SW Arch Meeting Minutes – August 2, 2012

Agenda

- Announcements and Updates
- Progress on Non-release projects
- Review of file hierarchy specification
- Fedora 3.5 and R6.2 Status
- Decommissioning lefty64

Announcements and Updates

Ron reviewed the major decisions for implementing DOIs. Briefly, these are the following: 1) we will assign DOIs to all resources, including previously ingested resources, 2) DOI suffixes will be opaque and generated automatically by the EZID service, 3) we will use the EZID test prefix for generating DOIs on the develop and staging servers, and 4) assigning DOIs to specific file datastreams will require further investigation. We also discussed how to handle DOIs for preprints and published articles. From a preservation perspective, there is justification for archiving the published article in RUcore although it is not clear whether we can legally do this. Assuming we do archive the published article, the user will have two DOIs and three possible renditions of the article as follows: a) RUcore DOI which provides access to both the preprint and the published article and b) the publisher DOI. Although not an architectural issue, there is considerably more discussion is required on how the workflow for journal articles.

There was a brief discussion of the planned network outage on August 12. After getting an update from Dave, we decided to post message alerts on RUcore portals and also on the websites for the five journals published by RUL. We also noted that there was a need to improve communication between IIS and the Software Architecture Working Group on these major events.

For the Fall semester, we will stay on our regular two week meeting schedule. The next meetings are as follows: Aug. 16, Aug. 30, Sept. 13, Sept. 27, etc.

Status of Non-release Projects

Because of priorities to prepare for the network outage on August 12, the jpeg thumbnail and JPE video projects are still pending. Jeffery will verify that an active policy statement applying to another datastream in the object will not block the addition of a jpeg thumbnail. (Note, this was verified on the development server later in the day). Jeffery will also resend to Dave the list of IDs for the JPE videos that need to be pulled into the production server. Dave noted that the week of August 6 will pretty much be dedicated to developing the plan and providing support for critical library services during the network outage on August 12.

File Hierarchy Specification (previously known as Complex Objects)

Chad reviewed the file hierarchy specification, covering options for capturing hierarchy and order and reviewing additional scenarios. The decisions related are briefly summarized:

- We will proceed with Option 2, recording hierarchy in the ID attribute, i.e. ID="1.1", ID="2.1", etc. which appears to offer best approach for ease of reading.

- It was noted that we will need a WMS capability to edit/replace datastreams in order to preserve the integrity of the file hierarchy structural map. Generally, it would not be advisable to allow someone with a general purpose xml editor to edit the structural map because of the possibility of incorrectly editing the ID or ORDER fields.

- For directories on the local server, the structure can be readily discerned in order to produce the structural map. We agreed that directories originating on client workstations should be tar-ed. WMS will un-tar these directories in order to produce the structural map.

- Regarding directories with mixed file types, we agreed that the research data scenario is the highest priority. In this scenario, all files can be considered archival and re-formatting for preservation is needed. We will use the exif tool to capture technical metadata for these file types however manual effort will be required for those file types not supported by exif.

- We agreed that WMS will create a structural map for all objects regardless of whether or not there is a file hierarchy present.

- We acknowledged that it was important to support a table-of-contents view for books and possibly other content models. We did implement an early TOC disseminator which used the structural map in the SMAP1 datastream.

Chad will update the specification to reflect the above decisions and minor revisions resulting from the addition of several more scenarios and minor changes in the user interface for downloading datasets.

Fedora 3.5 on Staging and Testing of R6.2

To proceed with R6.2 Dave will roll back Fedora 3.5 on staging and re-install Fedora 3.4.2. We will then test R6.2 on staging. However, give the priority regarding the network outage, this work will not start until the week of Aug. 13. After we have tested R6.2, Dave will then re-install Fedora 3.5 so that we can resolve the remaining issue related to policies.

Other Items

Dave pointed out that objects with lower case IDs (i.e. pdf-1 vs. PDF-1) are creating errors and statistics are not being recorded for these objects. Jie and Chad will investigate and resolve this issue. Dave also pointed out that, with the decommissioning of lefty64, we will need to re-point handles and

also edit the object metadata to reflect this change. Lefty64 is in need of a space cleanup; Isaiah will review and delete unnecessary files.

Agenda Items for Next Meeting

- Status of Non-release projects
- Status of R6.2/Fedora 3.5 installation and testing
- Object definitions - review
- Decommissioning lefty64
- Pending agenda items
    - RUetd – WMS update
    - Enhanced UI for the landing page (a possible framework) – Chad and Jeffery
    - WMS – creation of relationships (rels-ext) for data projects

rcj – 08/017/2012