

Library Lab Proposal – October 21, 2011

Harvard University Archives (HUA)

Megan Sniffin-Marinoff, University Archivist; David Read, Records Analyst

NOTE: Significant revisions to the proposal are noted in bold italics throughout. Changes include: the addition of explicit activities in Phase I to reach out to peers at Harvard, MIT, and Tufts; annotations to the software demo; and the inclusion of software, added to the cost share for the HUA.

Enhancement of Open Source Software for IRMA (Improved Records Management Administration)

The Harvard University Archives (HUA) is applying for funding for a pilot project to take the first steps to continue development and implement into production at Harvard an open source records management tool to enhance user and system access to records retention schedules and assist in the identification, capture and management of University records in all formats.

Records Management, as defined by the Association of Records Managers & Administrators (ARMA), is “the systematic control of records throughout their life cycle.” The proactive and systematic identification, capture, and management of University records provide the policies and framework which build the core of the permanent archival collections at Harvard. Equally important, records management significantly enhances the University’s operational efficiencies and assures the demonstration of regulatory compliance. The Records Management Services (RMS) unit of the HUA, assists (either directly or indirectly and variously through other records management units) the over 6,000¹ University staff and administration whose day – to – day activities keep the University running, with the identification, management and preservation of appropriate University records.

Records management responsibilities and requirements at Harvard must be made easier for University administrators and staff to understand, access, and employ. Currently the University records retention schedule (see definition below) access interface has limited browsing and search capabilities and provides user access to only a portion of the full schedule set of University records retention schedules, as office specific schedules are not currently offered online due to confidentiality restrictions. Efforts are underway at the University Archives to enhance records management education and outreach significantly. We must, however, address better access to the University records retention schedule for University staff and administration as well as for the systems creating and managing these records.

A retention schedule is a policy document that lists the names of record series (groups of like records) produced or received during University business, along with their retention periods and disposition (destroy or transfer to archives). Enhancing access to the University records retention schedule will allow the University to better employ these university policies, which will facilitate regulatory compliance and better identification, management, and capture of permanent and non-permanent University records in all formats. Ultimately, the capture of these permanent records is the way in which University collections of archival records are successfully built over time – a significant difference from the way in which library collections are developed and collected. As we enter the 375th year of documenting Harvard history with accelerating changes in record-keeping practices, streamlining this activity is more important than ever.

RMS is in the process of moving from a legacy model of user engagement and assistance that was principally reactive, manual, and focused on a slow cycle of retention schedule building towards one of:

- Proactive user engagement

¹ Staff and administration count increases to approximately 10,200 when HBS and HMS are included. Harvard University Fact Book, 2010-2011.

- Automation
- Regular updating of records retention schedules
- Innovation

Current examples of this shift in approaches include:

- Reaching out to users and hosting regular record management workshops: (Requires authentication) <http://isites.harvard.edu/icb/icb.do?keyword=k7521&pageid=icb.page317626>
- Active leadership and participation by RMS staff in innovative website and email archiving projects :
Email archiving: (requires authentication) <http://isites.harvard.edu/icb/icb.do?keyword=k38759>
Website archiving: <http://hul.harvard.edu/ois/systems/wax/>
- Shift to online retention policies

In line with this current shift to proactive user engagement, automation and innovation, the HUA is proposing a project that will address key problem areas that will:

- Engage records creating offices earlier and more actively with RMS with an online survey (pre-interview) tool
- Provide better user access to University Records Retention Schedules (multiple points of access to schedules / true full text search / examine adding online access to office specific retention schedules)
- Create efficiencies for Records Management Services (web based approach to user survey and interview data collection and management)
- Improve University risk management and compliance as related to records (better user and systems access to records retention metadata)
- Proactively bring permanent records to the Archives (better user and systems access to records retention metadata)
- Examine metadata sharing between records management and permanent archives collection systems (Application Programming Interface (API) or standardize export/import)
- ***Examine authority naming for University departments using Encoded Archival Context (EAC)***
- Examine records retention metadata sharing between RMS and University information systems that create and manage electronic records (Service Oriented Architecture [SOA] approach to electronic records management)

The proposed project will build on a current open source software project (“Records Authority”) funded by the National Historic Publications and Records Commission (NHPRC) and based at the University of Denver (DU). This project will be an undertaking in user driven innovation by taking University records management stakeholders feedback to push the NHPRC project forward with enhancements to both user and systems access to records retention metadata, laying the foundation for true electronic records management at Harvard University. The benefit to the University of this undertaking will be better user and systems access to and employment of university records retention policy. This will in turn enhance compliance and the identification, capture and management of both permanent and non-permanent University records in all formats. The Records Authority project was developed initially by David Read, a current Harvard RMS team member, who served as the initial project Principal Investigator while at DU. This project and tool is being selected by HUA because as NHPRC has concluded by it’s funding of the project, the tool fills a void for a low cost, flexible and web based solution for user engaged and systematic retention schedule building and access.

