Open Repositories 2015: Fedora Interest Group Presentation Proposal

A Fedora 4 Repository from the Ground Up at the University of Maryland

Joshua A. Westgard
University of Maryland Libraries

The digital collections repository of the University of Maryland Libraries celebrated its 10th anniversary in January 2015. While this homegrown system — built on Fedora 2 and supported by two locally developed metadata schemas — has served us well, after 10 years it is clearly showing its age. Over the years a combination of staff turnover and other, higher departmental priorities have conspired to prevent the upgrade to Fedora 3 from becoming a top priority. With the release of Fedora 4 this year, however, the stars seem to have aligned to give us the opportunity to re-envision our digital collections repository and make a fresh start with the new system.

In this presentation I will describe our ongoing efforts to design and implement a Fedora 4 based repository system in 2015. By the time of the conference in June, we will be roughly halfway through an initial year-long development process. The high-level objectives of our project are: (1) to leverage the repository improvements of Fedora 4, (2) to migrate some existing services and applications to the new system, and (3) to develop new features that were unrealized in the development of the previous repository and its interfaces.

Our approach to the development process has been to begin by standing up, as early in the process as possible, a production instance of Fedora 4, with only minimal customization (specifically just enough to enable role-based authentication against existing campus services). We will then begin to test this production instance with various sample collections of content. Our goal is to develop Fedora as a first-class service, with fine-grained access controls built in to the repository itself, though some further access controls may be added at the application layer. While at this stage we are still uncommitted to a particular front-end application, we will almost certainly adopt either Hydra or Islandora rather than opting for a purely custom solution.

Our goal is to use Fedora 4 to manage not only the full range of content types: images (including photographs as well as digitized images of both printed and manuscript materials),

audio, video, and born-digital archival materials, but also research data, metadata collections, and derivative files such as OCR, hOCR, and transcription files. Whereas our current Fedora 2 system has been used exclusively to manage *access copies* of materials in our digital collections, we intend to use the new system not only to serve access copies but also to manage multiple backup copies of our preservation assets, and to leverage a full suite of fixity checking tools. In my presentation I intend to briefly describe our experiences standing up this production instance and then devote the bulk of the presentation to describing the content models we have developed for our materials.