Zephir Advisory Group

October 21, 2015

Present: Ryan Rotter
               John Mark Ockerbloom
               Kathryn Stine
               Todd Grappone (recorder)
               Tim Cole
               Chew Chiat Naun
               Jon Rothman

Absent: Gary Charbonneau

Ryan proposed that try a different format for our meetings (perhaps one that allows for desktop sharing), so the next meeting will be held on Google Hangouts, which Kathryn will setup.

1. Zephir Update

Updates: ZAG would like to have access to the Zephir a quarterly record loading reports. For now, the Q3 summary numbers are:

> 494,310 records loaded (total)
> 340,439 are updates, 153,871 are new
> 5 shadows loaded (records updated by Zephir staff to force closure of
>   in-copyright works mistakenly opened as public domain)

ACTION: Kathryn will share Zephir quarterly record loading reports with ZAG.

There is not a requirement to have contributing institutions routinely update the records they'd submitted to Zephir.

In Q4, the Zephir team will be looking at revising how Zephir can support HT metadata corrections, a process predicated on current metadata correction policy. Zephir staff currently reports on who is responsible for the records that need correcting when multiple institutions hold (and have submitted records for) a given title. Zephir has set up an item data API and this could play a role in the corrections workflow, though there are permissions issues to address prior to opening this up. The Zephir team will be scheduling a call with the Bibliographic Corrections Group of the HathiTrust User Support Working Group to set requirements.

Bulk of development work in Q4 is to complete migration to AWS. The team is also migrating to GitHub and updating documentation.

CDL Mass Digitization staff, HT CRMS and gov docs registry staff set up a phone conversation also with Zephir staff to compare notes on metadata. They looked at how the metadata they have generated is related to that in Zephir. Questions about whether Zephir could be a metadata store of record and whether HT should have its own record to enhance came up as both the CRMS and gov docs projects include interesting data that could be compared with/integrated into Zephir. The question of whether HathiTrust could/should have its own record to correct/enhance needs to be addressed in a policy.

KS: Authority data could be very useful in Zephir.

ACTION: Kathryn to share notes from this call with ZAG.

JMO: How is clearance happening after CRMS?
KS: I'm not sure what plans may be, but requests for addressing copyright issues come in through the HathiTrust User Support Working Group (HTUSWG), and would likely continue to be addressed there; some involve metadata corrections. CRMS volunteers and gov docs registry staff contribute requests through the HTUSWG (for both rights and metadata corrections).

TC: Makes a point about authority linking. Zephir should look not only at how to change records but how local records can be connected to authority via linked open data. Could we use VIAF identifiers in Zephir?
JR: There is an alternate direction under discussion. Designating WorldCat records as official record. This could relate/be applied to authority data.
KS: The question of changes being overwritten is one we have to deal with. If there were a HT record, how would we maintain? There is no Tech Services group in HT, though the Bib Corrections subgroup of the HTUSWG is at ten members and does some diagnosis when handling correction requests.

KS: Let's make this a future agenda item: Authority data and policy issue. ZAG will outline the policy issues for MUSAG and PSC. ZAG will flesh out the Operational impetus for the policy.

JMO: We should look beyond VIAF as well. What are the implications for maintaining metadata?

ACTION: We'll add "authority data and policy issues" a discussion of the implications for maintaining metadata (including that sources as linked open data?) to the next regular ZAG call agenda (11/16).

ACTION: Kathryn will share a link to the presentation that she and Stephanie Collett (Zephir Technical Lead) delivered at code4lib 2015 on the challenges identified in consuming linked open authority data.

Fire drill at CDL interrupted the call. Some were able to resume for a very brief MUSAG update and to tentatively schedule a follow-up call.

2. MUSAG update

Todd reported that the MUSAG had their first, kickoff call on October 9th, which focussed primarily on the business of the group.
The group went over anticipated MUSAG/ZAG interactions and identified that Todd and Patti Martin (CDL) will be the conduit between the two groups. Patti will bring action items from the MUSAG to Kathryn and the ZAG for weighing in.

Some of the initial work of the MUSAG is of relevance to the ZAG:
- Conduct an environmental scan to look at HT metadata stores
- Develop a Metadata Use and Sharing Policy - this is relevant to the earlier ZAG discussion on whether there could/should be a "HathiTrust record." A draft of the policy has been circulated to MUSAG members.

ACTION: Kathryn will forward MUSAG notes from Patti Martin to the ZAG.

The following agenda item will be addressed on a future call:

3. SLA 4.3.1-4 policies and procedures
   - Discussion: Common questions, Procedural issues, Policy implications
     - **Please continue to review, comment on, and add to [this document](#), which compiles commonalities among our work thus far and will help frame our next steps.**
     - Consider development request scenarios
   - Next steps
     - Identify policy areas to bring to the attention of MUSAG
     - Propose revisions/refinements to [Pre-ZAG Procedures for Zephir Data and Functionality Requests](#)
     - Re-group on specific policy/procedural areas?

ACTION: Kathryn will schedule a tentative continuation call on Wednesday, November 4th,, also to be at 10am PDT/1pm EDT.