

---

---

Appendix 7

**Walkthrough Applying  
The “Preserve Electronic Records”  
Model**

Version 5.1

---

---

Preservation Task Force

**Walkthrough Team:**

Kevin Glick  
Hans Hofman  
Rich Lysakowski  
Ken Thibodeau  
Bill Underwood  
Bruce Walton

8 October 2001

## Introduction

A walkthrough of the "Preserve Electronic Records" model has been conducted.<sup>1</sup> The primary purpose was to test and improve the model using data from one of the InterPARES case studies.

The Preservation Task Force used the IDEF(0) methodology to analyze the problem of preserving authentic electronic records, and the lowest-level activities of the resulting model are believed to be a solution to the problem. The task force wants this model to be of the highest quality, to be understandable to archivists and computer system developers, and apply to a broad variety of electronic records.

Walkthroughs are an effective way to improve the quality of documents that describe the analysis, design, code and user interface of a system. A walkthrough is a peer-group review of any information system product. There are a number of types of walkthroughs—activity model, data flow model, and user requirement walkthroughs that are concerned with the functionality of the system; design walkthroughs that are concerned that a system or program design meets functional requirements; code walkthroughs that are concerned that program code satisfies the program design; and test walkthroughs to ensure the adequacy of test data for a system.

The invention of walkthroughs should probably be attributed to both Gerald Weinberg and Michael Fagan. Weinberg used the concept of a walkthrough in a *Programmer Team* approach to developing software.<sup>2</sup> Fagan developed a process called *Formal Inspections* to deal with the problem of reducing errors in the development of large software systems.<sup>3</sup> Yourdon embraced Fagan's ideas and developed the concept of *Structured Walkthroughs*.<sup>4</sup>

The objectives of the walkthrough of the Preservation Task Force preservation model were:

- To test the IDEF(0) model for the preservation of electronic records (version 5.1) with data from a specific case—in other words, to demonstrate that there is a real-world interpretation of the model.
- To create examples of a preservation strategy, preservation action plan(s), and targeted preservation methods that are linked to the body of records in an actual case.
- To more precisely specify the definition of an activity as a transformation of inputs to outputs.
- To identify the data elements of the inputs and outputs to activities of the model.

The walkthrough was conducted using version 5.1 of the "Preserve Electronic Records" model. Lessons learned in this walkthrough were applied to develop version 6, the final version of the first InterPARES Project.

---

<sup>1</sup> The walkthrough was conducted at a Preservation Task Force meeting during InterPARES Workshop No. 9, 15–17 October and 20 October 2001, Rome.

<sup>2</sup> G. M. Weinberg. *The Psychology of Computer Programming* (New York: Van Nostrand Reinhold, 1971).

<sup>3</sup> M. E. Fagan. "Design and Code Inspections to Reduce Errors in Program Development," *IBM Systems Journal* 15, 3 (1976): 182–211.

<sup>4</sup> E. Yourdon. *Structured Walkthroughs*, 4th Ed. (Englewood Cliffs, NJ: Yourdon Press, 1989). The state of the art is described in D. Freedman and G. Weinberg. *Handbook of Walkthroughs, Inspections and Technical Reviews*, 3rd Ed. (New York: Dorsett Home Publishing, 1990).

## The Walkthrough

The walkthrough organization was relatively informal. The roles were:

- The *presenter*, who "put on the table" the preservation model that was being reviewed.
- The *reviewers*, who had a good understanding of the preservation model and raised issues and suggested solutions to problems.
- The *case study expert*, who answered questions posed by the reviewers about the data from the case study.
- The *secretary*, who recorded the discussed facts and issues and took and distributed the minutes.

The method used in the walkthrough was 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 suggesting possible solutions.

## The Case Study

Case Study 26, the New York State Workers' Compensation Board (WCB) Electronic Case Folder System, was used in the walkthrough of the "Preserve Electronic Records" model. The following items collected or prepared during the case study were used in the walkthrough:

- a) WCB Electronic Case Folder Workflow
- b) Questions about Business Context
- c) Case Study Overview
- d) *Case Study Interview Protocol (CSIP)*
- e) *Template Element Data Gathering Instrument (TEDGI)*
- f) Diplomatic Analysis of Case Study
- g) Interview transcript
- h) Round 3 and 4 Pre-interview
- i) Description of System (NY Workers' Compensation Web site)<sup>5</sup>
- j) ERWIN data model of Electronic Case Folder

There was no Appraisal Report or Terms and Conditions of Transfer for this case.

The discussion in this section is structured around the activities of the preservation model. For each activity, the definition and inputs and outputs to an activity are reviewed. The reader will

---

<sup>5</sup> <<http://www.wcb.state.ny.us/>>

notice that for some activities there is a partial discussion, or no discussion at all. Those activities were only partly reviewed or not reviewed due to time constraints.

In the notes of the walkthrough, sources of information from Case Study 26 are indicated in parentheses after data elements. The endnotes in this paper indicate possible modifications of the IDEF(0) diagrams based on the walkthrough. They correct some errors and inconsistencies and suggest some simplifications.

## **A1 Manage the Preservation Function**

### Activity Definition:

"... preservation is managed by producing a comprehensive preservation framework consisting of sets of *preservation strategies* and *preservation actions* plans, each linked to a specific body of electronic records selected for preservation, along with the technological infrastructure and *preservation methods* needed to implement the action plans. ..."

The reviewers identified the following data elements for the three inputs to this activity.

### Inputs:

#### 1. Information about Electronic Records Selected for Preservation

- a) Record Creator's Name: New York State Workers' Compensation Board (Source: *TEDGI* 3.1)
- b) Transfer Agent's Name: Same as above
- c) Identification of Records

Title: New York State Electronic Case Folder (Source: *CSIP*)

Description: Series of case files for adjudicating benefits of disabled workers (Source: *TEDGI*).

