Software Architecture Working Group   **February 17, 2009**

Minutes of February 12, 2009 Meeting

Agenda

1. Quick updates and Announcements
2. R5.0 status
3. Finalize ETD spec
4. Review update of annotation spec
5. Compile/Config options (brief discussion).

Announcements and Quick Updates

- There are many new collection coming into RUcore/NJDH including material related to Peter Still (a slave), newspaper photographs from Camden dating back to the 1800s, Griffiths materials, Civil War journals, a new e-journal – New Jersey History, and more faculty collections.

- In the recent NJVid meeting, it was decided to add Flash streaming as a presentation format. This capability would not be available until R5.1.

R5.0 Status

Yang reported that WMS testing has begun and he is investing effort in fixing bugs. He also indicated that ingest is working but is also generating errors. Jeffery and Yang will explore the source of the error message. Yang has yet to address WMS edit and the use of findObjects. In testing, it was reported that there are problems accessing Djvu and handle creation is not working. The Djvu environment should be moved to lefty64. Jeffery and Sho will work on these issues as a high priority. We also discussed the creation of the xml file containing the sha1 signature. Jeffery has written a script to create this file, however we decided that we couldn't insert this process in-line with ingest because of the time to compute signatures for large files (as much as 2 minutes). As a manual process for R5.0 video ingest, Isaiah will execute this script when he locates the archival master on mss3. However, we do need to think about a more automated pre-ingest environment – something to think about for R5.1.

Kalaivani reported that we have all the recent metadata changes and Yang indicated that WMS changes will add another week to the code complete date. We currently shooting for February 26, however this is tentative given some uncertainties with WMS edit and findObjects. We also agreed that ETDs will continue to be handled in R5.0 as they were in the previous release. In particular,

supplemental files will be tar-ed as part of the archival master. Finally, Yang will need to address the migration of the collection database in WMS.

ETD Specification

We had an abbreviated discussion on ETDs, covering only clarification on what we will do in R5.0. Given the mutual dependencies between WMS and RUetd, we decided that a separate specification will be needed for the R5.1 release.

As indicated above, we will treat supplemental files in R5.0 as part of the tar-ed archival master (the same as in the previous release). We also decided that moving ahead with xacml policies for embargo was a little too messy to include in R5.0 so this feature will be pushed into R5.1. For R5.0, the embargo date and dissertation will be held in the ETD application. For R5.1, Rhonda will work with Shaun and the MDWG to identify a metadata field and format to contain the embargo date and the RUetd will pass this date as part of the object to WMS. Yang also reported on an issue with content models, indicating that the current WMS release associated a policy with the ETD content model which made it difficult to move ahead with our proposal to have the ETD and its supplementary files take on one of the pre-determined content models (i.e. manuscript, video, etc). As a result, we will continue to use "ETD" as a content model in R5.0 and move to this new policy in R5.1. As a final item, Ron pointed out that we will need to include the pdf text handling issue in R5.1. Jeffery reported that there is also a related problem in that illegal characters are appearing in the XML-1 datastream. This problem should be fixed along with the pdf text layer (i.e. getting the appropriate page markers inserted into the XML-1 datastream so that full text searching will work across all of the ETDs).

Release Summary

Ron proposed a re-organization of the R5.1 and R5.2 release schedule. Basically, the changes are to remove relationship services and the annotation capability from R5.1 and move them to R5.2. This is being done because of the extensive architecture and development required for these capabilities. It was also proposed that we move ahead with using rels-ext for ETD supplementary files in R5.1. It is assumed that the relationships will be relatively static and therefore not difficult to manage manually.

Other Items

- We didn't have time to review the configuration and compile options. Jeffery will update the document and forward to sw_arch for comments. We'll discuss at the next meeting.

- Ron reported that it is important that we baseline the metadata document and associate it with a release. Basically, the metadata spreadsheet is the RUcore database and we need to make sure we know which version we are using.

Pending for Next Meeting

- WMS requirements for R5.1
- ETD/WMS requirements for R5.1
- Signature checking enhancements for R5.1
- Review of the updated Annotation spec
- Compile/configuration options
- R5.1 pdf text layer with page markup
- Review of R5.1 schedule
- Global registry discussion
- Use of API/A
- Relationship services
- Shibboleth configuration for mss3

rcj – 02/17/2009