



TAPER
TUFTS ACCESSIONING PROGRAM FOR ELECTRONIC RECORDS

Fourth Interim Narrative Report
Project Activities July 1, 2009 – December 31, 2009

Submitted
February 17, 2010

NHPRC Grant Number
RE10005-08

Krista Ferrante
Project Archivist

Eliot Wilczek
Project Director
University Records Manager



INTRODUCTION

The Digital Collections and Archives (DCA) at Tufts University is undertaking a National Historical Publications and Records Commission (NHPRC) Program Expansion Grant project to design, develop, and implement machine-readable submission agreements and records creator records. Work for this project started on January 1, 2008 and is scheduled to end on December 31, 2010.

SUBMISSION AGREEMENT SCHEMA

Based on an element list the DCA developed, Appleseed Software Consulting created a draft submission agreement schema in March 2009. The DCA and Appleseed have tested the schema in the submission agreement builder tool (SABT), which will serve as the operational environment for the schema.

Subsequently, Appleseed made minor revisions to the schema. The DCA plans to release a 1.0 version of the schema when the SABT goes into production.

SUBMISSION AGREEMENT BUILDER TOOL

Development of the SABT is nearing completion. The SABT will serve as the online environment for records creators and DCA staff to complete submission agreements. Appleseed has used the functional requirements outlined by the DCA to create the SABT. Appleseed produced a development version of the tool in September 2009. Since then, project staff have conducted testing of the various components of the tool. After each round of internal evaluation, the DCA has provided Appleseed with feedback, and in turn the consultant has made adjustments to the tool. The DCA's prime criteria for the SABT requirements have been ease of use and sustainability. The project staff's goal is to create a tool that can provide the DCA with enough information to fully document accessions and forecast the arrangement and description challenges that any set of accessioned archival records may present to the DCA. The test version is currently located at <http://moby.tufts.edu/taper>. The project staff anticipates that the DCA will be able to make the SABT a production service by April 2010.

Developing the requirements along with feedback and questions from Appleseed have prompted the DCA to have several internal discussions concerning how the department will implement the SABT and integrate the tool and submission agreements into its broader accessioning and arrangement and description processes. This included making several decisions regarding the DCA's descriptive metadata creation workflow and the architecture of its metadata management environments.



RECORDS CREATOR RECORDS

The DCA plans to have the submission agreements reference externally defined activities, resources, rules and standards. Within the scope of the TAPER Project, the DCA is developing records creator records as one of the resources referenced by a submission agreement. The DCA will use the records creator records to identify users of the SABT and to populate the submission agreements with accurate information about the entities transferring archival records to the DCA. Specifically, the DCA is able to generate lists from the records creator records that can be imported into the SABT. The DCA is also using the University's LDAP directory to populate the submission agreements with accurate information about records producers that log-in to use the SABT. Accurately identifying the records creators will help the DCA predict and plan for records it is suppose to accession. More broadly, records creator records will enable the DCA to improve the flexibility and quality of its archival description. The DCA documented its incorporation of records creator records into the TAPER Project in its August 12, 2008 revised plan of work.¹

SCHEMA

In order to develop and implement records creator records that will store information about records creators—particularly those that transfer records to the DCA, the project staff has investigated standards such as ISAAR (CFP) and EAC-CFP along with institutional strategies for their implementation. After careful evaluation, the DCA has decided to use the EAC-CFP schema instead of developing its own schema as it had originally planned.

The DCA came to this decision for many reasons:

1. Compatibility with MARC authority records. Project staff have met with librarians from other Tufts libraries about generating local authority files for individuals and departments at Tufts. Since the other libraries do not have local MARC authority records describing Tufts entities, they have expressed interest in the records creator records and their compatibility with MARC.
2. Ability to draw relationships between records creator records in order to demonstrate context.
3. Ability to draw relationships between records creator records and collection material in order to improve archival description.
4. The close adherence to the standards outlined in ISAAR (CFP). ISAAR (CFP) provides a robust framework for describing corporate bodies, families and people. EAC (CFP) offers the ability to build a description in XML that follows that standard, while also providing enough flexibility for local implementation.

¹ "Revised Plan of Work" August 12, 2008, http://dca.tufts.edu/downloads/plan-of-work_revised_tufts.edu. Document listed at <http://dca.tufts.edu/?pid=136>.