Model or Description of the Business Processes that generated these records: "Workers' compensation insurance provides weekly cash payments and the cost of full medical treatment, including rehabilitation, for covered employees who become disabled as a result of a disease or injury connected with their employment. It also provides payments for qualified dependents of a worker who dies from a compensable injury or illness. In administering this program, the Workers' Compensation Board receives and processes workers' claims for benefits, employers' reports of injury, and medical reports from physicians and other health care providers. The board adjudicates and resolves all issues and makes awards and findings as rapidly as possible to ensure that an entitled claimant receives benefits and medical treatment promptly. Hearings are conducted before law judges, or, on review or appeal, before panels of three board members." (Source: NYS WCB Web-site Description of the Mission of the Organization)

A business process analysis of the business processes was used to determine improvements in the business processes, and needs for changes in regulations and in designing the NYS Electronic Case Folder system. It was not collected as a part of the case study. (There may be different types of cases, depending on this business process model.)

### Document types:

- Claims for benefits
- Employer's reports of accidents and illness
- Correspondence
- Medical reports from physicians and other health care providers
- Insurance carrier's reports

(Source: supporting documents)

Volume:

There are over 300,000 open cases (Source: Supporting documentation)

File or Data Structure:

Relational Database describes the records and their relationships.

- d) Disposition (Scheduled date of transfer): (Source: Supporting documents)
  - e) Media on which current records are stored: Net HT Series 30GB 12" OSCAR optical WORM disks
  - f) Information requirements to support a presumption of authenticity
2. Information about transferred and accessioned records<sup>i</sup>
- a) Record Creator's name
  - b) Transfer Agent's name
  - c) Date of transfer
  - d) Identification of records
    - Title
    - Description
    - Volume
    - File or Date Structure
    - Technical Information (e.g., XML SQL, file formats, encoding)
  - e) Information supporting presumption of authenticity of records transferred by the records creator. This consists of evidence for each of the following requirements:
    - A.1.a Identity of the Record*
    - A.1.b Integrity of the Record*
    - Requirement A.2 Access Privileges*
    - Requirement A.3 Protective Procedures: Loss and Corruption of Records*
    - Requirement A.4 Protective Procedures: Media and Technology*
    - Requirement A.5 Establishment of Documentary Forms*
    - Requirement A.6 Authentication of Records*
    - Requirement A.7 Identification of Authoritative Record*
    - Requirement A.8 Removal and Transfer of Relevant Documentation*
- In lieu of a presumption of authenticity, a verification of authenticity should be provided.
3. Management Information about Preservation: This will only be available after the first transfer has been processed because it's feedback information. However, an archive would probably use feedback it had about series with similar characteristics in developing the strategy and plan.

**A1.1 Determine Preservation Requirements**

**A1.1.1 Identify Types of Archival Properties that must be Preserved**

Input:

Information about Electronic Records Selected for Preservation

The reviewers identified case study data corresponding to the three outputs of this activity.

Outputs:

- a) Classes of Records:
  - Claims for Benefits
  - Employers' reports of accidents and illness
  - Correspondence
  - Medical reports from physicians and other Health Care Providers

Insurance carrier's reports  
(Source: supporting documents)

- b) Types of Record Arrangement: Indexed on case file identifier. Secondary indexes on other attributes of the case file. (Source: ERWIN Data Model of Electronic Case Folder System) (To articulate the preservation strategy for these records, the preserver would need to determine: Are these attributes data elements in the database? Do they have any relationship to specific business processes, such as audits?)
- c) Types of Archival Bonds: Contents of case file ordered by document number. Document numbers are created when a document is imported into the electronic case folder system. The File Net High-Performance Image Import (HPII) system creates this number.

#### **A1.1.2 Determine how Records are composed from Digital Components**

Case study data were identified for one of the outputs of this activity.

##### Output:

Record Composition Requirements:

- 1. Documents in NYS Workers' Compensation Board Electronic Case File system are page images represented as TIFF 6 files. Page images must be reproducible (Source: *TEDGI*, Supporting Documentation).
- 2. Metadata about documents and cases are represented in a relational database managed using Sybase (Source: Supporting Documentation).
- 3. The TIFF 6 files are stored as files external to the database.

#### **A1.1.3 Determine How Records are Arranged**

##### Output:

Requirements for arranging records: Must be able to present cases in order of case file identifying number.

#### **A1.1.4 Determine How Archival Bonds are Expressed**

##### Output:

Requirements for instantiating archival bonds: must be able to order documents in a folder by document number.

#### **A1.1.5 Synthesize Requirements for Preservation**

##### Inputs:

Classes of Records

- Record Composition Requirements
- Types of record arrangement
- Requirements for arranging records
- Requirements for instantiating archival bonds

Case study data were identified for the output of this activity.

##### Outputs:

Specified Requirements for Preservation

- a) Classes of Records:
  - All classes of records are stored as document images.
  - Page layout and appearance must be preserved.
  - Each TIFF file contains multiple pages.
- b) Types of Record Arrangement:
  - Records must be arranged into case files according to case file ID and document number.

- Case files must be arranged by case file ID and by an index on case file ID.
- c) Types of archival bonds:  
archival bonds are instantiated on the basis of case file and document numbers.

#### **A1.1.6 Determine Basis for Certifying Authenticity**

##### Activity Description:

Guided by Archival Science and Institutional Requirements (both tunnelled to this diagram) and the Specified Requirements for Preservation applicable to a body of records selected for preservation, determine the basis for asserting the authenticity of the records. This basis will have two parts: information supporting the presumption of the authenticity of the records as transferred from the creator and information about how the preserver satisfies the applicable Specified Requirements for Preservation after the records are transferred to the preserver. The information requirements to support a presumption of authenticity of records up to their transfer is received as part of Information about Electronic Records Selected for Preservation and Information about Transferred and Accessioned Records, and will be designated for retention along with other information about the records. The types of information required to support an assertion that the preserver has preserved and reproduced authentic records become a control on subsequent preservation activities.

