Fedora and the Preservation of University Records Project

1.3 Concerns

Version
1.0

Date
September 2006

Digital Collections and Archives, Tufts University
Manuscripts & Archives, Yale University

An Electronic Record Research Grant funded by the
National Historical Publications and Records Commission

Digital Collections and Archives
Tisch Library Building
Tufts University
Medford, Massachusetts 02155
http://dca.tufts.edu

Manuscripts and Archives
Yale University Library
Yale University
P.O. Box 208240
New Haven, Connecticut 06520-8240
http://www.library.yale.edu/mssa/

Co-Principle Investigators
Kevin Glick, Yale University
Eliot Wilczek, Tufts University

Project Analyst
Robert Dockins, Tufts University

This document is available online at
http://dl.tufts.edu/view_pdf.jsp?urn=tufts:central:dca:UA069:UA069.004.001.00003
(September 2006)

Fedora and the Preservation of University Records Project Website at
http://dca.tufts.edu/features/nhprc/index.html

Funded by the
National Historical Publications and Records Commission
Grant Number 2004-083

Fedora and the Preservation of University Records Project

TABLE OF CONTENTS

OVERVIEW

This project makes a set of intellectual assumptions, or overall concerns, that inform all of its findings. Every requirement and step described in the Requirements for Trustworthy Recordkeeping and Preservation the Ingest Guide, and the Maintain Guide implicitly carry with them the following nine concerns: Audit, Authorization, Automation, Compliance, Documentation, Financial Sustainability, Metadata, Reporting, and Training. In order to meet the requirements or undertake the steps of the guides in a trustworthy manner, an Archive or institution must address all nine concerns.

AUTHENTICITY

The goal of records preservation is to physically and intellectually protect and technically stabilize the transmission of the content and context of electronic records across space and time, in order to produce copies of those records that people can reasonably judge to be authentic.

Authenticity is the trustworthiness of the record as a record—that the record is what it purports to be and has not been tampered with or corrupted in essential respects. A person cannot automatically presume the authenticity of an electronic record; he or she must weigh the evidence that the record either is or is not what it purports to be and either has or has not been modified or corrupted in essential respects—and then judge whether the record is authentic or not. Authenticity is not a component of a record but the judgment a person makes about a record. When reports from this project report refer to "authentic records" or "authentic electronic records" they are using shorthand for records that a reasonable person would judge as authentic.

In order to be able to reasonably judge a record as authentic, one must be able to establish its identity and demonstrate its integrity. One must ensure that electronic records are clearly identifiable, of demonstrable integrity, and that accidental corruption or purposeful tampering has not occurred since they were created and set aside. One can accomplish this by maintaining the records in a trustworthy electronic records system. A trustworthy records system ensures the preservation of a record's identity and integrity, protecting it from corruption and tampering. Therefore, a record created/captured and managed in a trustworthy records system can be presumed to be authentic. A person judges the trustworthiness of records systems and the authenticity of records.

Several recordkeeping and preservation requirement sets developed primarily by the archival and records management professions over the past two decades have specific requirements devoted to authenticity. However, authenticity—creating the likelihood that a reasonable person would judge records as authentic—is the goal of a records system; it is not a required attribute or activity of a system. Therefore, none of this project's reports explicitly discuss authenticity as a specific requirement or step, instead it is the implicit, core aim, and purpose of every requirement and step.

LIST OF CONCERNS

Below are concerns that permeate every aspect of recordkeeping and preservation in an effort to reach the ultimate goal of reproducing authentic copies of records. Many recordkeeping and preservation requirements developed in the last twenty years express these concerns—especially audit, compliance, and metadata—as specific requirements related to but distinct from other requirements. However, unlike requirements or steps related to relatively discrete functions (such as records capture), requirements like audit, compliance, and metadata are not distinct functions, they are instead concerns that permeate every function. For example, records capture, disposition, and delivery are all functions that must be auditable.

Audit

Every action taken in a records system to create, collect, organize, categorize, maintain, preserve, retrieve, use, or execute the disposition of a record must be auditable. This means that every action must produce an account of itself that an external entity can audit, and the records system must support a process that can execute audits. A records system must ensure that actions performed on records, their metadata, and the system itself are auditable. Institutions must keep unalterable audit trails and preserve those audit trails for as long as the appropriate auditor requires them for review. Audits should reveal information on the nature of the action, the entity undertaking the action, and the time of the action.[1]

Authorization

Every action taken in a records system to create, collect, organize, categorize, maintain, preserve, retrieve, use, or execute the disposition of a record must be undertaken by a person or unit within an institution, or a designee of the institution, who has the authority to undertake that action. This includes a person's or entity's authorization to view records (read access) and a person's or entity's authorization to take actions upon records (write access). A person's or entity's authorization to view or take actions upon records is based on a person's or entity's rights, security clearance, position within an organization or society, or training. For example, a citizen has the right—the authorization—to view non-classified government records because of various government laws. A person's or entity's authorization to view records and in particular take actions upon records also depends on policies, procedures, and the state of records. For example, a person who is authorized to confidentially destroy certain records is only authorized to destroy those records when they have exceeded their retention period and they have had the proper final review. Institutions need to document, monitor, enforce, and update their documentation of authorizations, which can become quite complex.

