

Introductory Items

- We had a discussion in the steering committee regarding how we might deliver WMS to partners. The current approach would be to provide a pre-open source WMS that would probably be the current R4.2 version. Partners would need to understand that this version is for testing and for them to familiarize themselves with the software – both installation procedures and the code. We will not make changes to the software to remove Rutgers dependencies. They are “on their own” if they make changes, i.e. we will not support WMS if they make local changes. In the Fall, 2007 we would release the open source version which would have a more modular and configurable architecture, e.g. allowing them to use other PID technologies such as ARKs. Grace will work with Jeanne Boyle on the license agreement for the pre-open source release.
- The ad hoc project coordination group (Ann, Chad, Ron) addressed the issue of how we should manage non-RUcore projects. We identified four major areas: WAAND, CF5 to CF7 migration, pending projects (e.g. Latin American Pamphlets) and partner collaboration. We are proposing a project coordinator for each of these areas. Discussions with Grace will occur after she gets back from vacation.
- In the steering committee, we acknowledged that the current long list of collections on the advanced search page of both RUcore and NJDH is not sustainable (the list will continue to grow). We proposed separating out the institutional repository collections (ETDs, faculty submissions) onto a separate portal as a near term solution. In the longer term, we will need to develop a more comprehensive method of dealing with lots of collections on the user interface.

R4.2 Status

We agreed to hold to the schedules as proposed by Kalaivani in a previous email (see below). In order to make these dates, we will move the DNG metadata enhancements to R5.0. Also, the pipeline virus scan for faculty submissions will be move to R5.0. Virus scanning will be done by manually invoking the command for faculty submissions in the R4.2. Given the late arrival of some of our requirements, we decided to add a target date for “requirements complete” for the next release. Isaiah will resurrect the targets document and provide for updates.

The release and testing dates for R4.2 are the following: a) code freeze – May 25, 2007, b) lefty testing from May 30 through June 20 with following intervals: i. Testing 5/30 – 6/08, ii. Bug fixing 6/11 – 6/15, iii. Re-test on lefty 6/18 – 6/20, c) delivery to mss2 – June 22, d) mss2 testing – July 5 through July 10, e) Notification to partners – July 10, f) mss3 installation – July 11 through July 16, g) reminder to partners – July 16, and h) pubic release – July 17, 2007.

Collection Identifiers

We continued discussion of how to handle collection identifiers since Yang pointed out a difficulty with detecting whether a collection had already been generated. We need to be able to do this to automatically generate collections for faculty submissions. After a prolonged, but very useful, discussion we made the following decisions:

- All collection identifiers will be generated automatically.
- For faculty departments, we will use the LDAP organization code. There was considerable discussion about what was actually represented in the organization code. For example, it was thought that a librarian at Kilmer was shown with a “Kilmer” orgcode rather than an RU Libraries code. (Post- meeting investigation indicated that this was not the case and that the orgcode should work fine).

- We modified the collection identifier format from the previous meeting to appear as follows: a fixed string or “rucore” concatenated to three subfields where subfield 1 is a single digit representing the type of collection, subfield 2 is a five digit field that is either all zeros or contains the LDAP organization code and subfield 3 is the 5 digit noid sequence number. For subfield 1, permissible values at this point are 0 – regular collection, 1 – faculty collection, 2 – ETD collection and 9 – virtual collection. In a subsequent sw_arch list posting, Yang posted the collection identifiers for the virtual ETD, RU and RU faculty collections – see below.

ETD collection: rucore90000000010

Rutgers University collection: rucore90000000020

Rutgers University Faculty collection: rucore90000000021

An issue was also raised which Rhonda will pursue with appropriate committees and working groups. The issue was whether or not people other than faculty can submit (e.g. students or staff members).

Collection Hierarchy

We decided that we did not have time to make any significant changes in how collections are represented on the advanced search pages. Chad will provide pointers from the ETD and faculty submissions pages that will position a search in these respective collections. The more comprehensive collection hierarchy display will be implemented in R5.0

Checksum Service

Jeffery and Dave will work to get the checksum verification service working in R4.2. Although we can probably verify all checksums for 10,000 objects in one overnight run, the service needs to be configurable so we can spread it out over several days.

Quick R5.0 Discussion

We didn't have a lot of time to discuss candidate R5.0 features. We did propose to set up Fedora 2.2 only as a test and investigation platform. Other than using the checksum API to create and verify checksums, it doesn't look like there are any other important features in Fedora 2.2. The proposal is to develop R5.0 on Fedora 2.1 and then jump from 2.1 to 2.3 in the Spring, 2008. Fedora 2.3 will offer a number of new capabilities including event management/messaging and CMDA (content model disseminator architecture). Ron is trying to schedule a call with Sandy Payette to see if we're missing anything by jumping from 2.1 to 2.3.

There are a lot of candidate features for R5.0 and there will need to be considerable more discussion. Certainly the PDF server and WMS open-source will be strong candidates for R5.0. Our target release timeframe will likely be mid to late Fall, 2007. We also decided that it would be prudent to experiment with compound objects in Fall and implement the capability in a Spring, 2007 release. Compound objects will be needed for ETD supplemental files, however we expect to have very few of these initially. In the interim, all supplemental files will be added as datastreams to the simple object.

Other Items

Kalaivani reported that she had repaired some 70 plus objects in mss3 that were not passing the checksum verification. We expect that now that all mss3 objects will pass. The 1782.2 handle server now appears to be working properly. Because of Ron's vacation, the next sw_arch meeting is scheduled for Thursday, June 21. Members of sw_arch are encouraged to work through Chad and Ann Montanaro if a meeting is needed before June 21.