

## Software Architecture Meeting, Thursday February 21<sup>st</sup> 2008

(Ananthan, Beard, Ellis, Geng, Hoover, Jantz, Liew, Marker, Nakagama, Triggs, Yu; Inma Megia attended as a guest of the committee)

---

### Agenda

---

- 1) R4.5 update
- 2) NJDH searching/collection object/empty collection - from testing
- 3) Cyber-infrastructure update
- 4) Storage architecture
- 5) Disk failure - what actions should we take
- 6) OJS journal export and metadata

### General Announcements

---

- We welcomed SCILS work study student Inma Megia, who has been working with Rhonda and Linda and will occasionally attend SW ARCH meetings as part of her learning experience.

### Agenda Items

---

- 1) R4.5 update

Kalaivani reported on the current status of the 4.5 release. Testing has been completed on mss2 and no problems have been experienced with the Faculty Deposit Module, RUcore, Fedora management, and the WMS. The handles on mss2 were successfully changed and tested, and the signature checker was run on the mss2 objects. The signature checker identified 15 failed test objects on mss2, most of which were attributed to the casual nature of such objects. Kalaivani is still investigating several of these, but they will not hold up the release. There was only one issue (discussed as agenda item 2) with the NJDH search interface, but a suitable workaround was found that should allow the release to go forward. The present plan is that announcements to partners about the release and attendant downtime will be made today (February 21<sup>st</sup>), the production system will be taken down for installation next Wednesday evening (February 27<sup>th</sup>), and the system will be brought back at or around noon on the following day (Thursday, February 28<sup>th</sup>).

- 2) NJDH searching/collection object/empty collection - from testing

During testing on mss2, it was discovered that the NJDH search interface would not function as expected without at least one object belonging natively to the root collection (NJDH). This was a known issue, and such an object had been created for the NJDH collection prior to the last release, but the new, tighter specifications for assigning collection membership in this release (the collection ID *must* appear in a mods:identifier element with a type attribute of "local" nested in a mods:relatedItem element) caused the test object in question, which did not have a mods:relatedItem element, to drop out of the collection. The solution was to repair the XML object by inserting the needed mods:relatedItem element and then to re-index the collections. Investigations showed that the corresponding object on mss3 did in fact have such a mod:relatedItem element, and that the problem should not occur on the production server. Even so, the problem highlights some issues that continue to arise from the historical ambiguities that attended the creation and definition of some of our collections. After the release, Kalaivani and Rhonda will examine the collection issue more carefully and explicitly map the NJDH collection structure.

### 3) Cyber-infrastructure update

Ron gave an update of the recent Cyber-infrastructure meeting and how it might affect the work of this group, specifically future releases. We anticipate moving from Luna Insight to ArtStor. It was decided that we should plan on a “requirements release” to coincide with the 4.6 release, with a Requirements (for R5.1) deadline of May 30. A number of needed requirements documents were identified, which will be specified at the end of these minutes. Isaiah and Kalaivani will be pulling together this year’s ARL stats from various sources, but it will be desirable to automate this process in the future. The issue of submitting grant-funded publications to PubMed Central was discussed. It was estimated that for the next few months at least, there will be only a “trickle” of such documents. Chad will create a simple web form allowing files to be uploaded to Rhonda, who will handle submissions to NIH individually for the time being. This is expected to be a learning experience that will feed back into an automated submission feature from the repository that would be delivered in the 5.1 release. At present we anticipate the 4.6 release in June/July, the 5.0 release (including major structural changes but no new features) in late September, and the 5.1 release in December of this year, though we will be revisiting these dates as work progresses. The open source WMS, scheduled for release in 4.6, will be included in the 5.0 release.

### 4) Storage architecture

Ron presented a storage architecture document with a proposal for handling very large files, such as archival video masters. The proposal is to ingest objects with “external” archival datastreams that will be maintained on a backend server. It is hoped that this approach can be made compatible with Fedora’s future development of “managed external” datastreams. We identified the need for a backend server that could be used both for development and production, rather like the current OCR and handle servers. We formed a Storage Architecture Subgroup to look into these issues further. Ron proposed that initial membership should be Dave, Sho, Isaiah, and Jeffery, as well as any others from SW\_ARCH who might be eager to join such a group. The group will address backup, redundancy (which must also work for NJVid), dealing with large files, and a review of recent hardware failures. It was noted that any new equipment must be ordered by May because of budgetary considerations.

### 5) Disk failure - what actions should we take

Dave followed up on his dealings with Sun. We hope to be able to convert our service, time, and materials expenses into a service contract with Sun. Sun feels that the machine would not have “panicked” as it did if more recent firmware had been in place. We need a firmware upgrade, which will require more downtime. Any such downtime for the firmware upgrade will be scheduled after the 4.5 release. It was concluded that we need to return to this issue.

### 6) OJS journal export and metadata

Shaun is working on generating XML exports from OJS that will go into WMS using the model of the ETD interface. The “handshake” between OJS and WMS must be established. Ingestion into Fedora will be done at the article level. As with the earlier prototype, the handles for OJS objects will point to the OJS abstract pages rather than the normal showed page for Fedora objects. The ETD model will be followed for generating searchable full texts. That will be done, as with the ETD

texts, within the WMS. Rhonda and Shaun will work on requirements, and we expect the release to be in 5.1.

#### Requirements documents due May 30, 2008

---

The following requirements documents were identified:

- 1) Statistics for various projects (Isaiah, Kalaivani, Chad, Jie, Rhonda)
- 2) Storage architecture (Dave, Sho, Isaiah, Jeffery)
- 3) OJS export to WMS and Fedora (Shaun, Rhonda)
- 4) Submission to PubMed Central (Rhonda, Mary Beth)
- 5) RUCore/NJDH enhancements (Rhonda, Linda, Mary Beth)
- 6) Marc export enhancements (Kalaivani, Mary Beth)

#### Next Meeting

---

The next Software Architecture Meeting is scheduled for Thursday March 6<sup>th</sup> at 9:30 A.M. at the SCC. An initial, proposed agenda is below:

- 1) Pending items from today's meeting:
  - PubMed submissions
  - statistics requirements for next release (Chad and Jie)
  - rels-ext and xacml experiments (Jie and Jeffery)
  - digital docs requirements (Mary Beth)