	Highly Desirable	Allow researchers the ability to search for null or undefined fields (DoD C2.2.6.8.6)
F.19	Highly Desirable	Allow the researcher to specify partial matches and designation of "wild card" fields or characters. (DoD C2.2.6.8.3)
F.20	Highly Desirable	Allow searches using Boolean and relational equators: "and", "and not", "or", "greater than" (>), "less than" (<), "equal to" (=), and "not equal to" (<>), and provide a mechanism to override the default order of preference. (DoD C2.2.6.8.4)
F.21	Highly Desirable	Provide a citation block for all records that varies according to the indication and type of use restrictions
F.22	Highly Desirable	Support casual browsing for content through the use of finding aids or other narrative high-level descriptions

G. Management Tools

Tools must be provided to the repository archivist that will allow for effective management of the repository; therefore, the system will:

Ref #		Function/Feature
G.1	Highly Desirable	Keep a parent child relationship tree for office -- organization reporting structure/administrative history
G.2	Highly Desirable	Audit the usage of individual records, accessions or any other chosen aggregation of records (such as by department, year, record series, etc.) specified by the repository archivist
G.3	Highly Desirable	Allow for ad-hoc reporting on any combination of audit logs or database fields specified by repository archivist (WADA)
G.4	Highly Desirable	Indicate for each record in the system, what disposition authority applies, the issuing organization, policy number, version, date of issue/effective, etc. (UPitt II.D.3)
G.5	Highly Desirable	Have the ability to attach images (or as the case may be copies of digital versions) of signed MOU/Submission agreements at the agency or accession level (OAIS 2.1.4)
G.6	Highly Desirable	Maintain copies and versions of any data-crosswalks developed as part of the MOU/Ingestion process within the repository
G.7	Highly Desirable	Provide the repository archivist with a dashboard that displays real-time status of the various modules used within the repository
G.8	Highly Desirable	Provide a secure login page for producer access that allows records transfers, contact updates, department access to restricted records, etc.
G.9	Highly Desirable	Maintain a list of each file type, the version, the modality (text, numeric, image, sound, video), the encoding scheme (ASCII, Unicode, etc.), and compression method within the system and the software required to view it along with an indicator to the location within the repository that a copy of the rendering software is stored
G.10	Highly Desirable	Provide an API to agencies to integrate the transfer of records from agency systems to Digital Archives (AuERMSspec A.2.20)
G.10	Highly Desirable	Allow records to be assigned to more than one disposition schedule (AuERMSspec A.2.37)

H. Metadata Creation and Handling

Long-term preservation of trustworthy records requires both descriptive and preservation metadata to be indelibly linked to the record; therefore, the system will:

Ref #		Function/Feature
H.1	Necessary	Have a flexible metadata assignment schema that can be modified by the repository archivist
H.2	Highly Desirable	Will track changes made to the metadata schemas in a change log
H.3	Necessary	Identify the IP Address and domain from which the record originated with sufficient specificity to identify the organization responsible (UPitt I.A.2)
H.4	Necessary	Uniquely identify the transfer of records with date, time and necessary sequence identifiers (UPitt I.A.3)
H.5	Core	Assign a unique computer-generated record identifier for each record and component of the record (DoD C2.2.3.5)
H.6	Highly Desirable	Not permit modification of the metadata fields indicated as not editable (DoD 2.2.3.8)
H.7	Highly Desirable	Track all the metadata standards and version used within the repository and provide a unique identifier to the specific version of the standard indelibly attached to the record
H.8	Highly Desirable	Associate terms used to describe or index the record by the producer with any the normalized fields created in the repository (UPitt I.B.2)
H.9	Highly Desirable	Capture, populate, and/or provide the producer with the capability to populate the metadata elements pertinent to the accession before transferring the records; and ensure that fields designated mandatory are in the proper format before transferring the record (DoD C2.2.3.10)
H.10	Core	Provide metadata for each component of a record in an accession and their relationship to other components, as well as the accession as a whole (Archival Bond)
H.11	Highly Desirable	Indicate on the record if any use restrictions exist (UPitt II.A.2)
H.12	Highly Desirable	For each email ingested into the system, capture both the intelligent name and actual email address of sender and recipients (DoD C2.2.4.2)
H.13	Highly Desirable	Allow for repository archivist defined metadata fields

