



THE FLORIDA STATE UNIVERSITY
COLLEGE OF COMMUNICATION & INFORMATION
School of Library & Information Studies

FLORIDA MEMORY PROGRAM LONG-RANGE PLAN: INTERIM REPORT

Paul F. Marty, Ph.D. <marty@fsu.edu>

Associate Professor, School of Library and Information Studies,
FSU College of Communication and Information

Nicole D. Alemanne, MS <nalemanne@fsu.edu>

Doctoral Student and Research Associate, Information Institute
FSU College of Communication and Information

Charles R. McClure, Ph.D. <cmcclore@lis.fsu.edu>

Director, Information Institute and Francis Eppes Professor
FSU College of Communication and Information

For

Jody Norman
State Library and Archives of Florida

March 8, 2010

FLORIDA MEMORY PROGRAM LONG-RANGE PLAN: INTERIM REPORT

Introduction

In December 2009, the School of Library & Information Studies (SLIS) and the Information Use Management & Policy Institute (Information Institute) of Florida State University began work on the *Florida Memory Program Long-Range Plan* to be completed on or before June 15, 2010. The purpose of the project is to produce a written report describing a long-range plan for the Florida Memory Program and detailing the necessary steps and potential solutions for the Program's continued growth and success over the next three years.

Project Description

The Florida Memory Project¹ provides online access to historical resources housed in the State Library and Archives of Florida. Currently more than 550,000 digitized records in the form of audio files, video clips, textual material, and photographs are available to website visitors. However, the project faces a number of issues that will make it increasingly difficult to provide service at the level necessary to fulfill its mandate. Specifically, collections are housed in a variety of different databases with software configurations that are dated, pose security issues, and have significant search limitations. In addition, there is a need to develop a more robust plan for the long-term storage of digitized master files. The long-term plan will recommend solutions that will allow the Florida Memory Program to continue to grow and operate successfully, including recommendations for improving the project's information architecture, interface design, and technology infrastructure.

The project team is conducting a series of evaluations and assessments to determine the products, technologies, standards, and services that will best meet the needs of the State Library and Archives of Florida as they work to develop the Florida Memory Program over the next three years. These activities comprise four specific tasks: needs assessment, systems analysis, comparative evaluations, and solutions identification. Drawing upon the outcomes of each of these activities, the project team will prepare a final report providing recommendations for strategic issues including growth, storage, software selection, hosting, and data migration.

The context in which this study is being conducted is one of limited resource availability throughout Florida government and at the State Library and Archives of Florida. The proposed plan will need to recognize these and other constraints to be successful. This interim report details project progress to date, and provides a preliminary needs assessment and initial set of recommendations based on that assessment.

¹ <http://www.floridamemory.com/>

Summary of Progress

Currently, the project team has completed the orientation phase and is nearing completion with the needs assessment and systems analysis phases. These phases are based on the following goals as described in the scope of work:

- *Orientation Phase.* The project will begin with a series of onsite visits to the State Library and Archives of Florida to orient the project team to the activities of the Florida Memory Program to ensure everyone is in agreement as to the purpose of the Florida Memory Program and the needs and expectations of the staff in terms of developing a long-range plan.
- *Needs Assessment.* The project team will conduct a needs assessment to determine the technology, information, and personnel needs for the successful growth of the Florida Memory Program over the next three years. This needs assessment will document the current situation and describe specific needs for future growth and long-range planning.
- *Systems Analysis.* The project team will perform an analysis of the systems and technologies currently in place at the Florida Memory Program, and document the State Library and Archives' current plans and expectations for the future growth, expansion, and use of these systems, including relevant hardware, software, and data records. The project team will also document the current volume of data stored in these systems and the expected growth of these data over the next three years, as well as current levels of data use and expected usage changes during this time period.

The project team is on schedule to complete the needs assessment phase by the end of March 2010 and the systems analysis phase by the end of April 2010.

Interim Needs Assessment and Recommendations

The Florida Memory Program offers a broad range of exceptional resources for residents in the State of Florida and beyond; the State Library and Archives have assembled and prepared an outstanding collection of data documenting the history of Florida and providing valuable services for the citizens of the state. Staff members at the State Library and Archives continue to work to improve the website and its collections, but have, on their own initiative, identified several hurdles that must be overcome and several changes that must be made if the project is to keep moving forward.

The challenges facing the Florida Memory Program stem from the way the project's exhibits and collections have been constructed over multiple years, using multiple technologies and multiple data standards, many of which are conflicting or non-standardized. In order to move the project forward and develop a website fulfilling the project's mandate and the staff's vision, current operations must be streamlined, consistent standards implemented, database systems developed across the various collections, and technology employed that better supports the work of the staff in managing the collections and the service of meeting the needs of the citizens of Florida.

A complete systems overhaul would unleash the staff's creativity and allow it to develop and deploy collections and exhibitions more quickly. This would help them better meet the needs of the general public, and enable both staff and users to find things faster and better without having to struggle with outdated technologies and systems.