##### Controls:

Archival Science

##### Benchmark Requirements:

Information supporting the presumption of authenticity of the records as transferred from the creator.<sup>6</sup> This consists of evidence as to how the creator addressed each of the following:

*A.1.a Identity of the record*

*A.1.b Integrity of the record*

*Requirement A.2 Access Privileges*

*Requirement A.3 Protective Procedures: Loss and Corruption of Records*

*Requirement A.4 Protective Procedures: Media and Technology*

*Requirement A.5 Establishment of Documentary Forms*

*Requirement A.6 Authentication of Records*

*Requirement A.7 Identification of Authoritative Record*

*Requirement A.8 Removal and Transfer of Relevant Documentation*

If the evidence does not support a presumption of authenticity of the records as retained and transferred by the creator, their authenticity will have to be verified on a case-by-case basis.

Baseline Requirements: How the preserver satisfies the applicable requirements for preservation after the records are transferred to the preserver.<sup>7</sup> This consists of demonstration that each of the following requirements is met.

*B.1 Controls over Records Transfer, Maintenance, and Reproduction*

*B.1.a Unbroken custody of the records is maintained.*

*B.1.b Security and control procedures are implemented and monitored.*

*B.1.c The content of the record remains unchanged after reproduction.*

*B.2 Documentation of Reproduction Process and its Effects*

*B.2.a The date of the records' reproduction and the name of the responsible person.*

*B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver.*

---

<sup>6</sup> Authenticity Task Force, *Requirements for Assessing and Maintaining the Authenticity of Electronic Records*, Appendix 2.

<sup>7</sup> *Ibid.*

- B.2.c The impact of the reproduction process on their form, content, accessibility and use.*
- B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identify and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.*
- B.3 The Archival Description of the fonds includes ... information about changes the electronic records of the creator have undergone since they were first created.*

Inputs:

Specified Requirements for Preservation

- Information about Electronic Records Selected for Preservation
- Information about Transferred and Accessioned Records

Outputs:

- Basis for Certifying Authenticity of Transferred Records
- Authenticity Requirements for Preservation

**A1.2 Select Preservation Technologies**

For any aggregate of electronic records selected for preservation, preservation technologies must be used which are appropriate and adequate for reproducing the records, maintaining the archival bonds among the records, and satisfying the benchmark requirements for authenticity. Given continuing change in information technology, the selection is not likely to be a one-time decision. Rather A1.2 is a dynamic process dependent on changes in types of digital components of electronic records, the state of the art of IT (computer storage, operating system, database technology) in which new preservation options must be identified and evaluated and new preservation methods and technological infrastructure acquired.

**A1.2.1 Identify Preservation Options**

Case study data were identified for each of the inputs and outputs.

Inputs:

Information about Digital Components of Electronic Records

- a) Information about current media on which digital components are stored: WORM
- b) Relationship between TIFF files and relational database: Presuming the logical model of the database cannot be reduced to a flat file, is it essential that this structure be maintained in relational form? If the structure is essentially that of a graph, could it be replicated in other methods, e.g., XML DTD, data warehouse, or OO database? Could the TIFF files be integrated into the database?

Output:

Preservation Options

The institution might require keeping document images as TIFF files and storing digital components on DLT tapes. That would dictate a decision in this case to keep the digital components as TIFF files and to transfer the files to different media, e.g., from WORM to DLT tapes.

**A1.2.2 Evaluate Preservation Options<sup>ii</sup>**

Choice of storage media will be based on per unit cost, market penetration, transfer rate, longevity, and storage device costs which are determined by the State of Information Technology and budget, which is an institutional constraint.

Choice of archival storage system will be based on business needs and budget, e.g., frequent access to some bodies of records and infrequent access to other bodies of records might dictate the choice of a hierarchical storage system.



### **A1.2.3 Select Preservation Methods**

#### Input:

##### Evaluated Preservation Options

Case study data were identified for the output of this activity

#### Output:

Selected Preservation Methods: For example:

- a) Format for document images: preserve in TIFF format with ability to reproduce documents with TIFF viewers.
- b) Format for textual documents: preserve in PDF format with ability to reproduce documents with PDF viewers.
- c) Method for reconstituting case files and, if needed, records.
- d) Metadata about records or digital components should be stored in a relational database.
- e) The schema for relational databases should be represented in SQL.

### **A1.2.4 Acquire Capability to Apply Selected Preservation Methods**

Examples of the Technology Infrastructure needed for this case was identified.

#### Outputs:

1. Technology Infrastructure: Includes computers, media, storage systems, and generic software such as Operating Systems, Database Management System, Packaging Software.
  - a) Viewer software: Acquire viewer for TIFF that operates on current computer hardware and operating systems.
  - b) Format conversion software and hardware: Acquire hardware and software to convert WORM to DLT. Decide whether the creator—in which case it becomes a condition of transfer—does that conversion.
  - c) Acquire software to migrate obsolete format to current format.

The definitions of activities that used preservation methods were used to identify the kinds of preservation methods that would be needed.

2. Targeted Preservation Methods<sup>iii</sup>

#### Definition:

Software used to implement a preservation strategy that is related (targeted) to a body of records, types of electronic records, and classes of digital components). The source of information about the body of records, types of electronic records, and classes of digital components is the Specified Requirements for Preservation. Note: see information about the method and the body of records, types of electronic records, and classes of digital components to be added to the output of A.1.2.4 to produce Targeted Preservation Methods.

Targeted Preservation Methods (from Definition of Activity 1.2)

- Method for checking the integrity of the transfer process.
- Methods for checking the completeness of the transfer: Are all required case files included? Are all records in each case file present? Are case ID and document numbers uniquely assigned? Do data in the database match corresponding attributes of the TIFF files?
- Methods for managing and storing of digital components in digital files and on physical media.
- Methods to reconstitute and reproduce records.
- Methods to enable others to reproduce the records.