H.14	Highly Desirable	Allow for metadata fields to be predefined as mandatory
H.15	Highly Desirable	Notify producer and repository archivist of any records that do not have the required metadata and move records into holding area
H.16	Highly Desirable	Allow for inheritance of metadata fields from parent objects (AuERMSspec A.1.63)
H.16	Highly Desirable	Allow system admins to override or amend metadata entered by authorized individuals (AuERMSspec A.1.65)
H.17	Highly Desirable	Capture transmission metadata information and associate with accession (AuERMSspec A.2.51)
H.18	Highly Desirable	Validate metadata associated with record against approved metadata schema (AuERMSspec A.2.57)

I. System Security

Records accepted into the custody of the repository must be protected from unauthorized alteration, addition or deletion; therefore, the system will:

Ref #		Function/Feature
I.1	Core	Allow only authorized individuals to create, edit and delete components and their identifiers (DoD C2.2.1.1)
I.2	Highly Desirable	Restrict individuals allowed to designate the metadata fields that are to be constrained to selection lists (DoD C2.2.1.2)
I.3	Highly Desirable	Ensure only authorized individuals are allowed to define and attach business rules and/or access logic to any metadata field including user-defined fields (DoD C2.2.1.5)
I.4	Core	Provide the capability for only authorized individuals to modify the metadata of stored records; and not allow the editing of metadata fields that have been specifically identified as not editable. (DoD C2.2.3.22)
I.5	Highly Desirable	In conjunction with its operating environment, shall not allow audit logs to be edited
I.6	Highly Desirable	Support pass-through authentication from an archives maintained directory, for access to system resources, including servers, applications, firewalls, databases, storage devices, etc. (WADA)
I.7	Highly Desirable	Allow for individual, unique log-ins rather than departmental IDs (WADA
I.8	core	Allow for system log-in with non-display character passwords (WADA)

I.9	Highly Desirable	Allow for system security to be assigned at group levels with varying degrees of system access (WADA)
I.10	Core	Allow for the system administrator to establish security levels
I.11	Necessary	Log all security violations and send immediate alerts to key personnel selected by the repository archivist (WADA)
I.12	Core	Allow for administrator-defined and maintained application security, including security for application modules and transactions, as well as access levels controls on all applications -- such as read-only, end-user, and system administrator. (WADA)
I.13	Core	Prevent unauthorized access to the repository
I.14	Necessary	Support generic anonymous/read-only account for searching repository from a web-interface (WADA)
I.15	Core	Provide that following any system failure, the backup and recovery procedures will ensure data integrity by providing the capability to compile updates and provide the capability to rebuild from any backup copy through use of the backup copy and all subsequent system audit trails
I.16	Necessary	Maintain an audit history transaction table that will track all additions, modifications and deletion to any records and be unalterable, even by an administrator
I.17	Highly Desirable	Log all user log-ins and attempted log-ins.
I.18	Necessary	Create copies of records and their metadata that can be stored off-line and at separate location(s) to safeguard against loss due to system failure, operator error, natural disaster or willful destruction through export of all metadata pertaining to each record, along with a base64 version of all digital components, into an XML file. (WADA, DoD C2.2.9.2)
I.19	Highly Desirable	Any changes to security levels (even those done by an administrator), on records, applications or servers, will automatically generate an email notification to personnel selected by the repository archivist
I.20	Highly Desirable	Allow researchers to create personalized logins in order to sign up for specialized notifications when particular, identified types of records have new accessions and to save searches or records of interest
I.21	Necessary	Use a system generated PKI infrastructure to verify the identity of the computer transferring records to the repository prior to initiating transfer

I.22	Highly Desirable	Provide a workflow for authorization of any access restriction changes made to any record
I.23	Highly Desirable	Allow system admin to set parameters for account lockout in event of failed login attempts (AuERMSspec A.3.3)
I.24	Highly Desirable	Provide centralized management console for access and security controls for users, records and other entities (AuERMSspec A.3.7)
I.25	Highly Desirable	Ensure that all security protocols are inherited from parent to child unless overridden by authorized account (AuERMSspec A.3.33)
I.26	Core	Ensure that security protocols are pulled from the centralized list prior to each event (AuERMSspec A.3.35)
I.27	Core	Ensure that the more restrictive security requirement is enforced in the event of overlapping security protocols

