InterPARES 2 Project

International Research on Permanent Authentic Records in Electronic Systems

Case Study Proposal
Validation of the InterPARES Preservation Model
Using Records and Data from a NASA Spacecraft Mission Operation
Focus 2 - Sciences

William Underwood
Georgia Tech Research Institute

June 2002

Rationale (Research Questions)

Can the PTF Model of the activities for Preserving Authentic Electronic Records be applied in the real world? In other words, "Is it a valid model of the activities necessary to preserve authentic electronic records?"

One of the merits of the InterPARES Preservation Model is its comprehensiveness and generality. It provides a generic preservation strategy (or framework) that can be used by archival institutions to develop their own preservation strategies depending on their institutional requirements and the specific bodies of records they must preserve. If specific management decisions were included in the model itself, it would compromise its generality. However, it needs to be demonstrated that this general preservation model works in specific cases. This would contribute to the belief on the part of archival institutions that the preservation framework actually works, and would provide them with an example of how to go about applying the framework in their own archival institutions.

It is proposed to validate (or falsify) the IDEF0 model for Preservation of Electronic Records (Version 6) with data from a specific case of business records and for a specific archival institution. Falsification amounts to identifying activities, controls, inputs or outputs of the model that are not coherent, or need further specification. If this occurs the model is refined until it is possible to apply the model to the specific case. If a model is falsified, an attempt should be made to revalidate the model with another case.

Research Methodologies

A walkthrough of an activity model is one way of reviewing the model in order to validate it or falsify it. During InterPARES 1, Case Study 26, the New York Worker's

Compensation Board (WCB) Electronic Case Folder System, was used in a walkthrough of the Preserve Electronic Records model.[1] Parts of the model were validated. However, the case studies were not designed with the objective of testing or validating the Preservation Model. For instance, there was no Appraisal Report or Terms and Conditions of Transfer for this case. Furthermore, case study data for an archive that would be preserving the records was not collected.

To gain the kinds of information that will be needed in the walkthrough, one first needs to design a data collection instrument. One of the obstacles that the PTF encountered in conducting the first walkthrough was the lack of a data model for the inputs and outputs of activities. Thus, it is suggested that a preservation data model be constructed as a prerequisite to conducting a second walkthrough of the Preservation Model.

The method used in the walkthrough is to iteratively step through each of the lowest level activities in the Preservation Model:

(1) Reviewing the activity definition and the input, output and control definitions.
(2) Identifying data elements of labels on input and output arrows.
(3) Defining the transformation of inputs to outputs.
(4) Determining values of the data elements that are related to the specific body of records.
(5) Recording the results and any problems or issues that arise and suggest possible solutions.

The following is a possible organization for a walkthrough. The roles are

- The *presenter*, who "puts on the table" the Preservation Model that was being reviewed.
- The *reviewers*, who have a good understanding of the Preservation Model and interpret the model definitions and links to ask questions of the case study expert to identify data corresponding to inputs and outputs, and raise issues and suggest solutions to problems.
- The *case study expert*, who answers questions posed by the reviewers about the data from the case study.
- The *secretary*, who records the discussed facts and issues and takes and distributes the minutes.

Description of the Case Study Subject

The case study subject will be the records and data of a NASA Spacecraft Mission Operation and a NASA Space Science Data (and Records) archive. The Consultative Committee for Space Data Systems (CCSDS) Panel 2 has developed a reference model

[1] A Walkthrough of the PTF IDEF0 Model for Preserving Electronic Records, Appendix to Preservation Task Force Final Report.

for a digital archives.[2] This reference model is being considered as an ISO Archiving Standard.[3] Panels 1 and 2 are considering end-to-end (plan mission, operate mission, utilize mission information) representations of space data and records using XML. The NASA National Space Science Data Center (NSSDC) is at NASA Goddard Space Flight Center. They are implementing the "utilize mission information" as an archival function. Dr. Underwood is a member of the XML Working Group of CCSDS Panel 2, has briefed the InterPARES Preservation Model to the Panel 2.

Welch and Karlin[4] describe an IDEF0 process model accomplishing a spacecraft mission. "The top level functions to accomplish the mission are plan mission, operate the mission system and utilize the mission information." "The planning function (A1) consists of three major sub-functions, the operations design, the development, modification and maintenance of procedures and databases and the planning of what is desired of the rest of the system to support both testing and in-flight operations." "The operate mission system function (A2) consists of four major sub-functions, scheduling mission activities, performing the mission activities, evaluating the activity performance as well as the general health and safety of the mission system and transporting and handling mission data." "The utilize mission data function (A3) primarily consists of navigation, spacecraft bus operations and science or payload analysis. The science or payload analysis includes science data processing, analysis of quick look products, data archive, and presentations and papers discussing the scientific findings."

A case study would be conducted for a specific NASA mission (probably excluding "plan mission"). The case study would have a scientific focus, the case study data would be collected with the walkthrough of the preservation model in mind. Since this includes the judgement of the preserver as to whether the records selected for preservation can be presumed authentic, it would include collection of data that can be used for "Testing of the ATF's method of assessment with the Benchmark Requirements for Presumption of Authenticity."[5]

Since, it is necessary to collect data on the business activities of the space mission that generate the records and data, and to collect data on the reliability and authenticity of these records, the case study data is also of potential use for addressing research questions in the Domain 1: Records Creation and Domain 2: Record Authenticity and Reliability.

Research Team

Lead Investigator: Bill Underwood, American InterPARES, Co-chair Working Group 3.2
Co-investigators: (Members of Working Group 3.2)
Research Assistant: Georgia Tech and UBC graduate students

---

[2] CCSDS 650.0-R-2: Reference Model for an Open Archival Information System (OAIS). Red Book (Draft Standard). Issue 2. June 2001.
[3] http://ssdoo.gsfc.nasa.gov/nost/isoas/.
[4] D. Welch & J. Karlin. Functional Model for Spacecraft Operations, *SpaceOps96.*
[5] Dr. Underwood has suggested this as another case study.

Timeline:

1. Develop preservation data model and relate to PTF activity model for Preserve Electronic Records (It is not necessary to complete this task in order to start the remaining tasks).
2. Design case study
3. Human subjects approval
4. Make arrangements with case study subject
5. Progress Report, InterPARES Workshop, Sept. 2002
6. Collect data
7. Conduct the walkthrough with case study data
8. Interim Report, InterPARES Workshop, Feb. 2003
9. Follow up with case study subject
10. Final Report, InterPARES Workshop, Sept. 2004