Document Title Revised Plan of Work

Document Number 0.1

Version Final

Date 09/22/03

NHPRC Grant Number 2004-083

1. Articulate functional requirements for active recordkeeping systems that will transfer records to an archival repository. Beginning with the functional requirements expressed by the Indiana University Electronic Records Project (Bantin 2002), the co-principal investigators will survey the recordkeeping requirements literature (from previous NHPRC-funded projects and national and international standards) and select those requirements that are most applicable to the university archives community. This activity is necessary for two reasons: (1) Before designing new systems or evaluating existing systems for active records management, it is necessary to define what types of functionality such a system should possess to support the transfer of records from it to an archival repository; and (2) "No one list [of functional requirements] has been endorsed by the [archival] profession." (Bantin 2002) **(Kevin Glick/Eliot Wilczek) Month 1**

2. Articulate functional requirements for archival repository preservation systems. Beginning with the preservation system functional requirements expressed by the PERM Project (*Functional* 2002), the co-principal investigators will survey the preservation system requirements literature and select those most applicable to the university archives community. These requirements are necessary in order to preserve electronic records and ensure their continued accessibility and authenticity over time in the preservation repository as the creating technologies become obsolete. **(Kevin Glick/Eliot Wilczek) Month 1**

3. Articulate the Ingest Electronic Records process. In order to accession electronic records, it is necessary to bring in the digital components of electronic records, the accompanying information related to the creation, preservation, and reproduction of those records, and transmittal information. It is also necessary that such transfers comply with terms and conditions of submission established by the archives, or that the archives can execute a preservation action plan to bring the records into compliance with those requirements. (*Findings* 2002) While this process has been described and modeled, most university archives have not yet been able to plan or carry out the transfer of electronic records accessions as described above. This project intends to help university archives by developing a step by step guide to the interaction between university archives and records

Kevin Glick
Yale University
kevin.glick@yale.edu
203.432.2202

Eliot Wilczek
Tufts University
eliot.wilczek@tufts.edu
617.627.4588

producers, by developing simple tools to facilitate the ingest process, and by undertaking the ingest process employing the Fedora system.

3-1. Develop a step-by-step guide to the interaction between university archives and records producers. We will create this guide according to the instructions for creating a producer-archive community standard conformant with the Producer-Archive Interface Methodology Abstract Standard. (*Producer* 2002) The Abstract Standard document explicitly describes the sequence of steps that an archives must follow to plan and carry out the transfer of information objects to the archives (these steps will not be listed here). We will create a guide that will conform to the requirements of a producer-archive community standard for the university archives community compliant with the Producer-Archive Interface Methodology Abstract Standard. The activities necessary to create the guide include:

3-1-1. Analyze each action of the producer-archive Abstract Standard. It is necessary to evaluate each phase and activity described in the Abstract Standard to determine its applicability to the university archives community. Each action will be accepted, eliminated, or changed to suit the community. **(Kevin Glick/Stephen Yearl/David Gerwirtz/Eliot Wilczek/Project Analyst/Greg Colati) Month 1–Month 6**

3-1-2. Define terminology. We will draft the guide in language that reflects university archives assumptions and practices. We will also create a glossary to clarify words with multiple or ambiguous meanings along with an equivalence table mapping the glossary terms to terminology from the Abstract Standard. **(Kevin Glick/Eliot Wilczek) Month 1–Month 6**

3-1-3. Create an information model of the community. We will define the community's main information objects and relationships through information modeling. **(Kevin Glick/Eliot Wilczek) Month 1–Month 6**

3-1-4. Reference related standards. We will note any standard relevant to this community and identify missing standards that need to be developed. **(Kevin Glick/Eliot Wilczek) Month 3–Month 6**

3-1-5. Identify tools. We will identify any tools that may or must be used to perform each of the actions described in the Abstract Standard (for example, an application to collect the necessary metadata in a form acceptable to the archives repository). **(Kevin Glick/Stephen Yearl/David Gerwirtz/Eliot Wilczek/Project Analyst /Robert Chavez/Greg Colati/Thorton Staples) Month 1–Month 6**

3-2. Develop simple tools to facilitate the transfer of electronic records from producer to archive. This will be an iterative process that will overlap with

creation of the step by step guide, the producer-archive projects, and the articulation of the maintain electronic records process. The Fedora Project Analyst and Gewirtz will be primarily responsible for this activity with assistance from Yearl and Chavez. **(Project Analyst/ David Gerwirtz/Stephen Yearl/Robert Chavez/Thorton Staples) Month 1–Month 9**

3-3. Undertake six producer-archive projects. Yale and Tufts will each undertake three (3) separate producer-archive projects, transferring each accession of electronic records into Fedora according to the guide described above. The producer-archive projects represent a variety of records types, records formats, and producer-archive relationships. Tufts will transfer: (1) administrative documents (correspondence, reports, financial spreadsheets) generated in desktop applications by the Administrative Office of the Arts and Sciences Library; (2) datasets and reports produced by the Office of Institutional Research; and (3) an electronic publication from the Alumni Office composed of a mix of text, graphics and images. Yale will transfer: (1) records documenting the meetings of the Yale Corporation, the university's governing body, that are organized in a Records Management Application; (2) administrative documents generated in desktop applications by a Library administrator who is no longer employed by Yale; and (3) CAD drawings of university buildings created by an independent contractor as part of a renovation project. All of these accessions have been appraised to have archival value. These different types of accessions will greatly enhance our understanding of the Ingest Electronic Records process. These projects will allow us to test and consider revisions to the guide according to real-world experiences. **(Kevin Glick/Stephen Yearl/David Gerwirtz/Eliot Wilczek/ Project Analyst/Robert Chavez/Greg Colati) Month 7–Month 10**