J. Purging Records from the System

Rules, regulations, policies, procedures and legislation change, technological and human error occurs, and the need to keep records of historical, legal and fiscal value is periodically re-appraised necessitating the ability to remove records from the system; therefore, the system will:

Ref #		Function/Feature
J.1	Highly Desirable	Identify the governing policy, rule, regulation or law that authorized the purging of the record from the system, and the identity of who performed the authorized
J.2	Core	Remove all digital objects associated with a record when the record has been purged from the system so that they cannot be restored through OS, file system, or forensic means
J.3	Highly Desirable	Provide the ability to purge records from the system, while maintaining the metadata that was associated with the record to provide evidence of existence (InterPARES record profile concept)
J.4	Highly Desirable	Track the reason why the record was removed from the system
J.5	Core	Allow only authorized individual to purge records from the system
J.6	Highly Desirable	In the event of multiple disposition schedules, ensure that the longest retention is satisfied prior to the disposal of the record
J.7	Core	Allow only authorized individuals to make changes to the disposition schedules and disposition assigned to the record (AuERMSspec A.4.7)

J.8	Core	Updates to disposition schedule must be effective immediately and apply retroactively
J.9	Core	Dispositions must be calculated in real time and may not be artificially advanced (AuERMSspec A.4.31)
J.10	Highly desirable	System must provide a mechanism and support legal holds to suspend disposition on identified records, accessions, disposition schedules or contributors

Appendix D: Overview of the Digital Archiving Software Options Compared

About Alfresco

http://www.alfresco.com/

Alfresco is an open source Enterprise Content Management (ECM) system that manages all the content within an enterprise and provides the services and controls that manage this content. At the core of the Alfresco system is a repository supported by a server that persists content, metadata, associations, and full text indexes. Programming interfaces support multiple languages and protocols upon which developers can create custom applications and solutions. Out-of-the-box applications provide standard solutions such as document management, records management, and web content management.

As an entirely Java application, the Alfresco system runs on virtually any system that can run Java Enterprise Edition. At the core is the Spring platform, providing the ability to modularize functionality, such as versioning, security, and rules. Alfresco uses scripting to simplify adding new functionality and developing new programming interfaces. This portion of the architecture is known as *web scripts* and can be used for both data and presentation services. The lightweight architecture is easy to download, install, and deploy.

- **Alfresco applications**
  Alfresco applications are built on the content application server and rely on the server to persist, access, query, and manage content.
- **Content applications**
  Alfresco can be used for building most ECM applications. Aside from the major applications such as document, image, records, digital asset, and web content management, there are a number of specific applications and use cases that add value to the enterprise.
- **Alfresco web tier and Surf**
  Alfresco provides ECM capabilities as data services, user interfaces, and user applications. The user interface capabilities are provided by applications and application components using Alfresco's web tier, Surf, originally developed as a faster way to develop content applications using scripting and REST architecture. Alfresco contributed Surf as a project to the Spring community for use with other Spring web tier components, such as Spring MVC, Spring Webflow, and Grails.

- **Alfresco programming models**
  A number of programming models are available for building an application using the Alfresco content application server.
- **Guiding design principles**
  Alfresco's founding engineers designed a product to support modern ECM requirements based on principles that are still in use today.
- **Architecture**
  There are many ways to slice and deploy Alfresco, however most deployments follow a general pattern. Ultimately, Alfresco is used to implement ECM solutions, such as Document Management (DM), Web Content Management (WCM), Records Management (RM), and Digital Asset Management (DAM). Across those solutions may also be elements of collaboration and search.

About Archivematica

http://archivematica.org/wiki/index.php?title=Main_Page

Archivematica is a comprehensive digital preservation system. Archivematica uses a micro-services design pattern to provide an integrated suite of free and an open-source tool that allows users to process digital objects from ingest to access in compliance with the ISO-OAIS functional model.

