Software Architecture Meeting, Thursday, July 8th, 2010

Ananthan, Beard, Geng, Marker, Mills (recorder), Triggs, Yu

Agenda

1. **Announcements and updates** (all)
2. **JOH ingest** - Ingest will have to wait until after R5.2, i.e. late this year (Beard, Marker)
3. **Annotation** - EPPN, MP4, indexing (Mills)
4. **Changes in collection and resource metadata** (Ananthan, Marker)
5. **Collection structure** (Ananthan, Marker)

Agenda Items

1) Announcements (all)

Report from the July 7th CyberInfrastructure was reported. Those items included the MP4 Video Format specification was reviewed and approved. At the meeting changes to the search interface presented to the user were discussed and approved. Three major portals; Scholarship, Research and Cultural & Historical Resource, with every object in the repository belonging to at least one was approved. A specification outlining the details of the implementation will be written. Finally from the meeting, the coming year's goals will focus on user experience and outreach.

2) Jazz Oral History (JOH) ingest - Ingest will have to wait until after R5.2, i.e. late this year (Beard, Marker)

79 of the 82 oral histories archival datastreams will exceed the 2GB file size limit. Support for this will need to be included in R5.2 along with possibly restricting access to the presentation versions to only the Dana Library. The audio files are currently split and need to be concatenated as well.

3) Annotation - EPPN, MP4, indexing (Mills)

After speaking with NJViD implementation of multiple components of the annotation tool need to be revised. NJViD will not be using XACML to restrict annotations to certain EPPN. They are implementing a database and adding that granular restriction checking to the retrieval process. Also, NJViD has moved to MP4 video file support exclusively. They still need to re-encode videos that have FLV and MOV formats to having just a MP4 datastream. NJViD also wants to stream all video, not just restricted video, to their end users. Given these two major facts and many minor factors the Annotation Tool development completion is still far out. It was recommended that R5.1.2 just support annotation objects, and not the tool itself. This will allow pre-produced annotations, supplied to NJViD, to be created in a batch process and viewed. Individual annotation creation will come at a later time after all the nuances are work through.

4) Changes in collection and resource metadata (Ananthan, Marker)

Changes to collection and resource metadata for R5.2 were discussed, noting additional collection metadata will be added to the MODS section of the object. No objections were raised; a final specification will be posted.

5) Collection Structure (Ananthan, Marker)

Feedback from Agnew was reported. In the future a RUcore user will be presented with the option of search across three major portals; Scholarship, Research and Cultural & Historical Resource. All resources in RUcore will need to be identified as being part of at least one of those portals; resources can be part of more than one major portal also. A specification is being drafted by Ananthan and Marker that outlines changes that need to be made for R5.2.

The current collection hierarchy is not consistent between the multiple applications that depend on it. Namely the collection structure found in the WMS does not match the one used for the search portals. This is because both applications are not pulling that information from a single source. Marker noted that the hierarchy currently found in the search portals is the more ideal of the two. WMS has experienced issues with bugs in the collection management module. It is thought for R5.2 the bugs found in WMS will be resolved and the collection hierarchy will be aligned to that of the search portal. This does not address the larger problem of there still being two data sources which will be out of sync. In a future release, possibly R5.3, the two data source issue will need to be resolved by creating a single source of that data both applications can draw upon. A specification will be written for R5.2 that outlines changes to software that need to be made to support the immediate need of synchronization.

Next Meeting

The next Software Architecture Meeting is scheduled for Thursday, July 22nd at 9:30 A.M. at the SCC.

Topics for next meeting

Outstanding R5.2 Specifications

The following are R5.2 specifications that need to be finalized and/or approved
  * Collection Structure Specification, currently being drafted

Update on Google Scholar

SMAP revision to make METS compliant and SMAP datastream review