Targeted Preservation Methods (from A3, A3.2, and A3.3)

Storage Methods:

- Method for placing record components in storage, e.g., packaging and preserving integrity of record components)
- Storage update method
- Monitoring method
- Problem correction method
- Retrieval method
- Methods for Updating Components.

Targeted Preservation Methods (from A4)

- Record reconstitution method
- Presentation method
- Packaging Method (for dissemination).

### **A1.3 Specify Preservation Strategies and Actions**

#### **A1.3.1 Articulate Preservation Strategy**

Activity Definition:

Controlled by the Specified and Authenticity Requirements for Preservation applicable to a body of electronic records selected for preservation, and limited by the State of the Art of Information Technology. Use Information about this body of Electronic Records Selected for Preservation and Information about Transferred and Accessioned Records from the same body of records, along with Preservation Technology Specifications, which describe the Targeted Preservation Methods applicable to these records, to specify and output a comprehensive Preservation Strategy for preserving the body of records.

Controls:

- a) Specified Requirements for Preservation: Are the criteria for evaluating execution of the preservation strategy in the Specified Requirements for Preservation?
- b) Authenticity Requirements for Preservation
- c) Evaluation of Execution (Feedback from A1.4)

Inputs:

- a) Information about Electronic Records Selected for Preservation
- b) Preservation Technology Specifications, i.e., Information about Targeted Preservation Methods
- c) Information about Transferred and Accessioned Records<sup>iv</sup>

Output:

Preservation Strategy

Definition:

A coherent and comprehensive approach for preserving a body of records selected for preservation, derived from archival and institutional requirements, taking into account Evaluation of Execution of current and prior Preservation Strategies and reflecting the state of the Art of Information Technology. A preservation strategy includes objectives for maintaining components of electronic records and related metadata and information over time and for reproducing the records in authentic form, in the order imposed by the records creator, and criteria for evaluating execution of the preservation strategy. The strategy includes specifications for handling exceptions to its standards, and identifies the targeted preservation methods to be used.

Activities A2, A3, and A4 provide a Preservation Framework from which a Preservation Strategy specific to a body of records is constructed. A Preservation Framework is a generic preservation

strategy for achieving preservations goals such as: authentic records, retrievable records, and reproducible records.

To determine the elements of a preservation strategy, we considered how the preservation strategy is used as a control on all remaining processes in the preservation model. In parallel with this examination, we sought to determine the elements of a preservation action plan (output of A1.3.2), as they were inputs to many of the remaining processes.

Preservation action plans are not transformed by the activities to which they are input. The actions of an activity are being triggered by the execution of the preservation action plan. Preservation action plans are a sequence of preservation actions or triggers.

When a Preservation Action Plan is executed it sends control (preservation) actions to other activities of the model. It might help to have a separate activity called "Execute Preservation Action Plan" that has Preservation Action Plans as an input and Preservation Actions as an output that triggered the rules and methods of preservation activities.

The reviewers concluded that having a preservation strategy as a control and preservation actions plans as inputs to subactivities of A2, A3, and A4 may be redundant. Rather the preservation strategy may only be needed as a control on A1.3.2 and as an input to A1.3.3.

The reviewers were unable to characterize a preservation strategy for a specific body of records apart from the Generic Preservation Framework and the preservation actions for a specific body of records. This leads us to suggest a change in terminology.

<b>Preservation Model Version 5.1</b>	<b>Suggested Terminology</b>
Preservation Framework	Preservation Strategy
Preservation Strategy	Preservation Plan
Preservation Action Plan	Preservation Action

Activity 1.3.1 is a very complex activity and is probably worthy of decomposition.<sup>v</sup>

### **A1.3.2 Plan for Implementing Preservation Strategy**

The reviewers identified data elements making up the Terms and Conditions for Transfer and some sample data from the case study.

#### Outputs:

- a) Terms and Conditions for Transfer:
  1. Record Creator's Name: NY State Workers' Compensation Board
  2. Transfer Agent's Name and Authority: John Doe, Records Manager
  3. Identification of Records
    - a) Identification Number (of Record Series)
    - b) Title: Electronic Case Folder System
    - c) Description: Series of case files for adjudicating benefits of disabled workers
    - d) Document Types: Claims for Benefits, Employer's reports of accidents and illness, Correspondence, Medical Reports, Insurance Carrier's reports
    - e) File or Data Structure: Relational Schema
    - f) Types of files or selection criteria for including/excluding files in a transfer: (e.g., closed files, history files)
    - g) Volume (number of files, number of media, size of files): 300,000 cases, one million files, nine million document images.
  4. Expected transfer date or period: Position of the transfer in the records life cycle: (e.g., one year after closeout of a file)

5. Medium or means that will be used to effect the transfer: DLT Tape
6. Technical information that should accompany the transfer
  - a) Metadata for the digital components should comprise: document type, digital component, the format, preservation history, and the originating technological environment.
  - b) All digital components that are document images that are not in TIFF format should be converted to TIFF format.
  - c) The schema describing the metadata for the records in the Electronic Case Folder System should be represented in SQL.
  - d) Metadata for the records should comprise: the provenance (including a description of the business process procedures); the documentary context; the benchmark requirements for authenticity; information about how the records relate to the particular digital components that constitute the records; information about the different types of documents, the archival description of the records, and their arrangement.
7. Information supporting presumption of authenticity of records transferred by the records creator. This consists of evidence for each of the following requirements:
  - A.1.a Identity of the record*
    - A.1.a.i Name of author, Name of addressee: The ECFS data model permits the association of author's name, addressee, name of action or matter, and chronological date with each document.
    - A.1.a.ii Name of action or matter.
    - A.1.a.iii Chronological date.
    - A.1.a.iv Expression of Archival Bond: When documents are imported by FileNet, a case file is ordered by document number.
    - A.1.a.v Indication of attachments: Document preparation and mail transmittal preparation rules address how attachments are kept in the case folder.
  - A.1.b Integrity of the record*
    - A.1.b.i Name of Handling Office: NY WCB.
    - A.1.b.ii Name of Office of Primary Responsibility: Same as above.
    - A.1.b.iii Indications of types of annotations: FileNet supports annotations, but they are not used.
    - A.1.b.iv Indication of technical modifications: Paper documents are scanned into document images in TIFF 6 format and maintained on WORM disks.
  - A.2 Access Privileges:*  
Access to ECFS is controlled via passwords, job titles, workgroups, geographic location and business need.
  - A.3 Protective Procedures:*  
Loss and Corruption of Records: There are back-up copies of the WORM disks and transaction logs.
  - A.4 Protective Procedures:*  
Media and Technology: WORM disks are guaranteed for more than one hundred years.
  - A.5 Establishment of Documentary Forms:*  
Each form is described in a procedure manual that is managed in Lotus Notes.
  - A.6 Authentication of Records:*  
Authentication of document images in a case file is occasionally required in the adjudication process. They are presumed authentic because they are scanned images of paper documents and they are used in the normal course of business.

