Fedora and the Preservation of University Records Project

1.2 System Model

Version
1.0

Date
September 2006

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

Project Analyst
Robert Dockins, Tufts University

Fedora and the Preservation of University Records Project

TABLE OF CONTENTS

ISSUES WITH THE PRESERVATION OF ELECTRONIC UNIVERSITY RECORDS

The essential nature of the modern office at colleges and universities—complete with hybrid paper/electronic systems, digital environments established to support manipulation and repurposing of data at the expense of recordkeeping, obsolescence of hardware and software, media decay, the proprietary and idiosyncratic nature of applications, and other problems— makes it difficult for archivists to provide for the long-term preservation of authentic electronic records and maintain the accountability of the organizations and operations which those records are supposed to document. This leads institutions to create and maintain electronic records that they cannot automatically trust and depend on in the same way that institutions trust and depend on traditional paper records. In general, archivists have difficulty preserving electronic records that fail to be (1) accessible, readable, or intelligible due to compatibility and obsolescence issues; (2) identifiable and retrievable due to an incongruence of classifications and/or taxonomies; and (3) reliable in the accuracy of their content due to the ease of updating and altering records, either inadvertently or purposefully.

In order to address these issues, an organization must recognize that the goal of electronic 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. This is a continuous process that begins even before the moment of creation, and pervades every single recordkeeping activity.

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 refer to "authentic records" or "authentic electronic records" it is shorthand for records that a reasonable person would judge as authentic. One cannot judge the authenticity of a recordkeeping or preservation system, only its trustworthiness.

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 records system. A trustworthy records system ensures the preservation of a record's identity and integrity, protecting it from corruption and tampering. Therefore, a reasonable person can presume that a record created/captured and managed in a trustworthy records system is authentic.

The archival, records, and information management communities have used the terms electronic recordkeeping system in a number of ways. Some conceive of a recordkeeping system broadly as the entire framework of recordkeeping from creation to preservation and access, while others

conceive of it more narrowly, describing the specific computer application tasked to store and manage records. In addition, the elements that make up a recordkeeping system and the factors (controls) that influence recordkeeping have been described interchangeably. The imprecise use of the term electronic recordkeeping system in documents articulating requirements for such systems may have hindered records professionals' efforts to turn these documents into evaluation tools and detailed implementation and application development guidelines. In order to alleviate this problem, this project has explicitly defined a trustworthy electronic recordkeeping system and the composition of its elements. In addition it has also differentiated between the components of a records system and the records controls that influence and shape an institution's recordkeeping activities.

TRUSTWORTHY ELECTRONIC RECORDKEEPING SYSTEM

A trustworthy electronic recordkeeping system is the combination of all of the records components—people (natural and juridical), institutions, applications, infrastructure, and procedures—necessary for records to be created, collected, organized, and categorized to facilitate their preservation, retrieval, use, and disposition in a manner that provides a circumstantial probability of the authenticity of those records and a likelihood that a reasonable person would judge those records as authentic. An institution uses a combination of records components to help it meet its records needs and expectations that are articulated or manifested in requirements, policies, responsibilities, and practices.

INFLUENCE OF THE LIFECYCLE MODEL

The research work of this project is based largely on the conceptual underpinnings of the records lifecycle model, presuming that a Producer will create, acquire, use, and manage records in a recordkeeping system to suit its current business needs, and later the Archive will ingest some of those records into a separate preservation system that the Archive administers. In this model, the Archive acts as a neutral third party in the recordkeeping process, acting first and foremost on behalf of broader societal needs rather than on behalf of the Producer. As a neutral third party, the Archive has no stake in the content of the records and no reasons to alter records in its custody, and it should not allow anybody to alter the records either accidentally or on purpose. Many archivists have rejected the lifecycle model in favor of the records continuum concept, where recordkeeping is seen as a continuous process that is not time-based, separated into a series of clearly defined steps, or administered by completely separate juridical entities. Many Producers and Archives operate in a mixed world between these two models. For example, many Archives operate separately from a Producer, are part of the same organization as the Producer, and do not act as a neutral third party.

These two types of records systems (recordkeeping and preservation) are distinguished from one another in this project because such a distinction helps best describe the records environment of many colleges and universities in the United States. However, separate recordkeeping and records preservation systems are not necessary to make up a trustworthy records system at a college or university. Recordkeeping systems, because they must preserve records in order to "keep" them, naturally fulfill many of the same requirements placed upon records preservation systems. The opposite is not true. A trustworthy records preservation system cannot exist without a trustworthy recordkeeping system. It is not possible for a records preserver to preserve authentic records if such records were not created and managed in trustworthy recordkeeping systems fulfilling the recordkeeping requirements.[1]

---

[1] It would be possible for a preserver to preserve authentic copies of records whose authenticity might be in question. For example, the preserver can not alter a forgery so that it can be presumed authentic, but the preserver can faithfully preserve authentic copies of that forgery.