

Minutes of December 7, 2010 Meeting

Agenda

1. Updates and Announcements
2. Data prototype for OVPR
3. Google scholar update
4. Status of Fedora 3.2.1 and Shibboleth on mss3
5. Specification for submission to NIH
6. Thumbnails for PDFs

Updates and Announcements

- In the recent CISC meeting, Grace indicated that we have a December 15 commitment to the OVPR web. The commitment involves demonstrating that we can ingest data with appropriate metadata and provide the capabilities for access and preservation. More on this below.
- Also, from the recent meeting with Pazzani, the ability to find articles in Google Scholar is now a priority one for us. See item 2 below.

Data Prototype for the OVPR

Many of us will be involved in the next week or so to ingest data and provide access through a data portal. Grace will be receiving some data from the OVPR and we will also use the 3 compound objects that are in the data prototype on lefty64. The assumption is that this data and the prototype could become permanent so the consensus is that we should do this work on lefty64. Since we don't as yet have Shibboleth on mss3, some of the objects will need to be blocked with an embargo policy.

Google Scholar

Several months ago we initiated the Google Working Group to address the issues with indexing in Google and Google Scholar (GS) and to move expeditiously to fix the problem. The recent OVPR meeting has moved this issue to a priority one. At this point the most pressing issue is to demonstrate to Pazzani and faculty in general that their articles can be found on Google Scholar. To date, there are some 150 ETDs but very few articles from RUcore that can be found with Google Scholar. Although there are several parallel initiatives underway, this discussion focused on the need to provide a "quick fix" on mss3 in order to provide more meta-tags for journal indexing. Jeffery presented a proposal which involves adding volume, issue, and page number as meta-tags. In our review of other repositories, it appears that these tags are critical in getting GS to index articles. Rhonda will provide the update in the metadata for selected articles in order to test this feature. This will be a controlled experiment in the sense that we will be starting with zero articles indexed with GS and we will make only this one change. Rhonda will select 10 high profile articles in RUcore for markup. We should be able to deliver this fix to mss2 during the week of December 6. As part of this quick fix, we also agreed to add priority, frequency and last-modified to the sitemap. Although we don't have much data on these settings, the group felt that it would be useful to give them a try. As a part of our continuing Google effort, Jeffery will post a general question to the Fedora listserv.

Status of Fedora 3.2.1 and Shibboleth on mss3

Dave indicated that he will be ready to update mss3 on the morning of December 6. We will do a quick sanity test on mss3 before RUcore goes public. Progress continues with Shibboleth. Compilations on mss3 should be complete by Sunday, December 5 and then Dave will proceed to configuring shibboleth.

NIH Specification

Rhonda indicated that she will address the update in the next sw_arch meeting. She will arrange a meeting with R. Tedesco regarding the faculty survey.

PDF Thumbnails

We briefly discussed the process for providing thumbnails for two types of multipage objects: 1) one that originates as a set of tiff files and 2) one that originates as a pdf. Project managers for RUcore, NJEDL, and NJDH (i.e. Rhonda and Linda) agree that we can reliably use the first page of the document to generate the thumbnail. There may be some cases where the first page is not meaningful but, in general, the group thought that this would be rare. We can use ImageMagick and drive this process by content model, i.e. when implemented we should have thumbnails for books, documents, manuscripts, etc. As we did for maps, we will need to run a script to update all the existing objects. Changes in WMS and faculty deposit will be needed as well. Kalaivani will create bug fix requests for both WMS and Faculty Deposit. Although it would be great to put this fix into R5.2, the view was that everyone is pretty much occupied with completing the current planned capabilities. (In a post-meeting action, I put this capability in R5.2.1)

Agenda Items for Next Meeting

- Google Scholar
- NIH specification update
- OVPR Website and data prototype
- Planning for JPE migration
- Pending items
 - Djvu java applet
 - Performance of handle server