*A.7 Identification of Authoritative Record:*

The document images are the authoritative record unless the paper file is still available.

*A.8 Removal and Transfer of Relevant Documentation:*

There has not yet been a transition of active records to semi-active or inactive status, which involves a removal of records from the electronic system.

In lieu of a presumption of authenticity, a verification of authenticity should be provided.

b) Preservation Action Plans

The reviewers created some examples of preservation action plans for the specific body of records in the case study. The following are examples of preservation action plans that a creator might need to perform before transferring records to an archives.

1. When records are transferred from the record creator to the preserver, there must be an integrity check.
2. When data from a relational database are to be transferred, they will be transferred as flat ASCII files.
3. When a schema for a relational database is to be transferred, it should be represented in SQL.

The following is a possible Preservation Action Plan Associated with a Body of Records Selected for Preservation that can be used with Transferred and Accessioned Records.

Preservation Action Plan

1. Retrieve digital components for Claims for Benefits in Electronic Case Folder System that are ASCII text files.
2. Convert the ASCII Text files to TIFF multipage format using preservation method (TiffMaker).
3. Store the digital components converted to TIFF multipage format back to archival storage.
4. Store in the database the information that on this date the digital components for Claims for Benefits in the Electronic Case Folder System that were in ASCII text format have been converted to TIFF multipage format.

**A1.3.3 Assess Strategy and Plan<sup>vi</sup>**

Activity Definition:

Using Information about (the application of the Preservation Strategy and implementation of the Preservation Action Plan to) Transferred and Accessioned Records, determine whether and to what extent the Preservation Strategy and Preservation Action Plan(s) applicable to a body of electronic records selected for preservation have succeeded in satisfying Specified and Authenticity Requirements for Preservation. Use this assessment to produce an Updated Strategy and/or Updated Action Plans(s). If monitoring of electronic records selected for preservation prior to their transfer produces significant Updated Information about Electronic Records Selected for Preservation, revise the strategy or plan accordingly. Similarly, if applicable preservation methods have been changed, revise the strategy and /or plan to reflect the Updated Preservation Technology Specifications.

**A1.4 Evaluate Execution of Preservation<sup>vii</sup>**

Activity Definition:

"Review Management Information about Output of Electronic Records to determine if requirements for authentic copies are being satisfied and to characterize customer satisfaction."

"Produce a Report on Authenticity of Records to respond to any challenge to the adequacy and efficacy of the preservation process."

This is where the Preservation Framework (strategy) is demonstrated to meet the Authenticity Task Force's Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records.<sup>8</sup>

#### Baseline Requirements

The preserver should be able to demonstrate that:

##### *Requirement B.1: Controls over Records Transfer, Maintenance, and Reproduction*

The procedures and systems(s) used to transfer records to the archival institution or program, maintain them, and reproduce them embody adequate and effective controls to guarantee the records' identity and integrity.

This requirement is satisfied by: (1) activity A1.3.2 for creating Terms and Conditions for Transfer; (2) activity A2.2, which compares the transfer with the Terms and Conditions for Transfer; (3) activity A2.2.3, which carries out Preservation Action Plans that use Preservation Methods to bring digital components into compliance with the preservation strategy; and (4) activity A4, which reproduces the record from maintained digital components.

and specifically that:

##### *B.1.a Unbroken Custody of the records is maintained*

This requirement is satisfied by the activities for Selection of Creator's Records for Preservation, Terms and Conditions for Transfer, and Transfer and Preservation of Records.

##### *B.1.b Security and control procedures are implemented and monitored*

This requirement is satisfied in part by Access Control and Access Privileges of a DBMS (database management systems).

##### *B.1.c The content of the record remains unchanged after reproduction.*

This requirement is satisfied by Preservation Methods that preserve the content of electronic records.

##### *Requirement B.2: Documentation of [Update/] Reproduction Process and its Effects*

The activity of reproduction has been documented and that this documentation includes:

*B.2.a The date of the records' reproduction and the name of the responsible person;*

*B.2.b The relationship between the records acquired from the creator and the copies produced by the preserver;*

*B.2.c The impact of the reproduction process on their form, content, accessibility, and use; and*

*B.2.d In those cases where a copy of a record is known not to fully and faithfully reproduce the elements expressing its identity and integrity, such information has been documented by the preserver, and this documentation is readily accessible to the user.*

These requirements are satisfied by activity A1.2.3, Selecting a method to apply to a class of preservation objects, and by Preservation Action Plans that store a record of updates to digital components and by assessment of the effect of a Preservation Action Plan on the reproduction of form and content.

##### *Requirement B.3: Archival Description*

The archival description of the fonds containing the electronic records includes—in addition to information about the records' juridical-administrative, provenancial, procedural, and documentary

---

<sup>8</sup> Appendix 2.

contexts—information about changes the electronic records of the creator have undergone since they were first created.

