SW Arch Meeting Minutes – January 17, 2013

Agenda

- Announcements and Updates
- Status of R7.0
- Specification – sha256 for legacy objects - postponed
- DOI specification
- Review of R7.1, 7.2, and 7.5

Announcements and Updates

Yang reviewed the options for wrapping techMD – use METS, FOXML or do not use a wrapper. We decided that we would not use a wrapper. In a related item, Yang will prepare the schema for technical, source, and rights metadata. We will need to work with the MDWG to insure that the schema is kept up to date. Ron noted that the jpeg thumbnails have been updated for the outstanding 288 objects (per a recent email from Dave).

Status of R7.0

Kalaivani indicated that she expected testing to be complete on development by January 22. We would then move to staging with an expected release to the public on February 11. Isaiah will post notifications of a release update on February 4. We have scheduled the sw_arch meeting of Feb. 28 to celebrate with champagne, orange juice, and bagels.

DOI Specification

Ron reviewed the functional specification for using DOIs in RUcore. Our first implementation in R7.1 would involve replacing Handles with DOIs going forward, i.e. we will still need to schedule the work to update all legacy objects. There is one difference in how we implement DOIs versus Handles. The DataCite organization has requested that all DOIs be assigned a set of core metadata including Creator, Publication Year, Title, Version, Publisher, and ResourceType. In most cases, this metadata can be extracted from the object metadata automatically without manual intervention. Some further investigation will be needed to handle situations when there is no version number or when a resource does not have an author/creator (e.g. for a photograph). The assignment of metadata to a DOI also creates a situation in which we have two representations of the same metadata. As a result, any edits of these items would have to be reflected in two places. This will likely be a rare event and can be reasonably managed. Ron noted that we have an authorized account with EZID to begin our development and have been allocated one million DOIs per year. Complete API documentation is available at http://n2t.net/ezid/doc/apidoc.html. See especially the PHP examples at http://n2t.net/ezid/doc/apidoc.html#php-examples. To conduct initial testing, developers should use the following: user name: apitest ; password: apitest.

The DOI specification also proposed that we examine how to implement persistent IDs at the file (datastream) level and how to manage versions of objects and datastreams. Both of these capabilities require extensive discussion and are not scheduled for implementation in any RUcore release. However, we were able to begin the discussion of the issues which are briefly summarized here. Ron will update the specification and schedule additional reviews in sw_arch, CISC, and the Research data group.

Archival Resource Keys (ARKs). The basic proposal is to implement persistent IDs at the file (datastream) level using ARKs. This capability allows more fine-grained citation, e.g. a researcher can cite a file directly. In addition the ARK has a capability called "inflections" that is not available with DOIs and which will allow us to provide several presentation options. For example, for an individual presentation file, we could direct the user to either the resource or to an associated landing page using a single ARK with the appropriate inflection character. The landing page could display additional metadata that might not be presented at the object level. The EZID software release in January, 2013 will provide a partial inflection capability so, beyond the need for further discussion, we don't yet have the full capability available to us.

Versioning. A section on versioning of objects and datastreams was included in the DOI specification since we will also have to understand how to manage DOIs for versions. It is acknowledged that versioning has impact on many aspects of RUcore including metadata, faculty deposit, and the data portal. These notes provide a brief summary of our discussion. In RUcore, versioning can occur at three levels – the datastream, the object, and the project. A versioning action can arise out of three conditions: 1) an update has been made and we want to replace the previous resource, 2) an update has been made and the researcher wants to add the update while leaving the previous resource active, and 3) an administrative update has been made. In an example of #1 and a datastream replacement, the following steps would need to be taken:

- Add the new datastream using Fedora versioning and assign a new DOI
- Update version numbers in the metadata
- In the event that the previous resource had been cited (via the ARK capability noted above), the ARK for the previous resource would have to be redirected to an information page that indicates the resource has been updated, possibly a reason for the update, and directing the user to the most recent resource.

Although this scenario seems relatively straightforward, the procedures become more complex when we address the versioning of objects (as opposed to datastreams). As noted above, more discussion in different forums will have to be conducted before we can finalize the approach.

Review of Proposed Content for R7.1, R7.2, and R7.5

We reviewed the proposals for release content in the next three dot releases. R7.1 will focus on the DOI implementation along with related enhancements (e.g. moving to an RUcore url rather than using "mss3"). R7.2 will focus on the upgrade to Fedora 3.6.x and the associated testing. We also proposed that we do a PHP version upgrade as part of this release. R7.5 focuses on enhanced faculty services, video (Wowza) and security (authorization/authentication). There will be more discussion on this release content; the release summary is attached below and will be posted to the RUcore site.

Agenda Items for the Next Meeting

- R7.0 Status
- Specification for sha-256 and legacy objects
- Specification for configuring SOLR to add metadata fields without a code release
- Review of mss3 to RUcore migration specification
- RUetd – WMS update
- More on DOIs

rcj – 02/01/2013

RUcore Release Summary – Specifications and Target Dates (R6.x and R7.x)

1. Release 6.2 – Faculty Deposit Enhancements – Released September 17, 2012
2. Release 6.3 (Analytic) – Released September 28, 2012
3. Releases 6.3.1 (October) and 6.3.2 (released Nov. 7)

4. Release 7.0 – Large Files, Complex Objects, Faculty Deposit (Release Target – February, 2013)
Specifications

- File Configurability - Complete
- File hierarch/structural map - Complete
- Fedora checksums – Complete
- Background ingest with alerting - Complete
- Download statistics for the Faculty Portal - Complete
- Faculty Deposit User Interface (more enhancements) - Complete
- Complex object User Interface – download zip for object or select files from a directory display(see structural map specification) – Complete
- API for OAI-PMH

Development (Code complete – November 21, 2012)

5. Release 7.1 – URL/DOI release (Release Target – TBD)
Specifications

- mss3 to RUcore migration – complete
- DOI implementation specification
- Dynamic field indexing with SOLR
- UI for landing page

6. Release 7.2 – Upgrade to 3.6.x (Release Target – TBD)

- Install Fedora 3.6.x on staging
- Testing of R7.0 on staging
- PHP upgrade
- Post-release update of "legacy" objects to sha-256 checksums

7. Release 7.5 – Faculty Services, Video, Security (Release Target – TBD)
Specifications

- Authentication/Authorization (Scenarios available)
- WMS File handling, validation and metadata extraction (with exif, media tools)
- WMS – support for external relationships
- Versionioning (file, object, and project)
- Faceted browsing
- Faculty service enhancements
    - Upload an html file in faculty deposit
    - Dynamic bibliography
    - Statistics enhancements
- MP4 container and support for Wowza
- Jpeg 2000 and page turner
- Schema for technical, source, and rights metadata

8. Release 8.0 – EAD Support (Release Target – TBD)
rcj – 01/18/2013