

Records Management and Digital Preservation using Fedora and iCalendar

Richard Green and Chris Awre

We have had the chance at previous Fedora Users' meetings to present the University of Hull's JISC-funded RepoMMan project. The vision at Hull was, and is, of a repository placed at the heart of a web services architecture: a key component of a university's information management. In this vision the institutional repository provides not only a showcase for finished digital output, but also a workspace in which members of the university can, if they wish, develop those same materials.

The RepoMMan project set out to consider how a range of BPEL-orchestrated web services could be brought together to allow a user to interact easily with private workspace in a Fedora-based institutional repository and how the web services might ease the transition from a private work-in-progress to a formally exposed object in the repository complete with metadata. That project is now complete.

The University, this time working with partners Spoken Word Services at Glasgow Caledonian University, has been fortunate to secure further funding from the JISC for the REMAP Project. REMAP takes up more or less where RepoMMan left off and exploits the BPEL-orchestrated Web Services environment to add value to digital objects at the point they are moved to the outward-facing area of the institutional repository. In particular, the REMAP toolset will build into this process a system for ongoing records management and digital preservation (RMDP).

RMDP will be supported by building into the new digital object a set of flags which will be used by an external system to trigger records management and/or preservation events. Thus, for example, when a past undergraduate examination paper has been in the repository for five years the a date flag might trigger a message to the appropriate departmental secretary that the paper is likely out of date and should therefore be hidden from future students; the secretary may wish to do this or to extend its life by resetting the flag. It may be that if the secretary does nothing the system will take a default action, in this case to hide the paper. As another example, a date flag may trigger a process to undertake a format review of the materials in the repository: this automated review, likely using external preservation Web Services in our BPEL orchestration, may result in a message to administrators that there are some specific objects in a file format that is no longer widely supported and suggest that something should be done. The REMAP team is investigating a range of possibilities, relating to repository events, dates and status, which will be explored briefly during the presentation.

We are anxious that our work should fit with existing standards as much as possible and we therefore intend to use the iCalendar standard (or xCAL, its XML compliant representation) to create the 'flags' datastream. At the time of writing we are investigating whether these flags might best be processed through a bespoke system based on a MySQL database or whether the Apple Darwin Calendar Server might provide the functionality we need; we should be able to report on this aspect of the work in April.

The work is based on a user-needs study at the University of Hull, complemented by comparative work at Spoken Word Services. This study identified a wide range of uses to which flags in a Fedora object might be put, a number of which lie outside REMAP's RMDP remit. This presentation will concentrate on the RMDP work but will mention briefly these other possibilities in the hope that others might follow them up. The REMAP Project is due to complete in March 2009.