This requirement is satisfied by: (1) requiring the transfer contain evidence supporting the presumption of authenticity of the transferred records, (2) maintaining records of refreshment of digital media and updates to digital components, and (3) activity A3.3, Update Digital Components, and specifically by preservation action plans that document updates to digital components.

Each of the Requirements for Supporting the Production of Authentic Copies of Electronic Records is satisfied by some set of activities of the Preservation Framework. This leads us to conclude that the Preservation Task Force's "Preserve Electronic Records" model provides a framework for preserving authentic electronic records. Within that framework, a variety of preservation strategies can be developed by archival institutions that are dependent on the characteristics of the selected, transferred and accessioned records; institutional requirements; and the current and changing state of information technology. The framework guides the development of strategies that can satisfy the Authenticity Task Force's baseline requirements.

## **A2 Bring in Electronic Records**

### Activity Definition:

"Following the direction established in the preservation strategy for a given body of records selected for preservation, the 'bring in' or ingest function applies preservation methods(s) targeted to that body of records to implement the preservation action plan for those records by processing each transfer of electronic records into accessioned electronic records. ..."

### **A2.2 Verify that the Transfer Is Authorized**

#### Inputs:

Registered Transfer

Preservation Action Plan<sup>viii ix</sup>

#### Outputs:

Rejected Transfer

Conforming Transfer

#### Activity Definition:

Verification that a transfer is authorized is based on comparing the terms and conditions for transfer established as part of the Preservation Strategy during Appraisal, with the information accompanying the Registered Transfer.<sup>x</sup>

The following is a more precise specification of the transform of inputs to outputs.

Compare the terms and conditions for transfer with the information accompanying the Registered Transfer to determine whether:

- a) The transfer was sent by an authorized person,
- b) It comprises the records specified for transfer,
- c) It includes required information about the records and their digital components, and
- d) It includes required information about the basis for asserting the authenticity of the records as received.

If the terms and conditions of transfer are satisfied, the conforming transfer is passed to the next step where its contents are examined. Otherwise, the transfer is rejected or the submitter is asked to address any problems identified.



## **A2.3 Examine Electronic Records**

### Input:

Preservation Action Plans<sup>xi xii</sup>  
Conforming Transfer

### **2.3.1 Map Records and Digital Components within Transferred Material**

#### Output:

Mapped Records and Digital Components<sup>xiii</sup>  
Rejected Transfer

#### Activity Definition:

In accordance with an institution's Accessioning Policy and the applicable Preservation Strategy, using the Technological Infrastructure in place, determine how the records, their digital components, and the information about them included in a Conforming Transfer are identified, and where they are located in the digital files and other materials received. Identify any records or components that should be in the transfer but are not found and determine if there are any records or components in the transfer that should not have been transferred. If critical deficiencies are found, terminate processing and output the Rejected Transfers. If no critical problems are encountered, output the Mapped Records and Digital Components, along with related information.

The following is a more precise specification of the transformation of inputs to outputs.

Examine the digital files and digital components of records in the transfer, along with accompanying information to:

- a) determine how the records, their digital components, and the information about them in the conforming transfer are identified.
- b) determine where they are located in the digital files and other materials received.
- c) identify any records, aggregates of records or components that should be in the transfer but are not found.
- d) identify any records, aggregates of records or components in the transfer that should not have been transferred.

Produce a mapping of information about electronic records transferred to the digital files that were transferred. Identify and locate the digital components included in the files(s), and link them to the records that they constitute.

### **A2.3.2 Verify that the Records in the Transfer can be Preserved and Reproduced.**

#### Activity Definition:

Ensure that transferred records can be preserved and reproduced in accordance with the applicable preservation strategy. Determine if all of the records that should be in the transfer can be reconstituted and presented; if all digital components of these have been received and are in the formats stipulated in the terms and conditions of transfer; if archival aggregates established by the records creator can be re-established; if the archival bonds among records can be expressed or instantiated, and also what basis exists for asserting the authenticity of the records as transferred. Records that can be preserved and reproduced are output as Preservable Records. In the case of a record that cannot be preserved in accordance with the preservation strategy, identify the digital component(s) that prevent or impede implementation of the strategy and forward them for actions that will enable preservation of the record. Digital components modified as a result of such actions are returned for verification of the success of the modification. If it was necessary to modify the applicable preservation strategy to enable preservation of the records, the secondary verification will be on the basis of the revised strategy. The verification process will take into account any cases where digital components were not successfully



modified, along with other problems discovered with the transfer, to determine whether, under the institution's accessioning policy, the preservable records— including records deemed to have only minor problems—should be accessioned or the transfer should be rejected.

Input:

Mapped Records and Digital Components<sup>xiv</sup>

The following is a more precise specification of the transformation of inputs to outputs.

The Mapped Records and Digital Components are used to examine the digital files and digital components of records in the transfer, along with accompanying information.

If:

- a) all digital components are in the formats stipulated in the terms and conditions of transfer and there are methods for reconstituting and reproducing records from these digital components;
- b) the archival aggregates established by the records creator can be re-established;
- c) the archival bonds among records can be expressed or instantiated; and
- d) there is adequate evidence for a presumption of authenticity of the records as transferred, then output the Preservable Records.

Otherwise, if actions are required to preserve some of the individual records transferred and the archival sets in which these records belong, then output these Digital Components of a Record that cannot be preserved.

When Conforming Digital Components are returned as a result of taking actions needed to preserve the Record, combine them with the rest of the preservable records and output the preservable records.

When Non-Conforming Digital Components are returned as a result of taking actions needed to preserve the Record, reject that portion of the transfer.

**A2.3.3 Take Action Needed to Preserve the Record**

When there are Digital Components of a Record that cannot be preserved, and the conditions of the Preservation Action Plan for this Transfer apply to these components, take the preservation actions and output the Conforming Digital components, or output the Non-Conforming Digital Components.

