

Software Architecture Meeting, Thursday April 3rd, 2008

(Ananthan, Beard, Geng, Hoover, Liew, Marker, Mills, Nakagama, Triggs, Weber, Yu)

Agenda

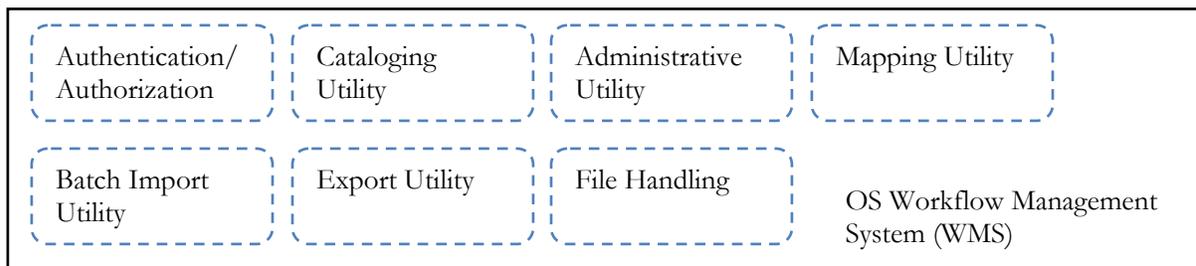
1. Core set of controlled vocabularies to deliver with the open source WMS along with name space discussion - Marker; Weber has been invited as well.
2. 4.5.1 Update - Ananthan
3. JPEG2000, Aware server, report, see email from earlier in week - Beard
4. Summary of Storage Architecture discussion covering:
 - storing video archival DS externally in R5.0. File will be stored on local file system and will not require equipment purchase.
 - making all archival DS external in R5.1. Might require equipment purchase.
 - initial downtime policy discussion
5. Scheduling of requirements for R5.0/5.1 - All
6. Open Source File Headers - Mills
7. Quick items (Google indexing of faculty collections , demo objects on mss3)

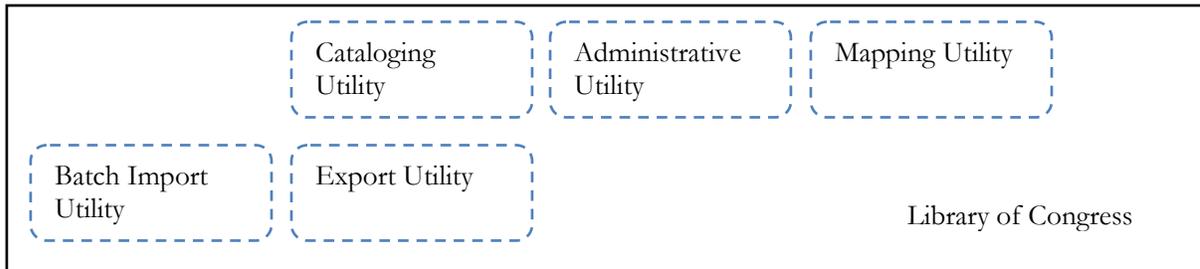
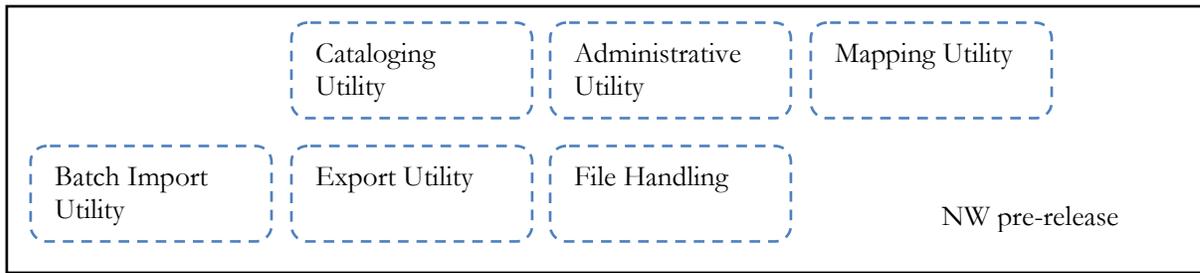
Postponed until next meeting

- Exporting ETDs to ProQuest - Ellis

General Items

- Next meeting is Thursday April 17th.
- Current status of open source WMS development was discussed. Yu made a point to state his current development plans put a possible open source delivery of WMS at June. Development of the Administration/Authorization module has not started. Yu needs at least one month of development time to finish this module. Yu noted that since the whole delivery timeline is dependent solely on him it is putting, as was agreed by the group, undue pressure on him as a developer. Yu requested help with future development. Yu feels the current development of the WMS A/A module needs to be done by him because bringing in an additional developer at this time would slow down the development. This is due to the fact this additional programmer would need time to get them up to speed with the WMS design pattern and philosophy before developing the module. Mills stated his frustration in the past with getting more programmers on board with WMS development as he feels he has gotten mixed messages from the different parties involved in accomplishing this. Mills also mentioned that even though it might be too late to currently have a programmer work on the A/A module future modules will need to be developed and this inclusion of an additional programmer needs to be done as soon as possible to alleviate future pressures on people. Mills will look to schedule a meeting with Agnew, Jantz, Yu, Ananthan and himself to discuss this.
- Terminology when speaking about the open source WMS has created confusion. The WMS and its variants are made of a set of modules, sections of code that perform specific tasks. See diagrams below:





Some confusion is coming from the use of releasing what is being called a “cataloging utility.” This is in direct conflict with the name of a module also called “cataloging utility” and is causing the confusion. When stating we will release the cataloging utility what is really being discussed? Just releasing the module “cataloging utility” will be pointless and not be worth anything. The “cataloging utility” application needs the inclusion of the modules noted in the Library of Congress diagram, the Cyberinfrastructure Steering Committee needs to make note of this development.

1. Core set of controlled vocabularies to deliver with the open source WMS along with name space discussion - Marker; Weber has been invited as well.

- No case by case evaluation of what will be delivered as a core set of vocabularies will be done. MDWG will take the existing set of MIC vocabularies; excluding and editing this set to create the core set that will be delivered with the open source WMS. After MDWG is done Yu will dump the controlled vocabularies, creating the SQL statements for optional use by open source implementers. MDWG will sort through in April. In mid-April Northwestern will get a snapshot of whatever the current status is of this vetting.

- LC has a current delivery date of May '08.

- A namespace needs to be delivered with the Open Source controlled vocabulary for validity. We have no schema so one will need to be created. There is not enough time currently to prepare a schema before the first open source release of WMS. We should however start working towards generating a schema. We currently do not validate our MODS before ingest. We will need to decide if we want to start to validate against a schema and at what point of the workflow will we want to do this. Triggs noted Fedora will validate against a schema at ingest. Marker mentioned future conformance to PREMIS and MIX could be achieved as well. Yu noted the source, rights and tech MD did not come from METS, descriptive did however. The group needs guidance from Cyberinfrastructure regarding this and Mills will speak with Jantz regarding this.

2. 4.5.1 Update - Anathan

Anathan reported that testing of 4.5.1 will occur on April 10th. Developers will re-commit 4.5.1 code to cvsroot to refresh dates on delivered code.

3. JPEG2000, Aware server, report, see email from earlier in week - Beard

-Beard discussed a draft document of the JPEG2000 implementation presented to the group. Beard will look into purchasing extended software support instead of purchasing server hardware that might or might not be used in the implementation. It was too early to make informed decisions about how to spend money on hardware. The software license pursued will run on many operating systems and includes developer and production licenses. JPEG2000 generation will need to be looked into as well. There are some open source solutions that could be investigate. Beard estimates a soft rollout of JPEG2000 services in 2009.

-This brought to light an issue this group has been presented with two years in a row. We have been put in a position where end of year money could be used to purchase software and/or hardware and we have been ill-prepared as a group. In fall or early winter this group needs to start to identify what it would recommend for end of year spending so as not to be caught off-guard and an informed decision can be made.

4. Summary of Storage Architecture discussion

- A summary of what was discussed in the Storage Architecture meeting that occurred one week before was discussed. This included offloading signature checking to the “external device,” creating a read-only system when unexpected downtime has occurred, creating external video arch datastreams in R5.0. The file will be stored on local file system and will not require equipment purchase. No detailed discussion happened at that meeting as to how the WMS would manage the external files through its internal process. Nakagama mentioned Fedora 3.0 running checksums might be questionable. Marker mentioned that bringing in a consultant might be useful at some point. Hoover noted that unless we clearly understood what we wanted to do and could communicate that effectively a consultant might not be that much of a help. If future Storage Architecture meetings are to happen externally from Software Architecture very detailed agenda’s need to be distributed so that parties involved in the discussion will know to go to that meeting.

5. Scheduling of requirements for R5.0/5.1 - All

- Marker is working on the faculty survey specs and is setting up meetings with her group.
- Marker stated the NIH requirements are well in hand also.
- Ellis and OJS are compiling the OJS specs.
- Marker is working with Ananthan to develop the MARC specs.
- Marker is working on NJDH, RUcore and Partner Portal specs with User Services as well.
- Triggs might have a content model and disseminators spec ready for the April 17th meeting.
- Geng is working on the XACML specification.
- Geng and Mills will be working on the stats spec in April and hope to have something ready by early May. Mills will setup a meeting with Marker, Ananthan, Vince(NJEDL) and anyone else who has requests.
- Weber needs to be invited to the next Software Architecture meeting to get feedback for preparing the Administrative Space specification.

Topics for next meeting

1. Administrative Space spec/Digital Docs – Weber (Needs to be invited)
2. Open Source File Headers - Mills
3. Signature Failures and how are we handling them? - Ananthan
4. Quick items (Google indexing of faculty collections , demo objects on mss3)
5. Exporting ETDs to ProQuest - Ellis