

SW Arch Meeting Minutes – September 27, 2012

Agenda

- Announcements and Updates
- Status of Non-release projects
- Testing and Release of R6.2 and R6.3
- Jpeg2000 specification
- Checksums and applying sha256 to legacy objects
- Object definitions - continued discussion

Announcements and Updates

Ron noted that SC&I has received a grant from the Dodge foundation. It is not clear how they will be dealing with copyright issues. There also seems to be some interest in having students use WMS to prepare metadata for the videos resulting from the Dodge poetry festivals. Ron reported briefly on a NISO conference touching on a project that investigated issues in determining equivalence for datasets. Ron will send out summary notes to all.

Status of Non-release Projects

Dave has started running the jpeg thumbnail updates in batches of 500 and all appears to be executing normally. He did report on an anomaly in the presentation order of results in which the order might have been changed by adding the thumbnail. Jeffery will investigate further to see if there might be an issue with SOLR indexing.

The move of the JPE videos to production is complete. Kalaivani reported that the catalog records have to be re-pointed from lefty to the production server. (In a post-meeting email, Kalaivani indicated that the catalog records have been updated).

Given the above reported success, these non-release projects will not be reviewed further in sw_arch.

Status of R6.3 and R6.5

R6.3 (Analytic) has been installed on production. Chad will do a bit more sanity testing. We can now move on to address the testing of Fedora 3.5 on staging (our RUcore R6.5 release). We decided that we did not need the two weeks of “soak” time for the previous two dot releases so Dave can move ahead next week to install Fedora 3.5.

Checksums and SHA-256

We have previously reviewed the specification for applying sha-256 to archival masters in R7.0. Jeffery will post this specification. Jeffery also reviewed the mini-specification for updating all the legacy objects to sha-256, a process that would occur after release for R7.0. There are a possible three steps in the process: 1) update the existing archival masters (tar files) to sha-256, 2) remove the old sha-1

checksum from the techMD, and 3) unpack the tar files into separate archival masters to align with the policy adopted for R7.0 and apply sha-256 to each master. Although these three steps are fairly complex and have significant impact on the structure of each object, there appeared to be support for running steps two and three together and thus eliminating the need to run step one. For step 1, Jeffery has experimented with using the REST api to update the datastream attribute for the checksum. This process appears to work well and also requires the deletion of the previous master to avoid duplicate storage. For steps 2 and 3, we decided to have Jeffery develop an in-depth specification for review in our next meeting. He will also need to talk with Yang to address common issues such as naming the masters. If all the technicalities can be resolved (naming conventions, handling of derivative, etc), we will plan to run steps two and three as one update process.

Jpeg2000 Specification

Isaiah and Chad reviewed the jpeg2000 specification. The jpeg2000 format promises to offer us considerable flexibility in handling various jpeg resolutions and to also serve as an archival master at some point in the future. For example, the map portal and yearbook projects will be able to use jpeg images and a page turner in lieu of very large pdf images. There is still more research to be done to determine which open source libraries and server software to be used and to address performance issues. We will also need to answer questions regarding the structure of the object (e.g. a jpeg datastream for each page in a book or a “packaged” single datastream). We will plan to provide the jpeg2000 presentation capability in R7.5 (Spring, 2013) and the archival master capability will be scheduled for a later release. Chad and Isaiah will review the specification in the next CISC meeting.

Retention Policy

Kalaivani reviewed the spreadsheet for itemizes software that might store sensitive private information. Within the next week, we will supply Kalaivani with updates that will include WMS, dlr/EDIT, software libraries, Waand, and RUanalytic.

Other Items

Although the release summary changes frequently, Ron will start posting the summary for improved ease of access. We didn't have time to discuss the bugfix process or the continuing discussion on object types; these items will be scheduled for the next meeting.

Agenda Items for the Next Meeting

- Status of R6.5
- Specification for applying SHA-256 to legacy objects
- Process for reviewing bugfixes for R7.0
- Object definitions – continued discussion

- Pending agenda items
 - RUetd – WMS update
 - DOI specification
 - Enhanced UI for the landing page (a possible framework)
 - WMS – creation of relationships (rels-ext) for data projects

rcj – 10/08/2012