**A3 Maintain Electronic Records**

Activity Definition:

“Following direction established in the preservation strategy for a given body of records selected for preservation, apply preservation methods(s) targeted to that body of records to implement the preservation action plan for those records ...”

**A3.1 Manage Information about Records**

The reviewers identified data elements for the first input to this activity.

Inputs:

1. Information about Accessioned Electronic Records
  - a) Record Creator's name
  - b) Transfer Agent's name
  - c) Date of transfer
  - d) Identification of records
    - Title
    - Description
    - Volume

File or Date Structure

Technical Information (e.g., XML SQL, file formats, encoding)

- e) Information supporting presumption of authenticity of records transferred by the records creator. This consists of evidence for each of the following requirements:

A.1.a Identity of the record

A.1.b Integrity of the record

A.2 Access Privileges

A.3 Protective Procedures: Loss and Corruption of Records

A.4 Protective Procedures: Media and Technology

A.5 Establishment of Documentary Forms

A.6 Authentication of Records

A.7 Identification of Authoritative Record

A.8 Removal and Transfer of Relevant Documentation

In lieu of a presumption of authenticity, a verification of authenticity should be provided.

- f) Information about the media on which the records were transferred.

2. Basis of Authenticity of Transferred Records<sup>xv</sup>
3. Information about Updated Digital Components
4. Updated Storage Information: Location of files in Accessioned Electronic Records.
5. Retrieval Request

This activity requires the following functions:

1. The capability to store, retrieve, and update data.
2. A data definition language (DDL) for defining the conceptual schema (data items, records, primary and secondary keys, record relationships and integrity constraints), user views, and a data dictionary.
3. [Access Control and Access Privileges] Mechanisms for controlling access to data and for defining what actions may be taken by processes or users.
4. [Transaction integrity] The capability to define transaction boundaries, i.e., the logical beginning and end of transactions.
5. [Control of concurrent transactions] Safeguards to prevent erroneous results that can occur when two or more processes attempt to access a data record concurrently.
6. [Recovery Services] The capability to restore the database in the event of some system failure.

These functions can be provided with a DBMS.

### **A3.1.1 Maintain Information about Records**

### **A3.1.2 Retrieve Information about a Requested Record**

### **A3.1.3 Retrieve Information about Digital Components<sup>xvi</sup>**

## **A3.2 Manage Storage of Digital Components of Records**

### **A3.2.1 Place Record Components in Storage**

[file size constraint] When digital components are placed in archival storage, they should be stored as a file no larger than the capacity of the chosen media.

### **A3.2.2 Refresh Storage**

When the prescribed date for updating a storage medium occurs, use the prescribed storage update method to transfer files to the new storage medium.

### **A3.2.3 Monitor Storage**

When an environmental condition for archival storage (e.g., humidity or temperature) is out of range, signal the type of archival storage problem.

### **A3.2.4 Correct Storage Problems**

### **A3.2.5 Retrieve Components from Storage**

## **A3.3 Update Digital Components<sup>xvii</sup>**

### **A3.3.1 Migrate Digital Components to Current Formats**

[Migration Strategy Rule] When format of digital component is obsolete, use a Targeted Preservation Method to migrate the digital component in obsolete format to the chosen current format, and associate with this group of digital components documentation of the migration of digital components to the new format.

### **A3.3.2 Convert Digital Components to Standardized Formats**

[Conversion to Standard Format Rule] When format of digital component is proprietary, use a Targeted Preservation Method to convert the digital component in proprietary format to the chosen standard format.

### **A3.3.3 Transform Digital Components to Persistent Format**

When the preservation action for the digital components of an archival set is to transform them to persistent format, apply the method for transforming the digital components to a software and hardware independent format, e.g., XSL-FO.

## **A4 Output Electronic Records**

### Activity Definition:

“Following direction established in the preservation strategy for a given body of records selected for preservation, apply preservation method(s) targeted to that body of records to implement the preservation action plan for producing an authentic copy of a record in response to a request for it. ...”

### **A4.1 Manage the Request**

### **A4.2 Review Retrieved Components and Information**

### **A4.3 Reconstitute Record**

### **A4.4 Package Output**

When requested records are to be disseminated, wrap or encapsulate digital components of the electronic record, or the Reconstituted Electronic Record, with the associated metadata using the chosen Packaging Method.

## **Summary and Conclusions**

A primary result of the walkthrough was to clarify the concepts of a Preservation Strategy, Terms and Conditions for Transfer, Preservation Action Plans, and Targeted Preservation Methods. Examples were also created for preservation plans and preservation methods. Another result was the demonstration that the preservation model provides a framework for satisfying the Authenticity Task Force's Baseline Requirements Supporting the Production of Authentic Copies of Electronic Records.

Conducting a walkthrough of the preservation model without having a data model was very time-consuming. However, we now have a clearer idea of the data elements making up the inputs and outputs to activities. During InterPARES 2, the data elements identified can be used in constructing a preservation data model.

Because the definitions of activities in the preservation model were descriptive, rather than defining transforms of inputs to outputs, the product (preservation model version 5.1) was not ready for a test of the model against case data. The walkthrough of the model amounted to a review of the model—raising issues and problems and discussing them, providing feedback on the model mixed with continued problem solving.

The case data were not adequate for testing the product. The case studies in the InterPARES Project were not designed with a walkthrough of the preservation model in mind. However, the attempt to walk through the model with case study data provided us with insight into the kinds of information that will need to be collected in order to perform a walkthrough. During InterPARES 2, a case study can be designed to collect the data that are needed to support a more complete walkthrough and validation of the preservation model.

## Notes

---

<sup>i</sup> "Information about Transferred and Accessioned Records" should not include the term "accessioned" because the records are not yet accessioned. Also information to be included in "The Information about Transferred Records" is set out in the "Terms and Conditions for Transfer."

<sup>ii</sup> May not need requirements for preservation as a control.

