SW Arch Meeting Minutes – June 28, 2012

Agenda

- Announcements and Updates
- Progress on Non-release projects and projects requiring scripts
- Specification for statistics
- Fedora 3.5 on Staging and testing of R6.1.2
- Review of Release Targets

Announcements and Updates

Rhonda requested a report regarding how authority ("auth" attribute) has been used. Jeffery will produce a report. This issue led to the discussion of a more general reporting capability. Rhonda and Ron will talk off-line to see if we can get a specification for this capability.

In an informal meeting earlier this week, Grace requested that we raise the priority for implementing facets. She is particularly concerned about having a facets capability in VMC this year. We agreed to start investigating and prototyping facets for VMC, focusing on how SOLR/Lucene should be configured, metadata inconsistencies, and how to present facets in the user interface. Given our already large release R7.0, we will likely want to do a dot release for facets only.

Ron has prepared notes on implementing DOIs with input from Chad and Jeffery. Architecturally, this should be a fairly straight-forward implementation, however there are some policy related issues that must be resolved (e.g. do we assign DOIs to all of our resources or only to those resources that are considered scholarly). We'll review these notes in the next sw_arch meeting and then take it to CISC. The CDL organization has indicated that our annual bill will be $830 (rather than $2500). The reduction is due to the fact that we are implementing the capability for a library service as opposed to a university wide service.

Individual Object Statistics for the Faculty Portal

Jie and Chad reviewed the specification for statistics for the Faculty Portal. This capability will only be available initially for faculty deposits. The basic decisions regarding this capability are listed below:

- We will count visits to the full record page via rucore searches and visits to showfed via handle resolution from Google searching.
- We will list all datastreams (file types) separately rather than aggregating into one statistic. For example, we will show pdf-1 and pdf-2 separately.
- No data for embargoed datastreams will be presented.
- We agreed not to show data for cities.
- We will show data by country and will not break out results by .com, .edu, etc.

- For presentation, we will show both a pie chart and a table and limit to the top ten groups. The remaining data will be shown as "Other".
- We will not show data by institution (i.e. Cornell, Princeton, etc.)
- The link to the statistics display will be shown on the full record and showfed.
- For use of the commercial database, we recommended that we take the option of paying for 2 million queries for a price of $200.

Jie will revise the specification and we will review it quickly in the next meeting and then take it to CISC.

Fedora 3.5 on Staging and Testing of R6.1.2

Dave indicated that Fedora 3.5 has been installed on staging and we can proceed with testing. Everyone should begin testing and reporting results on the sw_arch list. Dave indicated that there appear to be some anomalies with policies and with the relationship API, noting that the pid form of the subject method in the relationships method is now deprecated. Rhonda will add a policy and test; Jeffery will run tests on the API.

Projects Requiring Scripts and Non-release Projects

Jeffery walked us through the mini-specification for four projects requiring the running of a script – Voorhees rights statements, misspelled qualifier, xml-1 datastreams, and jpeg thumbnails. The script for all these projects is ready to run and Dave will run three (all except thumbnails) in the next few days. Jeffery and Yang will verify that they are using the same procedure for creating thumbnails. If so, Dave can run this script also.

From a process perspective, we agreed to create a mini-specification for each project that requires a script. In our next sw_arch meeting we will review and agree on the script to correct NJDH rights statements. Since there are many possible projects requiring a script, this topic will become a standing agenda item.

Review of Release Targets

Ron reviewed the document for our release targets. For R6.6 (faculty deposit enhancements), the code complete date is July 16. We will do development and staging testing and plan to release to the public on July 30. We decided that a notice of downtime is not required for this release.

We also decided to move the code complete date for R7.0 to October 26 to accommodate the complex changes required for WMS. We'll target the release date for December although this may be difficult because of the year-end holidays. The facet capability has been added into R7.5 however, as indicated earlier, we will likely put out a dot release for facets only.

Other Items

Dave brought up several items. He is working to re-purpose storage from the old Solaris system. Given some recent confusion about adding the first object to a collection, it might be a good idea if we could clarify this process. Regarding JPE, we should be able to ingest the remaining 11 objects into the production server but need to verify that we have the appropriate links in the IRIS record. Finally, we need to discuss the decommissioning of lefty64 (this is an agenda item for next meeting).

Agenda Items for Next Meeting

- Status of Non-release projects and a new project – NJDH rights statements
- Quick review of Faculty Portal statistics specification
- Complex object implementation specification
- DOI Implementation
- Decommissioning lefty64
- Pending agenda items
    - RUetd – WMS update
    - Enhanced UI for the landing page (a possible framework) – Chad and Jeffery
    - WMS – creation of relationships (rels-ext) for data projects

rcj – 07/09/2012