

SW Arch Meeting Minutes – May 9, 2013

Agenda

- Announcements and Updates
- User Accounts
- Google Search Results
- Status of R7.1 and Fedora 3.6.2 testing
- DOI Metadata Profile
- RUcore Release Summary
- Content Models – next steps (postponed to the next meeting)

Announcements and Updates

We briefly discussed the advantages of using ORCID for personal identity management in RUcore. Generally, it was felt that we should investigate this technology and that it may be useful in our ongoing implementation of user accounts. Guests – Joseph Deodato, Jane Otto and Laura Mullen – joined us for discussions on user accounts and Google searching.

User Accounts

Joseph reviewed the User Accounts functional specification, a capability that is targeted for R7.3. Since we had reviewed this document in CISC, the comments here focus on some of the implementation issues. We will need to investigate both Grouper for more fine-grained authorization and InCommon (Shibboleth) to enable members of outside institutions to authenticate using their netid. For the first release of User Accounts, we will use a local implementation to authenticate guests. The InCommon (shibboleth) and Grouper capabilities are targeted for R7.5. We will look to Chad and Jeffery to draft the implementation specification bringing in expertise from other members of sw_arch as appropriate. Rhonda will work with Jie to draft the specification for managing profiles and collections (scheduled for review in the next meeting – May 23).

Google Search Results

Jane discussed issues that she was seeing in Google search results, primarily related to the Ralston Horse Behavior Research collection. We verified that the videos are all properly indexed by restricting the search to mss3. It appears that Google is finding the resources although the videos may not appear until the 3rd page of results. The consensus was that external links to these resources are one of the critical indicators that causes Google to push the resources to the top of the list. Also, some of the metadata might be tweaked to improve searching. For example, the word “equine” does not appear as part of the title of the collection.

There were two observations from this discussion. First, we noted that the collection must be set for Google searching by an RUcore administrator. Secondly, collection managers may not be aware that

this is a necessary part of the workflow if they want their collection to be indexed. To address this situation, we decided to provide an RUCore enhancement to enable the collection manager to establish Google indexing. Jane will enter this feature into software libraries.

Status of R7.1 and Fedora 3.6.2 Testing

Dave has installed Fedora 3.6.2 on development and upgraded PHP. Initial testing has resulted in numerous fixes. Regarding PHP, we have decided to clean up all the messages in R7.1 that result from deprecated PHP functions. Chad indicated that there is one lingering problem with embargoes that prevents access even when the embargo has expired. It appears that this is related to the “unknown” user not being in the Fedora users’s file. If “unknown” is added to the user file, access appears to work. We will implement this fix by continuing to use the designation “unknown” in the user file.

We decided on the following actions. Early in the week of May 13, we will do testing of various RUCore functions on development. Assuming all is well, R7.1 upgrades will be moved to staging (Fedora 3.6.2 is already there) late in the week of May 13. Dave noted that the Iselon firmware upgrade is scheduled for May 22 at 10:30. In order to avoid having two downtime notices, we decided to do the public release of R7.1 in conjunction with the Iselon upgrade. The schedule is as follows: a) beginning 8 pm on May 21, upgrade production to R7.1, b) early on the morning of May 22 there will be an IP restricted window for sanity testing of R7.1 on production, c) at 10:30, do the Iselon upgrade, and d) we will plan to be back online by 1 pm on May 22.

DOI Metadata Profile

Rhonda reviewed the DOI metadata profile, indicating that the MDWG recommended using the DataCite profile so that we can have repeated elements (repeated elements are not allowed with the DC profile). To use this profile, Rhonda reviewed the MODS to DataCite mapping conventions. These conventions are quite well explained in the specification and will not be repeated here. One issue with dates is worth noting. If there is no date, ingest will proceed with the creation of a DOI that is marked as reserved. When a date is added later, the DOI status can be changed to active. It was noted that the DOI metadata mapping has been approved by CISC and the Metadata Working Group.

RUCore Release Summary

Given recent changes from CISC and the work to address Fedora 3.6.2 incompatibilities, Ron reviewed the resulting changes in our release targets. For R7.5, DOIs at the file level will be required to support the versioning of articles. Also, we will provide an article cover sheet, to be included in the PDF, in R7.5. For authentication/authorization in R7.5, we will investigate the uses of InCommon (shibboleth). There may be a relatively small annual fee to join InCommon. Given the changes required for Fedora 3.6.2, we have re-targeted the release of R7.1 to May 22. As a result, the code freeze date for R7.2 is now

set at May 20 (as opposed to May 1). It appears that most developers can make this code freeze date (Jeffery will verify that SOLR dynamic field indexing can be included by May 20). An update of the release summary will be distributed separately.

Agenda Items for the Next Meeting (May 23)

- Content models – next steps - WMS configuration, updating FD, stats, techMD (Rhonda)
- User account specifications for managing profiles and collections (Rhonda, Jie)
- Further discussion on when to schedule sha-256 updates for legacy objects
- Continue the discussion of how best to eliminate or fix RPC to initiate OCR
- Review of the code freeze date for R7.2

rcj – 05/20/2013