Automation

Many actions taken in a records system to create, collect, organize, categorize, maintain, preserve, retrieve, use, or execute the disposition of a record must be able to be undertaken in a automated manner that would enable an institution to implement the action in a scalable

[1] Recordkeeping requirement sets that extensively discuss auditing include Indiana University, *Requirements for Electronic Records Management Systems*; *Model Requirements to Ensure the Creation, Maintenance, and Preservation of Electronic Records,* prepared for the IDA Programme of the European Commission by Cornwell Affiliates; and *Design Criteria Standard for Electronic Records Management Software Applications.*

production workflow. Archives and institutions cannot automate every action, but the more they can automate, the more they can feasibly manage the daunting volume of electronic records many institutions and Archives face. Strictly speaking, automation is not absolutely required for a trustworthy records system. However, most electronic records systems handle such a large volume of records that an entirely non-automated system could not keep up with the sea of records it is responsible for. The institution could only manage or preserve a small percentage of records in a trustworthy manner while leaving the remaining records unattended.

Compliance

All recordkeeping and preservation activities are undertaken within the context of a legal, regulatory, and administrative environment. Institutions must identify, track changes to, and comply with the laws, regulations, standards, best practices, and professional ethics that affect its recordkeeping activities.[2] People must understand the laws, regulations, standards, best practices, and professional ethics that affect their recordkeeping activities.[3] Recordkeeping applications must not include any features that do not comply with the laws, regulations, standards, best practices, and professional ethics that affect the recordkeeping activities of the institution that the application serves.[4] Institutions should be able to demonstrate their compliance with the laws, regulations, standards, best practices and professional ethics that affect its recordkeeping activities.[5]

Documentation

All records systems need documentation that describes how to execute every action taken in a records system to create, collect, organize, categorize, preserve, retrieve, use, or execute the disposition of a record. In essence, institutions need to create written procedures for all of their substantive recordkeeping and preservation activities. Institutions must determine the appropriate detail and retention of their recordkeeping and preservation documentation. This concern does not include the creation of audit trails of individual recordkeeping and preservation actions, which many research projects refer to as documentation.

Financial Sustainability

Every records system must be financially sustainable if it is to persist long enough to fulfill its recordkeeping or preservation goals. An institution needs to follow sound business practices and have long-term plans in place supported by short- and long-term financial planning. Ensuring preservation of digital resources requires substantial and ongoing financial commitments over time—potentially more so than for traditional records. Electronic records preservation is dynamic; responses to technological obsolescence or media decay must be taken quickly and the life expectancy of a preservation treatment is short because the technologies utilized evolve rapidly. Consequently, preservation strategies must be periodically monitored and reassessed as the technological environment that supports standards, protocols, and formats, etc. evolves.

---

[2] University of Pittsburgh, *Functional Requirements for Evidence in Recordkeeping* 1a, 1a1-3, 1c; Indiana 1.l, 1.1.1; MoReq 11.4, 11.5, 11.5.2-3, 11.5.5; Public Record Office, *Functional Requirements for Electronic Records Management Systems* A.10.1, A.10.2; *ISO 15489-1: Information and documentation – Records management* 5, 5a-e, 7.1.h, 8.2.4.

[3] ISO 8.2.4.

[4] MoReq 11.5.4.

[5] Pitt 1; ISO 5, 5a-e, 8.2.4.

Institutions must be able to bear the financial cost of any recordkeeping or preservation activity to ensure that it can devote to that activity the resources needed to ensure its trustworthiness.[6]

Metadata
Many of the actions taken in a records system to create, collect, organize, maintain, categorize, preserve, retrieve, use, or execute the disposition of a record along with the audit trails and documentation of these actions manifest themselves as metadata. Encoding actions as metadata is critical for enabling records systems to have regularized, machine-readable, automated, and scalable workflows. Metadata is critical for documenting actions for audit purposes and generating timely and accurate reports for records system administrators and managers.[7]

Reporting
Records systems must be able to produce reports on most of the actions taken to create, collect, organize, maintain, categorize, preserve, retrieve, use, or execute the disposition of a record for records systems administrators and managers. These reports come in many forms and may be machine-readable or human-readable, delivering information to a service or person. Reports may describe individual actions or aggregate many actions. Reporting capabilities enable managers and administrators to manage records systematically, monitor records and usage, detect problems such as data failure or unauthorized access, and plan resource allocation and preservation strategies.

Training
Every person undertaking an action in a recordkeeping or preservation system to create, collect, organize, categorize, maintain, preserve, retrieve, use, or execute the disposition of a record must have the appropriate training needed to execute the action successfully. A person must be trained to a level that allows that person to undertake an action in a records system in a trustworthy manner. A person's breadth and level of training may be closely tied to his or her authorization to perform recordkeeping or preservation activities.

[6] *Trusted Digital Repositories: Attributes and Responsibilities* (Mountain View, CA: RLG, 2002) and *An Audit Checklist for the Certification of Trusted Digital Repositories*, Draft for Public Comment (Mountain View, CA: RLG, 2005).
[7] Several recordkeeping requirement documents have significant requirements concerning metadata, including, Indiana; MoReq; PRO; and DoD 5051.2. The *Data Dictionary for Preservation Metadata: Final Report of the PREMIS Working Group* came out in 2005 as a set of preservation metadata.