For further background and current NHPRC project deliverables, including the slides and software demonstration presented at the Society of American Archivists (SAA) annual conference on 8/27/11 by current RMS team member David Read, see:

- <http://library.du.edu/site/about/urmp/recordsAuthority/grantInformation.php> (a compilation of the following three links)
- NHPRC project Deliverables: <http://library.du.edu/site/about/urmp/recordsAuthority/grantInformation.php>
- SAA session abstract (link to demo) and presentation slides: http://saa.archivists.org/Scripts/4Disapi.dll/4DCGI/events/eventdetail.html?Action=Events_Detail&InvID_W=1827 [See David Read PowerPoint]
- Software Demo from SAA 2011 presentation: http://www.youtube.com/watch?v=pJgQyw_nNrc

During the initial Investigation and Development in Phase 1, we will contact our partners in the Harvard Library who also handle records management at the local level (such as the medical and business schools) to discuss what their approaches to these issues have been and why. This is important as we consider the ways that this tool could potentially be customized at Harvard. We will also reach out to the Institute Archives at MIT (our current partner in the Library Lab Zone 1 Project) as well as the records manager at Tufts University to discover how they may be addressing parallel issues in other institutional contexts.

Resources

Archives: .5 FTE Project Manager/Analyst (David Read currently on staff through April 2012. This would require an extension of his current appointment, which is anticipated) David will continue during the grant period to have as a priority the writing of record schedules and other RMS team assignments while taking on this pilot project. Other duties would be covered with a 50% short-term employee or another alternative for short-term coverage. Input and assistance from a variety of other HUA staff from records management, collection services, and collection development will also be required.

Responsibilities for Project Manager/Analyst on project would be to: Coordinate the pilot and serve as point of contact with testing pilot group and other RMS staff. Lead weekly project status meetings with project developer. Collaborate with project developer on analysis report of current NHPRC beta tool to meet Harvard needs and customize tool to Harvard environment. Lead policy and access study for providing access to office specific retention schedules online. Collaborate with project developer and other University Stakeholders on study of potential of current/future/potential shared services approach (Service Oriented Architecture [SOA]) at Harvard. Write final assessment of benefits of project for all stakeholders.

Budget request: 50% salary/benefits coverage for release time (his salary then re-allotted for other short-term coverage of regular duties)

Additional HUA cost-share: Camtasia screen casting software for stakeholder demo: \$300

Library Lab: .5 FTE Developer with Linux, Apache, MySQL, PHP (LAMP), Solr development background.

Responsibilities on project would be to: Collaborate with Project Manager/Analyst on analysis of current NHPRC beta tool to meet Harvard needs and customize tool to Harvard environment. Create remote software testing environment. Develop and implement priority software enhancements. Maintain bug tracking software. Report on development/implementation at weekly status meetings with Project Manager/Analyst. Collaborate with Project Manager/Analyst on study of potential of current/future shared services approach (Service Oriented Architecture [SOA]) to electronic records management at Harvard.

Budget request: .5 Developer time for approximately 11 months – total estimate is 24 weeks of developer time

Plan of Work

Phase 1: Investigation and Development (months 1-3)

- *Analysis of current NHPRC beta tool by Project Manager/Analyst, HUA/RMS team, and Developer to meet HUA/Harvard needs and customize tool to this environment*
- *Meet with records management peers at local offices (i.e. medical and business schools) to discuss their approaches to similar issues, determine mutual needs, examine local tools in use, and determine if this beta product and approach would translate to their environment*
- *Meet with records managers/archivists at MIT and Tufts to discuss their approaches and solutions to similar issues in other institutional contexts*
- *Identify and recruit pilot group of interested University stakeholders. Create testing environment (current software) and create testing/enhancement suggestion plan. Get stakeholder feedback.*
- *Present to Library Lab proof of concept, including report on discussions with peers*

Phase 2 Development and Testing (months 4-6)

- Development and testing of enhancements with pilot group
- Evaluate policy and access options for office specific records retention schedules
- Explore current/future/potential shared services approach to Service Oriented Architecture (SOA)
- *Explore viability for future/alternative approaches to record scheduling (i.e. “big bucket” <http://www.aiim.org/community/blogs/expert/Big-Bucket-Approach-e28093-How-The-View>)*

Phase 3 Development and Testing (months 7-9)

- Development and testing of enhancements with pilot group
- Implement enhancements based on stakeholder survey and testing
- Brand enhanced tool with our Harvard brand

Phase 4 Pilot Production Roll out (months 10-11)

- Production roll out to pilot group
- Report on measure of benefits, with follow-up to all stakeholders

Potential enhancements of the original NHPRC project at the end of one year:

- API (system/machine readable records retention schedules)
- Implementation of shared metadata approach to authority control for department names using Encoded Archival Context (EAC)
- Investigate / implement sharing data with other HUL collection management systems Archivists Toolkit (AT)
- Survey tool enhancements (follow up user surveys)
- Browse schedule by Central Admin/FAS feature
- Implementation of “Did you Mean” suggestion enhancement to Solr public search
- Personalized “Retention Schedule Builder.” Saved search type feature that would allow users to build and modify a view of records retention schedules that pertain to them, *including office specific schedules*. Some of this functionality already exists in the tool on the Admin. side: see 3:25 of software demo:
http://www.youtube.com/watch?v=pJgQyw_nNrc&feature=player_embedded#!