In order to fulfill this vision, technology, growth, and storage issues must be addressed for three major areas: a) information architecture, b) interface design, and c) technical infrastructure.

The Department of Management Services is planning to implement an enterprise information technology (IT) management solution that will affect all Department of State IT operations. While this transition is in process and will have important implications for the long range planning of the Florida Memory Project, the study team is not addressing these transition activities in the interim needs assessment and recommendations.

Information Architecture

In order to develop a more robust information architecture for the Florida Memory Program, all of the existing collections data must be brought into a centralized system with common data standards throughout, including the use of standards for metadata, description, formatting, and controlled vocabularies and thesauri. Such a change would address at least two major usability issues with the current system: the lack of standardized tools for users as they browse and interact with multiple collections; and an inability to search across multiple collections.

To achieve this goal, the staff must first determine the standards they wish to employ, and then migrate all of their data from the current to the new standards. The methodology, time required, and costs involved when implementing such changes depend on the solution chosen. In addition, choices made when developing the back-end information architecture will affect the type of interface design solutions that are available (see below).

The project team has identified four different types of information architecture solutions, as well as the advantages and issues connected with each (Table 1):

Database Option	Advantages	Issues	Cost
Design 100% in house	Full flexibility to design to needs Scalability	Many staff or consultant hours to build In house support needed	Low cost / High personnel time
Develop in house using open source software	Free to download Flexibility/ease of use Possibly a unified database Availability of interface templates Availability of disaster recovery/migration utilities	Little outside support Possible scalability problems Possible sustainability issues	Low cost / Medium personnel time
Work with current supplier (e.g., Gale) to provide customized solution	Less staff time than in-house or open source solution Possible cost savings due to relationship	Possible sustainability issues if relationship with supplier changes Less flexible solution	Medium cost / Low personnel time
Purchase off-the-shelf solution from library, archives, or museum systems vendor	Full supplier support, including consultants to port databases to new package Stable company behind package	Most expensive solution in terms of fixed costs Least flexible solution Currently available packages may be designed for other uses	High cost / Low personnel time

Table 1: Database Solution Options, Advantages, Issues, and Cost

Each proposed solution offers different issues in terms of flexibility, time, and cost. Designing a solution completely in-house allows for the most flexibility, but would involve the greatest input of staff time and related costs. The use of open source software would involve significant staff time to modify the chosen system to meet the Florida Memory Program’s specific needs, but this would involve less time than that needed for the in-house solution. Working with a current supplier such as Gale would have the lowest cost and staff time requirements, but allows for much less flexibility. Finally, an off-the-shelf solution involves the least amount staff time and might offer the most integrated and up to date technology (including ease of ongoing system changes), but offers the least flexible, yet most expensive solution.

No matter which solution is chosen, someone on staff will need to analyze the metadata standards currently used for each database, decide what new standards should be applied, and crosswalk all pre-existing data to the new standards.

Issues that will need to be addressed before new databases for specific collections are developed and existing data migrated to new standards include:

1. Collections Analysis / Overview
 - a. Decide which aspects of the collections data will be migrated (e.g. records), and which will be left in separate systems (e.g. images)
 - b. Document the metadata standards currently in use across multiple systems by developing lists of current metadata schemas
 - c. Decide which collections and which records will be migrated to a common standard
2. Metadata Standards² Selection / Crosswalking
 - a. Select and/or develop metadata standards that meet user needs and can be applied to records across collections (e.g. Dublin Core³)
 - b. Determine how this standard can be extended to reflect unique collections
 - c. Prepare records to conform to chosen standards by identifying common fields in each database, and scraping the database-less databases (BroadSides, Physicians Journals & WPA Stories) to create new tables
 - d. Prepare to migrate data by developing crosswalks for each database

Only after this initial analysis is performed, can the actual work begin to migrate the data from the current databases to a newly developed or purchased system. To calculate the cost of each solution, the value of several variables must be obtained: the amount of data to be migrated; the number of staff hours it will take to install a new system, configure new databases, develop a data migration plan, and migrate the data; and the cost per staff hour, and software costs. With these values determined, the following formula can be used: *Cost of software package + (amount of data to be migrated * (number of staff hours * cost per staff hour))*.

Interface Design

Once existing data have been consolidated and standardized, it will be necessary to create a new interface / front end from which end-users will access the new data structures as part of the revised Florida Memory Program. The following table (Table 2) delineates advantages and issues related to the front end for each of the identified information architecture solutions.

