

SW Arch Meeting Minutes – December 5, 2013

Agenda

- Announcements
- Updates
 - NJEDL fix script
 - Extending DOI metadata
- Session timeout
- Discussion of content model drafts
- New reports and fixes for R7.3
- Release summary – code complete for R7.3

Announcements

Rhonda discussed the highlights from the panel discussion on Embargoes, Open Access, and ETDs, noting that graduate students can be misinformed about the purpose of the dissertation and suggesting that more emphasis should be placed on knowledge and scholarly communication. Ron reviewed the impact of Cabinet digital priorities indicating that the RUcore development for open access and faculty deposit is a high priority.

We decided to cancel the December 19 sw_arch meeting. Ron is involved in a candidate interview and there do not appear to be any pressing issues requiring meeting. We can continue to pursue the timeout issue (see below) via email and informal discussions. Our next meeting will be on Thursday, January 9, 2014.

Updates

Jeffery has tested the fix script for NJEDL on development with successful results. He will also update the reporting function and deliver to Dave to run next week on production.

Ron reported the feedback from Kalaivani on the CISC recommendation for proceeding with the extension to the DOI metadata. CISC recommends that we not spend time trying to determine if we should block indexing of certain collections (under the assumption that they might not be research quality). As a result, DOI metadata for all RUcore deposits will be entered into the DataCite Metadata store. In addition, we will extend the DOI metadata to include abstract and possibly additional values for type of resource. The MDWG will determine the appropriate mapping from MODS abstract to the DataCite “description” field and the necessary changes to type of resource. Ron will provide a mini-specification for implementation in R7.3. After we have implemented the extension, Jeffery will follow up with a script to do the DOI retroactive assignment. In the end, these relatively minor changes will provide additional access points (Thomsen Reuters Web of Knowledge, DataCite Metadata Store) and should result in additional downloads for RUcore content.

Session Timeout

Yang noted that we still don't have good information on the cause of the timeout problem. We have explored various fixes but still haven't been able to isolate the cause of the problem. Dave changed the php ini file system wide, but the problem remained. Developers should check unlink code and rm's to make sure that these functions aren't inadvertently deleting sessions. We should also check the system for cleanup of the tmp directory and the apache logs to determine if spiders might be creating problems. Dave will examine possible SUSE effects on session management. In the meantime, the best advice for WMS users is to frequently save their results.

Drafts for the Content Model Data Dictionary

Rhonda has the drafts for the Analytic (from Chad) and ETDs (from Kalaivani). Isaiah will provide the draft for Moving Images. Rhonda will integrate into a single draft and we will review it in our next meeting.

Fixes and Reports

Jeffery reviewed several possible fixes and report enhancements for RUCore, summarized as follows:

- Update the purge object function so it does not purge handles (Issue 2233). This is fixed and can be moved to test status for R7.3.
- Fedora content model checking (Issue 2184) has been disabled until we have updated all of our content models. At that time, we will reactivate Fedora content model checking. The disable fix can be moved to test status for R7.3.
- Jeffery will develop a script to assign "rep-devel" prefixes to all handles for object on the development server, thus eliminating any development handles pointing to production objects (Issue 1074). Schedule to be determined.
- We discussed the SOLR issue (1068) regarding searching of special double-byte characters (these might be math notations, Greek letters, etc.). We concluded that spending time on the search problem at this point is not useful, however there is a related bug outstanding (Issue 1858) suggesting that that certain OCR outputs, when migrated to the XML-1 datastream schema create XML parsing errors. Ron will work with Jeffery to examine several scenarios where this problem might be occurring.
- Enhanced reporting has been requested (Issues 1068 and 1069) to identify xml-1 and other datastream files that have a files size of zero. Jeffery will bring a specification for this enhanced reporting to the next meeting (Jan. 9) and the capability will be targeted for R7.4.

Jeffery will update software libraries to reflect the above changes.

Release Summary – Code Freeze for R7.3

We revisited the code freeze date for R7.3. Although much progress has been made (specifically on the A/A capability), we concluded that a more reasonable target date for code freeze would be January 17, 2014.

Other Items

Yang mentioned that he will need a VM partition to finish the development for openWMS. Dave indicated that this should be available next week. Also, Dave and Chad will determine the VM configuration needed for test and development of openETD.

Agenda Items for the Next Meeting (January 9, 2014)

- Review of Content model draft for Analytic, ETDs, and Moving Images (Rhonda)
- Specification for new reports to detect empty XML-1 and objects with file size=0 (Jeffery)
- Specification for DOI metadata extension (Ron with MDWG)
- Continued discussion of the session timeout issue
- SW_ARCH holiday celebration – bring in your leftovers

rcj – 01/06/2014