Users monitor and control the micro-services via a web-based dashboard. Archivematica uses METS, PREMIS, Dublin Core and other best practice metadata standards. Archivematica implements media type preservation plans based on an analysis of the significant characteristics of file formats.

Archivematica is free and open source software. The software applications integrated into Archivematica are each released under their own open source license. These are checked for license compatibility before they are integrated into the project. A full list of applications with their respective license is available on the external software tools page.

Any new software code created for the Archivematica project is released under a GPL version 2 license. The source code is available at archivematica.googlecode.com. All the system documentation found on this wiki is released under a Creative Commons license.

Code contributions, bug reports, wiki documentation updates along with questions and feedback in the chat room and discussion list are strongly encouraged and welcomed.

Each Archivematica software release builds on the OAIS use cases and implements a growing set of media type preservation plans. These are based on an analysis of open standards, open source normalization tools, and the significant characteristics of specific media types.

About DSpace

http://www.dspace.org/

DSpace is the software of choice for academic, non-profit, and commercial organizations building open digital repositories. It is free, easy to install "out of the box", and completely customizable to fit the needs of any organization.

DSpace preserves and enables easy and open access to all types of digital content including text, images, moving images, mpegs, and data sets. With an ever-growing community of developers, committed to continuously expanding and improving the software, each DSpace installation benefits from the next.

DSpace is an out-of-the-box open-source repository software package for creating repositories focused on delivering digital content to end users, and providing a full set of tools for managing and preserving content within the application. DSpace is the most widely used repository software platform, with over 700 installations worldwide representing a growing and active user community.

The DSpace application can recognize and manage a large number of file format and mime types. Some of the most common formats currently managed within the DSpace environment are PDF, Word, JPEG, MPEG, TIFF files. It is worth noting that, although out-of-the-box DSpace only auto-recognizes common file formats, files of any format can be managed by DSpace. DSpace also provides a simple file format registry where you can register any unrecognized format, so that it can be identified in the future.

DSpace is a set of cooperating Java web applications and utility programs that maintain an asset store and an associated metadata store. The web applications provide interfaces for administration, deposit, ingest, search, and access. The asset store is maintained on a file system, or similar storage system. The metadata, including access and configuration information, is stored in a relational database.

About OCLC Digital Archives

http://www.oclc.org/digitalarchive/

The Digital Archive provides a secure storage environment for you to easily manage and monitor the health of your master files and digital originals.

It provides a foundation for digital preservation of all your digital collections. It stores master files and digital originals in a secure, managed and separate environment whether digital collections are built using CONTENTdm or another local access repository. The Digital Archive provides tiered pricing to grow as the digital collections grow.

Secure, managed storage

The Digital Archive is a specially designed system in a controlled operating environment dedicated to the ongoing managed storage of digital content. OCLC developed specific systems process and procedures for the Digital Archive tuned to the management of data for the long term.

Automated monitoring and reports

From the time the content arrives, OCLC systems begin inspecting it to ensure that what is sent is what is retrieved in the future. On the day content is ingested to the Digital Archive, OCLC systems perform quality checks and record the results in a "health record" for each file. Automated systems revisit these quality checks periodically generating up to date reports on the health of the collection. OCLC provides monthly updated information for all collections in a personal archive report portal.

Simple, straightforward workflows

The Digital Archive provides a cost-effective, managed storage environment for digital master files that fits in with the several workflows for acquiring digital content.

- For users of CONTENTdm (either hosted or direct), the Digital Archive is an optional capability integrated with the various workflows for building collections. Master files are secured for ingest to the Archive using the CONTENTdm Project Client, the Connexion digital import capability and the Web Harvesting service.

- For users of other content management systems, the Digital Archive provides a low-overhead mechanism for safely storing master files.

About Washington State Digital Archives

http://www.digitalarchives.wa.gov/

The Washington State Digital Archives is the nation's first archives dedicated specifically to the preservation of electronic records from both State and Local agencies that have permanent legal, fiscal or historical value. Located in Cheney, WA on the Eastern Washington University campus, the facility was designed from the ground up to host this technically complex program. The web interface and database storehouse were custom designed specifically for the Digital Archives based on the OAIS model to hold the unique and very important electronic records found throughout the state, and to provide simple, straight forward access to researchers. This research report will describe the OAIS ISO standard and its various components, while interleaving descriptions of how the Digital Archives has taken this model and developed a useable solution using state of the art technology.