² <http://www.niso.org/publications/press/UnderstandingMetadata.pdf>

³ <http://dublincore.org/>

Interface Option	Advantages	Issues	Cost
Design 100% in house	Full design flexibility Can leverage work already done	Need staff or consultants who are trained in development and design Will take staff time away from current projects	Low cost / High personnel time
Develop in house using open source software	Availability of interface templates Flexibility/ease of use	In-house staff time needed to work with design templates Little outside support Possible sustainability issues	Low cost / Medium personnel time
Work with current supplier (e.g., Gale) to provide customized solution	Supplier will help with design Less staff time than in-house or open source solution	Less flexibility in design and tools Less flexible solution	Medium cost / Low personnel time
Purchase off-the-shelf solution from library, archives, or museum systems vendor	Design is built into package Full supplier support	Little flexibility in design and tools Currently available packages may be designed for other uses	High cost / Low personnel time

Table 2: Interface Design Options, Advantages, Issues, and Cost

While it will certainly be possible to migrate large parts of the current interface, the implementation of a new information architecture will provide the opportunity to upgrade the existing interface, thereby allowing for:

- A standard “look and feel” across the entire project that also offers the flexibility for each collection and exhibition to have specific design elements;
- Standardized data value formats, thesauri, and metadata schemas that allow searches to be conducted across all collections databases;
- Flexible database structures that allow new fields to be created on the fly with new search tools can be customized for individual collections; and

- Improved accessibility features that can bring the Florida Memory Program into compliance with the Americans with Disability Act (ADA) and bring the Florida Memory Program to more users.

These are significant improvements that would enhance the Florida Memory Project and increase access to the collections.

The cost of improving the interface design will depend on the type of information architecture chosen. To calculate the cost, the value of two variables must be obtained: the number of staff hours that it will take to design and develop the new front end and the cost per staff hour. With these values determined, the following formula can be used: *Number of staff hours * cost per staff hour*.

Technical Infrastructure

The successful development of a new information architecture and new interface design for the Florida Memory Program will depend greatly on having an improved, and extremely stable, underlying technical infrastructure. The selection of the systems and solutions that comprise this infrastructure will be based on answers to questions such as:

- Where are the master digital files currently being stored, and how many are there?
- How many new records or items will be digitized over the next three years, and at what rate will they be created?
- Where are the derivative files stored, and how often are they backed up? How many derivative files currently exist?
- Where are the project's databases located, and who maintains them? Will any changes to the current situation change the cost structure, and how?
- Are data migration plans in place? How often will data will be migrated to new storage systems? Will new file formats need to be supported?

Decisions made about the underlying technical infrastructure will affect the production and storage of raw data, the storage of digital masters, scanning procedures, and the ways in which data are backed up and migrated from one system to another. Long-term storage costs are more than a one-time expense and involve ongoing costs to maintain services and pay for data center operations. The following issues must be addressed to estimate costs:

- Service costs will be affected by the chosen information architecture solution, as data center costs will be determined by the number of databases used. An open source software solution allowing multiple tables to be consolidated to a single database, for example, could lower the total storage expenses by reducing the number of databases that need to be maintained.
- The current hardware storage solution (Buffalo TeraStations⁴) is a stopgap measure, and new hardware will likely need to be purchased⁵ or new distributed storage technologies (e.g.

⁴ <http://www.buffalotech.com/products/network-storage/terastation/>

⁵ <http://lockss.stanford.edu/lockss/Home>

LOCKSS) explored. The ultimate hardware decision will be guided by answers to questions above such as the number of current master files, the number of new masters that will be created, and the number of copies of originals that will be stored.

A formula to calculate basic hardware needs is: (*Current rate of new master file creation * change in creation rate*) + (*number of servers * cost per server*).

Next Steps and Summary

In the next stage of the project, the project team will focus on completing the needs assessment and systems analysis phases of the project:

1. *Needs Assessment and System Identification and Evaluation.* Needs assessment and analysis activities will be concluded by the end of March 2010. This Phase of the project will provide the opportunity for ongoing clarification and re-assessment of these needs informed by the findings of the project team as they identify and evaluate potential systems and solutions that could meet the needs and goals of the Florida Memory Program
2. *Systems Identification and Evaluation Phase.* This phase will continue through April 2010. Activities for the phase include the identification, comparison, and evaluation of potential software and hardware solutions that could serve the long-term needs of the Florida Memory Program. Multiple systems and diverse solutions will be addressed for all program needs, including system growth and data migration plans, recommended software for current and planned system features, and recommended technology solutions for long-term storage and hosting of the Florida Memory Program; each of these solutions will be documented and addressed in the final report. This phase will also include the identification and documentation of related projects that could serve as exemplars for the future development of the Florida Memory Program.
3. *Reporting Phase.* The official findings of the project team will be documented in detail in a final report, outlining a long-range plan for the Florida Memory Program over the next three years. This report will be delivered to the Division of Library and Information Services no later than June 15, 2010.

The project team has made substantial progress and remains on schedule. This progress is due, in large part, to the assistance and involvement of Ms. Jody Norman and her staff in this project. In the first two months of the project, the team has completed the orientation phase and made substantial progress toward needs assessment and systems analysis. An interim needs assessment and initial set of recommendations have been developed based on that work.

Going forward, the project team will focus on completing the needs assessment and systems analysis phases of the project. A final report that documents the team's findings and outlines a long-range plan for the Florida Memory Program over the next three years will be delivered to the Division of Library and Information Services no later than June 15, 2010.