5. Archival collection management tools such as ICA-AtoM are being developed to interoperate with the EAC (CFP) schema and adhere to the ISAAR (CFP) standard.
6. EAC (CFP) has the support of the archival community and there are many early adopters anxious to see this XML schema sustained.

The development of EAC-CFP therefore presents this project with several advantages. The DCA has also begun to engage with the EAC-CFP development community, providing comments and feedback about the schema to the SAA EAC-CFP working group.

MANAGING RECORDS CREATOR RECORDS

The DCA has been reviewing management environments for the records creator records. Currently, the DCA is keeping the records creator records in a spreadsheet. As project staff continue to identify more and more creators across the University this has become harder to manage. Project staff are working to move this information into the DCA's current collection management system, ReDiscovery. Since ReDiscovery was designed for creating hierarchical archival descriptions it is not ideal for building records creator records. However, it provides the advantage of keeping all of the DCA's description in the same place, which facilitates the DCA's integration of the records creator records into its broader archival description workflow.

Project staff have also looked into using other open-source tools, including ICA-AtoM to manage the creator records. ICA-AtoM would be an ideal environment for the records creator records because it has a module specially designed for describing corporate bodies, families, and persons. This module strictly adheres to ISAAR (CFP), is web based, allows patrons to directly browse the descriptions online, and provides direct authority links to the collection descriptions. Unfortunately, ICA-AtoM does not currently have a mechanism for exporting records creator descriptions. In addition, the ICA-AtoM database also lacks to ability to tag select records for export. These are both functions that ReDiscovery provides and that the DCA has come to rely upon in order to ingest its descriptive collection material into its Fedora-based digital repository. The ICA-AtoM working group does have plans to have the records creator records export into EAC (CFP) but does not currently have plans to provide a tagging mechanism. The DCA is exploring the feasibility of developing this tagging functionality on its own.

RECORDS CREATOR RECORD PRODUCTION

The DCA is reviewing several internal and external sources to identify various Tufts entities that should have a records creator record. This work includes prioritizing which records to create during the course of the TAPER project. Unfortunately, the University does not have a single authoritative list of departments and offices. Instead the DCA has relied upon a variety of resources, including:



- The DCA’s collection management system.
- The DCA’s collection and donor documentation.
- Department listing from HR and Purchasing.
- Listing of “Tufts University” as an author in Tufts’ library catalog.
- Library of Congress Authority Headings relating to Tufts.
- Published or unpublished department histories.

The descriptions of records creator records that project staff have derived from these sources have varying levels of detail. There are three types of Records creator records with varying levels of descriptive detail: Stub, Full, and Cluster records. A stub record only has a few basic fields; authorized name, unique identifier, dates of existence, sources. These constitute the mandatory fields for any records creator record. A full record will usually have a majority of its fields completed—capturing all of the reasonably available information about a records creator. A cluster is a series of full records with complete information detailing the relationships between them. Clusters are difficult to define because all of the Tufts entities described by the creator records are technically connected to its governing organization, Tufts University. For the sake of clarity, project staff have defined a cluster as a group of two or more full records with one or more direct connections that do not include the connection to Tufts University.

Due to the time consuming challenge of finding the dates of existence the majority of creators are simply listed by name. Project staff currently has approximately 560 creators listed. 18 of these creators are stub records, and 10 are full records. These 10 full records comprise two clusters.

SAA SESSION

Krista Ferrante, Project Archivist, organized an accepted panel for the 2010 Society of American Archivists annual conference in Washington D.C. The session is entitled “Tools and Processes for Trustworthy and Scalable Electronic Accessioning.” The session speakers are Mark Myers, Kentucky Department for Libraries and Archives; Leslie Johnston, Library of Congress; and Krista Ferrante.

WEBSITE

The DCA posted the following items to the TAPER Project website, which is at

<http://dca.tufts.edu/?pid=49>:

- Project Documentation <http://dca.tufts.edu/?pid=136>
This includes project application materials, interim reports to the NHPRC, and some working documents.



- Deliverables <http://dca.tufts.edu/?pid=137>
Project staff will post deliverables on this page.
- Announcements, Presentations, Articles <http://dca.tufts.edu/?pid=138>
Project staff will post project announcements, presentations, and articles on this page.
- Project Personnel <http://dca.tufts.edu/?pid=135>.