The OAIS model was written to provide a framework for any archive that has responsibility to preserve information that is contained in an electronic form and make it accessible over the long term to a designated community. The model itself does not distinguish between long term records and permanent, archival records; the difference between the two in technology terms in indistinguishable when viewed over a 40 year time frame, considering technology changes dramatically (generationally) every two to four years. So this model sets forth the archival requirements for preservation, common terminology and concepts that both IT and archivists need to share, and establishes the requirements of the system to be flexible and adaptive as technology changes over time.

Based on the information contained within the MOU, and a data sample representative of the data to be transmitted, the Digital Archives creates an orchestration that takes the given data, normalizes required fields (such as making the date fields consistent across all data) and adds additional metadata based on the agency of origin and record type. Some of this additional metadata can be assumed based on the point of origin. County Auditors send their entire Recorded Database and no other records, therefore, the records series is known and the contents of the data submission can be inferred. In addition, the security and sFTP information is also associated with the record, a unique accession number is assigned to the transmission, and date and time of receipt is also permanently associated with the record. By applying additional metadata based on point of origin and record type, Content Information is stored in the *record common* table in the database – the who, what, where, when, etc. Additional tables in the

database contain the original metadata sent with the SIP, a table for the images, another table for finding aids and a finally a table for the security information on what record series and fields are confidential.

Data Management in the Digital Archives is done within the same SQLServer database that the records themselves are stored in. This was done to keep ALL the information about a record together; when it came in, when it was processed, how it was process, who has accessed the record, how many times it has been ordered, etc. The intention was to create one indelible record about the record. Contained within the database are tables for security, user accounts, orders, and search tracking. These tables are not *part* of the original SIP, nor are they part of the AIP as it is stored in the database, rather these tables function as Package Descriptors and Content Information about the record, as well as being able to fulfill event based orders.

The security table tracks the access restrictions on any record brought into the system, and can restrict records at the record series, record column or individual record. This granularity was necessary based on the broad base of records to be accessioned into the system: all adoption records are sealed for 75 years, DD214 discharge papers cannot have the images or SSN displayed, and a judge may seal an individual case to protect a person's rights. In addition to being able to manage access to the data, the database can also *permit* access to any of the above records assuming they are a registered user from the originating agency. In order to serve its secondary purpose of providing business resumption capabilities to remote agencies, the Digital Archives has to permit unrestricted access to authenticated users from the originating agency, including access to publicly restricted information.

Orders within the system are tracked from the point of creation by the consumer all the way through fulfillment by the producer. At any point, a producing agency can create an ad hoc report for all outstanding orders to be fulfilled. As orders are fulfilled, the system generates order fulfillment notices that are sent to the project team, producer and the consumer. Along with order tracking, the Data Management database also tracks how searches are being conducted by the consumer: which search interface they are using, which record series they are searching on, what words they are using and which images are they viewing. The combination of order tracking and search tracking allow the Digital Archives development team to devote resources to improving access to those records that are most used.

Appendix E: Recommended Disaster Recovery Manual Elements

Section 1: Emergency Procedures and Contact Information

Staff Fire Call List:

Organization (Org) Chart (Chain of Command) w/ contact information for agency execs
Emergency contact information for staff during off duty hours

Procedures and Guides for handling different types of emergencies.

Natural Disaster
Power Disruptions
Extreme Weather Conditions
Building Issues
Water Leaks
HV AC monitoring, alerting, servicing, and failures
Physical Security (i.e. Doors, Locks, Windows)
Security Events
Physical Intrusion
Network Intrusion
Hostile People
Media Events (i.e. Reporters)

Section 2: Documentation

Diagrams of Network

- Network Architecture
- Racks Drawing
- Physical
- Logical
- Server Room Layout - Fiber and Power Path
- ISP Connection

Configuration of Equipment and Network

Network IP Address & Server Names
Server Build-Out Worksheets (one per server)
Public DNS Entries - spread sheet (ex: www.digitalarchives.or.gov = 65.12.234.99)
Network Firewall
- Show Run Config (configuration settings)
- Shun list