4. Articulate the Maintain Electronic Records process. In order to maintain electronic records it is necessary to maintain information about accessioned records, maintain the digital components of those records, and execute action plans for preserving them. (*Findings* 2002) This project will focus its efforts toward understanding the portion of electronic records preservation that is covered by the Data Management and Archival Storage process as defined in the OAIS Reference Model. (*Reference* 2002) This project hopes to apply this understanding to help university archives by developing a step-by-step guide to the Maintain Electronic Records process and by attempting to undertake this process employing the Fedora system. **Month 7–Month 18**

4-1. Draft a step-by-step guide for the Maintain Electronic Records process. This guide will be created in a similar fashion to our guide to interaction between university archives and records producers. However, there is no abstract standard methodology to work from.

4-1-1. Model and identify different phases in the process. Beginning with the general framework of Data Management and Archival Storage, in addition to the InterPARES model of Maintain Electronic Records, this project will identify the different phases in the process and define the

objectives of these phases. **(Kevin Glick/Eliot Wilczek/Project Analyst/David Gerwirtz/Stephen Yearl/Greg Colati/ Robert Chavez) Month 7–Month 14**

4-1-2. Define actions. Each phase of the Maintain Electronic Records process will be broken down into a series of actions that must be carried out during that phase and any expected results of those actions. The result of steps 4-1-1 and 4-1-2 will be a general methodology framework. This framework will not be a standard and the work necessary to develop one is beyond the scope of this project. **(Kevin Glick/ Eliot Wilczek/Project Analyst/David Gerwirtz/Stephen Yearl/Greg Colati/ Robert Chavez) Month 7–Month 14**

4-1-3. Analyze each action of the general framework. It is necessary to evaluate each the phase described in the general framework to determine its applicability to the university archives community. Each action will be accepted, eliminated, or changed to suit the community. **(Kevin Glick/Stephen Yearl/David Gerwirtz/   Eliot Wilczek/Project Analyst/Greg Colati) Month 10–Month 14**

4-1-4. Define terminology. We will draft the guide in language that reflects university archives assumptions and practices. A glossary will be created to clarify words with multiple or ambiguous meanings. **(Kevin Glick/Eliot Wilczek) Month 10–Month 14**

4-1-5. Reference related standards. We will note any standard relevant to this community and identify missing standards that need to be developed. **(Kevin Glick/Eliot Wilczek) Month 12–Month 14**

4-1-6. Identify tools. We will identify any tools that may or must be used to perform each of the actions described in the Abstract Standard (for example, an application to collect the necessary metadata in a form acceptable to the archives repository). **(Project Analyst/David Gerwirtz/Stephen Yearl/Robert Chavez) Month 7– Month 14**

4-2. Undertake six Maintain Electronic Records projects. Yale and Tufts will each undertake three (3) separate Maintain Electronic Records projects, managing the storage of the same six accessions described in section 3-3 with the Fedora system according the guide. The variety of record types and formats will necessitate different preservation strategies, greatly enhancing our understanding of the Maintain Electronic Records process. These projects will allow us to test and iteratively revise the guide according to our real-world experiences. In addition, this will allow us to compare the steps of the maintain process to Fedora's current ability to undertake those steps and from that produce a gap analysis of what tools or modifications, if any, Fedora needs to successfully and fully execute these steps. The Fedora Project Analyst and Gewirtz will be

primarily responsible for this activity with technical assistance from Yearl and Chavez and in consultation with Wilczek, Glick, and Colati. **(Kevin Glick/Eliot Wilczek/Project Analyst/David Gerwirtz/ Stephen Yearl/Greg Colati/Robert Chavez/Thorton Staples) Month 15–Month 18**

5. Report findings / critique Fedora. The project will produce a report on Fedora's ability to serve as an electronic records preservation system. Fedora will be judged by its ability to accommodate the records preservation system requirements. The report will propose future developments to Fedora that may be necessary. This report will form the basis for the development of a records-centered Fedora interest group to support the next phase of its development. We will post this report, along with all other findings, guides, and tools, on the project website and make freely available. The report will also serve as a basis for presentations and publications on the findings of this project. All members of the project team will work together on this activity, while Glick and Wilczek will be primarily responsible for writing the final report. **(Kevin Glick/Eliot Wilczek/ Thorton Staples) Month 16–Month 18**

6. Future work and related grants / implement findings. Further development of Fedora and implementation of complete preservation systems (including the output of records from the archives to a consumer) are beyond the scope of this grant project. No NHPRC funds will be used to undertake this work. If Fedora will only require modifications that can be reasonably implemented in order to bring records into the archives and maintain them according to the guides developed in this project, both Tufts and Yale will build the use of Fedora into their normal operating procedures with their own funding. If Fedora needs substantial modifications or is judged to not be capable of undertaking the tasks properly, Yale and Tufts will continue such research and seek additional outside funding.

References:

Bantin, Phil, *Functional Requirements for Recordkeeping Systems – Evolution of the IU Functional Requirements, 2002.*

Findings on the Preservation of Authentic Electronic Records, US-InterPARES Project, September 2002.

Functional Requirements for Preservation, Preserving the Electronic Records Stored in a Records Management Application (PERM Project), December 2002.

Producer-Archive Interface Methodology Standard, Consultative Committee for Space Data Systems, CCSDS-651.0-R-1, December 2002.

Reference Model for an Open Archival Information System (OAIS), Consultative Committee for Space Data Systems, CCSDS-650.0-B-1. January 2002.