Published on *LTER Information Management* (http://im.lternet.edu)

Meetings notes: third meeting, 11/13/2012

WG VTC Tuesday, November 13, 2012
Participants: Philip, Suzanne, (Duane separately).

The group discussed the layout of the reports and reviewed the technical requirements to produce the information in the suggested format. The WG has volunteered to report back to the full IMC in December.

Duane has spoken to Mark Sevilla and they have some ideas (see below):

"(1) Should data package reports be "PASTA-centric" or "Metacat-centric"?

(A) In a PASTA-centric approach, the reports focus on data packages that have successfully been inserted into PASTA, while leaving out details about data packages that are currently in Metacat but not yet in PASTA. We could include a simple metric based on the ratio of the number of data packages in PASTA to the number of data packages in Metacat for each site, but other than that, we would not be reporting on the details of data packages in Metacat or why they are not yet "PASTA-ready".

(B) In a Metacat-centric approach, we would generate quality report statistics for all data packages in the LTER Metacat regardless of whether or not they have been successfully inserted into PASTA. This is a more comprehensive approach, but also much more time- and resource-intensive. The original report generated by Margaret used this approach.

IM Exec may want to continue using the more comprehensive approach described in (B) rather than the more "light-weight" approach described in (A). However, I think it's worth mentioning both approaches because if (A) was acceptable it would be much easier to implement.

(2) Add native support in PASTA for a quality report database.

The original approach for generating Data Package Reports would involve parsing thousands of XML quality reports and aggregating/collating them into a final report. Mark and I have started to discuss adding native support in PASTA for database tables that store quality report statistics. We think this approach would make it much easier for the Data Package Reporting WG to generate reports. There will be more room to experiment with a variety of different database queries and generate a variety of report outputs. It also makes it easier to generate reports for a specified time period. Also, the WG would probably be able to generate its own reports directly from the database without relying on LNO to manage the process."

Current Issue(s): Problems with the way in which data access is recorded would result in sites being sent large numbers of e-mails. This would need to be disabled in some way prior to the reports being run.

Actions:

1. Report back to full IMC at December VTC – ALL.

Next Meeting (VTC): December 2012.

Source URL: <u>http://im.lternet.edu/node/1128</u>