Ethernet Switches
- Show Run Config
- Excel - showing switch ports in color by vlan memberships

Fiber Channel Switches - Show Run
Directory Service Accounts
- System/Service Accounts
- User Accounts
- Security Groups & members & purpose of group
- Utility e-mail accounts & purpose and where used
- Mail Distribution Groups
- Group Policies

Storage Documentation
- Install Notes, Names, Revisions, contact, Host Collection, Storage Serial #, HA Serial #, Host Agent Software, Array Agent Software, LUN Layout, Service Tags.
- Storage Processors- configuration

Hardware Load Balancers - Show Run Config

Section 3: Backups

Clone Images

Policies (Freq. - Once per quarter & before a major system change or rollout)

Data Backup Software Information

Policies (Freq. Daily Incrementals, Weekly Fulls, Retention Schedule, location of tapes)

RAW Data: 2 tape copies of RAW Data

Raw data is original copies of data received from agencies.

Off site & on site copy

Fill uncompressed - do not span tapes.

Reasoning: Do not span tapes - due to the Average Mean Time Between Failure factor. For example: If hypothetically the tape has an AMTBF of 500 hours. Some tapes will last longer than that and some will fail before that time - the "weakest" tape would cause the whole series set of 5 spanned tapes to fail.

RA W Data printed copies: Print screens of folder directory showing:

Agency, Record Types, Year, Month, Day of transfer

Purpose: To help quickly identify which tape set has necessary records in the event we need to access the records in the future and do a restore from tape.

XML Storage: 2 tape copies of XML Deep Storage

Off site & on site copy

XML is an "open standards" (ASCII text file). Put a copy of each RAW metadata record into self describing XML file and include the digital object using the base 64 encoding scheme.

Purpose: To help keep the records in a "long term" "open format" state so that, if needed, the system could be rebuilt using the XML files.

System Backup

Policies (Freq. Which systems get backed up, how often, and retention schedule)

Full Backups

System Recovery Disk

System State

DataBase Backups

Policies (Freq. of backups, retention schedules)

Section 4: Software

Software

A Printed copy of a listing of all software used in the system

'Archival Quality' DVD copies of installation software CDs & DVDs for:

- Operating Systems
- Special Drivers (i.e. HBA, SCSI drivers, SAN masking software)
- Archives Software
- Database
- Service Packs
- Virus scanning software
- All backup software
- Any other Application Software used in the system

Purchase Orders

- Printed Copies
- Electronic Copies

Licenses

- Printed Copies
- Electronic copies of the licenses

Software Activation Keys

- Printed Copies
- Electronic copy - (A LicenseKey.txt file with software activation keys locating in each folder of software)

Section 5: Manuals

Printed copies and electronic versions on backup tape of:

Documentation:

- User Manuals
- Installation Guides
- Performance and Specification Guides
- Security and Best Practices Guides
- Disaster Recovery Guides

Equipment Documentation for:

- UPS system
- Backup Generator
- Building Electrical and Data Wiring Diagram & Blue Prints
- Ethernet Switches
- Fiber Channel Switches
- Firewall
- Servers
- Storage Network
- Tape Library
- VOIP system
- Security Cameras
- Security System
- Card Access System

Section 6: Service Contracts, Maintenance Agreements, and Warranty Information

Purchase Orders

- Printed Copies
- Electronic Copies

Vendor:

- Contract Information
- Scope of coverage (i.e. on-site, phone, 4hour, next day, parts only, labor)
- Serial numbers
- Expiration Dates
- Support Phone Numbers
- Support Site Login (i.e. URL, login name and password)

Systems:

- UPS system
- Backup Generator
- Ethernet Switches
- Fiber Channel Switches
- Firewall
- Servers
- Storage Area Network
- Tape Library
- VOIP system
- Security Cameras
- Security System
- Card Access System

Add-on Items: (Manufacture Warranties)

- Hard Drives
- RAM
- USB Pen Drives
- Etc ....

Historical Maintenance Log

- Summary of maintenance performed
- Failures
- Upgrades
- Bios/Firmware upgrades
- Unexplained Quirks and Observations