

## SW Arch Meeting Minutes - September 29, 2011

### Agenda

- Announcements and Updates
- Progress and Target Updates
  - R5.2.2 (WMS)
  - R6.0 - New Platform
  - R6.1 - SOLR/Lucene
- SOLR/Lucene Requirements
- Software development changes required for Fedora 3.4.2
- Continuation of AV objects with transcripts discussion

### Announcements and Updates

Our celebration acknowledged the success of Google Scholar indexing. One of the success indicators is that we have seen an average for January through September of approximately 26,000 downloads per month. In November of 2010, we were only seeing about 10,000 downloads per month. There was a brief review of CISC goals that impact sw\_arch including jpeg2000, EAD finding aids, LDAP identify management, and the context object. Goals that are already well underway include the new server/storage/OS platform and the Analytic. It was noted that we will likely proceed with djvu (the applet) for displaying yearbooks. We will still need to do some more testing of the applet and the unbundled djvu. Related to large file ingesting, Chad has had success in trials of ingesting 6GB and 14GB files. Based on this success, Ron suggested that we should ingest the Cranberry object on mss2 to not only test the large ingest but to determine if we can move this data from lefty64. Chad and Kalaivani will investigate, time permitting.

### Progress and Target Updates

*Release R5.2.2 (WMS Only).* This release is focused on WMS only and Yang reported that code complete for bug fixes will be by October 15. The extension of this date from the previous target related to file handling and enabling aggregate PDFs (especially for ETDs). We will take some more time to think through a more general policy for all archival masters. Given the revised target, we concluded that it was best to merge R5.2.2 with R6.1 since the dates for these releases were converging. As a result, R5.2.2 does not exist now as a separate release.

*Release R6.0 (New Platform).* Dave and Sho reported on the progress to achieve sanity with the new platform. The convergence of the new platform and several releases is creating considerable complexity in our release process. To simplify this process, we decided that releases R6.0 (new platform) and R6.1 (SOLR/Lucene) will be deployed on Fedora 3.2.1 (the current version). The upgrade to Fedora 3.4.2 will be conducted at some point next year. We concluded that there were no major features or bug fixes that are critically important at this time. In fact we will probably upgrade to a version of 3.5.x in the future.

We have achieved basic sanity on the new hardware with the current code base (R5.2) and Fedora 3.2.1. Sho will have another server set up by October 3 or 4 to test the PHP upgrade, Shibboleth, and API/A deny. The developers will have to check all PHP code on this base to determine if any capabilities have been deprecated or if other problems occur. We expect to achieve this basic PHP sanity by October 7. Based on this milestone, Dave and

Sho will work to a development environment set up on a new server by October 12. It was noted that we should update our configuration document after we have stabilized the development environment on the new server (Jeffery is the editor for this document).

*Release R6.1 (SOLR/Lucene)*. Rhonda reviewed the SOLR/Lucene requirements from the user services group. She noted that the objective is to provide the same basic capability as provided by the Amberfish search engine, however the group has designated areas for future development including various browse modes and faceted searching. The specification can be finalized with a few minor additions regarding terms to be added to the SOLR schema. Chad and Jeffery indicated that the requirements are sufficient for them to proceed with the implementation specification. The specification will cover the portal ID database and the impact on portals, searching, statistics and WMS. The WMS changes will extend the code complete date for WMS to November 1. We will continue to hold to the November 1 code complete for all of the related software for the SOLR/Lucene release. Given the comments earlier, we will merge the WMS fixes (previously release R5.2.2) with the SOLR/Lucene release (R6.1) with a target to production by December 15 and a release to the public shortly thereafter (i.e. before the holidays).

#### **Software Changes for Fedora 3.4.2**

Sho provided an email summary of the required changes. Given that we have decided not to upgrade at this time to 3.4.2, we did not review these changes in detail. However, developers should review the notes, especially for areas such as the upgrades to the SOAP API for manipulating rels-int. We may want to use rels-int for linking multiple sections of techMD.

#### **Other Comments**

Isaiah raised an issue about when we would proceed to MP4. Ron proposed that we should review all release content and targets next time in preparation for a review with CISC.

#### **Agenda Items for Next Meeting**

- SOLR/Lucene Implementation Specification – Chad, Jeffery
- Target dates for presentation to CISC
- Continuation - Video/audio objects with transcripts
- Large File Specification
- Review of the url migration implementation specification

rcj – 10/10/2011