<sup>iii</sup> Can "Targeted Preservation Method" that is an output of A1.2 and a control on A2, A3, and A4 be eliminated because it is an input to A1.3 (even though named "Preservation Technology Specifications") where it is incorporated into Preservation Strategy, which controls A2, A3, and A4? A strategy is a statement of direction: where you want to go, how you will get there, and how you will know you arrived. The method is software. You can't carry out the strategy without it, but it is distinct from the strategy. A method may be invoked by several strategies, any one of which may change. For reasons of economy and effectiveness, an archive will tend to limit the number of methods it supports. As long as one strategy still requires a method, the software remains in use.

<sup>iv</sup> That a transfer conforms to the Terms and Conditions of Transfer is checked in activity A2.2, Verify that the Transfer is Authorized. The output of that activity is Conforming Transfer, which has the definition "A transfer of electronic records that satisfies the terms and conditions stipulated for the transfer." If the transfer doesn't conform, it is rejected. It appears that activities A1.3.1 and A1.3.2 only need to be performed for Information about Electronic Records Selected for Preservation in order to produce a Preservation Strategy, Terms and Conditions for Transfer, and Preservation Action Plans. The Assessment activity (A1.3.3) could assess the information about transferred and accessioned records and update the preservation strategy and actions plans without having to re-perform activities A1.3.1 and A1.3.2. However, when we got around to figuring out how the preservation strategy and preservation actions plans would be used in A3.3, Update Digital Components, it was realized that some activity must be monitoring changes in the State of Information Technology (new computers, new operating systems, new formats, new standards) and this activity (-ies) would need to revise strategies, acquire targeted preservation methods (viewers, format converters) and then some activity would need to trigger (or plan) retrieval of types of digital components needing preservation actions and trigger the preservation acts themselves. In other words, a walkthrough needs to show the preservation planning and execution process (sequence of a activities) for selected records, for transferred records, and for maintained records in the face of changing technology. It may be that the activities of A1.2 and A1.3 are involved in all three processes. The key question is: Which activity monitors changes in Information Technology? It will trigger new options, selections, acquisitions, and preservation actions.

<sup>v</sup> For instance, develop general preservation framework, develop preservation strategies for specific bodies of records, and then one to push the triggers for preservation actions.

<sup>vi</sup> Should there also be a feedback from A1.3.3 to A1.3.1? If so, to be consistent, A1.3.3 should have one output, assessment, which feeds back to both A1.3.1 and A1.3.2. The updates are then carried on the existing strategy and plan outputs.

<sup>vii</sup> There is an output of A1.4, Evaluation of Execution, which loops back to control A1.3. In converting the IDEF(0) diagrams to PDF, that feedback loop was superimposed on the Target Preservation Method that is output from A1.2. The overlap should be fixed. A1.4 produces as an output Evaluation of Execution that is fed back to A1.3, but it is not shown as a control on A1.3.3 where it would be used.

<sup>viii</sup> Preservation Action Plans in A2.2 are used to trigger planning actions for verification.

<sup>ix</sup> Activity A2.2 only addresses the information about the transferred records. Preservation Action Plans should not be input to A2.2, but only used in A2.3 where the objects of the transfer are examined.

<sup>x</sup> This implies that the Terms and Conditions for Transfer must be an input to A2.2, or that it is a part of the Preservation Strategy, which is a control on A2.2. It also implies that activity A1.3.2 should take much of the information this in the "Information about Records Selected for Preservation" and put it into the "Terms and Conditions for Transfer." It is suggested that the Preservation Strategy that is used as a control on this

---

activity, and most of the other activities in A2, A3, and A4, be just a (Generic) Preservation Strategy (or Preservation Framework) and that Terms and Conditions for Transfer be a specific input for each Transfer.

<sup>xi</sup> The reason that Preservation Action Plans are needed in A2.3 is that the creator, as part of the Terms and Conditions of Transfer, may not have been required to perform all preservation actions that are required by the preserver's Preservation Strategy for that particular body of records.

<sup>xii</sup> The arrows in A2.3.2 and A2.3.3 directed from the lower edge of the activity box to the bottom of the page represent links to activities included in other diagrams. Why is the arrow labeled A4, Output Records, directed out of A2.3.2 needed? Digital Components of Records that cannot be Preserved are forwarded to A2.3.3. Why is the arrow labelled A3.3, Update Digital Components, needed? A2.3.3 has all the resources that are needed to update the Digital Components, without going to A3.3. The resources consist of the Preservation Action Plan and Targeted Preservation Methods. This may entail making Targeted Preservation Methods a control on A2.3.3, unless Targeted Preservations Methods are part of the Preservation [Control] Strategy.

<sup>xiii</sup> In the definition of the output "Mapped Records and Digital Components," the clause "and determining whether the information about the records and the digital components in the transfer is appropriate and sufficient for preservation and reproduction of the records" should be deleted because that function is performed in the subsequent activity A2.3.2, Verify that the Records in the Transfer Can Be Preserved and Reproduced.

<sup>xiv</sup> It may be necessary to have the Conforming Transfer as an input, because the Mapped Records and Digital Components are just a map, not the records and digital components themselves. Furthermore, the definition of A2.3.2 states that the basis for asserting the authenticity of the records as transferred will be examined, and this requires the Conforming Transfer, not just the map.

<sup>xv</sup> The Basis of Authenticity of Transferred Records does not need to be a separate input to A3.1 because it is a part of the Information about Accessioned Electronic Records (see description of inputs for A1 in this document).

<sup>xvi</sup> Plan for Updating Digital Components is an input to A3.1.3 so that components can be retrieved for updating.

<sup>xvii</sup> Which activity triggers the activities in A3.3, Update Digital Components? Is it activity A1.3.2, Plan for Implementing Preservation Strategies? Instead of Preservation Action Plans being inputs, perhaps these are the Intervention Triggers that occurred in prior versions of the preservation model.