

2019-10-04 Cornell LD4P2 Meeting notes

Date: 04 Oct 2019

Attendees: Jason, Steven, Tim, Adam, Lynette, Simeon

Regrets:

Agenda & Notes

Review actions from 2019-09-27 Cornell LD4P2 Meeting notes

- [E. Lynette Rayle](#) will describe current state of wikidata work (as identified on [Prioritization page](#)) and then ask [Christine Fernsebner Eslao](#) what we might consider "done" for the next pass on wikidata, then perhaps create new and more specific issues for later improvements
 - 2019-10-04 Not done yet, still need to check in with Christine (have been working on other items higher in prioritization)
- [Tim Worrall](#) to update discogs code to use standard bnode format
 - 2019-09-13 Tim about half way through this
 - 2019-09-27 Have updated discogs code to output ntriples as well as n3, will make PR for this and also the update to use bnodes
 - 2019-10-04 Got back from Justin a snippet of n3 that now works with Sinopia, and have now replicated this load of n3. Sinopia team are updating the label to list all formats that can be accepted. Have made PR which is awaiting review from [E. Lynette Rayle](#). Have found a new issue with loading Discogs data in that certain extra labels and types break the loading of data need to identify minimal example(s) to post bug report for Sinopia and potentially use for a workaround. This seems like something that is a general issue and will be a problem with other data sources such as SVDE. General idea is that extra triples should be ignored or loaded without showing, should not cause load to fail
- [Tim Worrall](#) will create issue to describe that Sinopia doesn't accept n3, it accepts only ntriples see above, that has changed
- [E. Lynette Rayle](#) also to switch output to ntriples instead of n3
 - DONE in Tim's PR, awaiting merge
- ALL - review , KPAOW [Usability results write up](#)
 - 2019-10-04 Also draft partway done for KPAOW [lessons learned and possible future steps](#)
 - Huda will walk us through these document Oct 18

Status updates and planning

- Blacklight LD @ Stanford
 - Awesome organization by Jessie and Huda, attendees needed lots of support and facilitation after 2.5 days (it was a pretty intense meeting and people did a great job participating), only a few people stayed for the half of the 4th day
 - Jessie, Mark, Chris (Stanford) and Huda are going to BL meeting next week and will take ideas forward
 - Succeeded in synthesizing scenarios with a good mix of BL and other folks (e.g. VuFind, EBSCO, Ghent autocomplete, LC)
 - Specific suggestion to implement schema.org beyond microdata to have a JSON-LD serialization as focus of BL output
 - Outputs
 - [Blacklight LD Working Meeting Google Drive](#)
 - [Blacklight LD Working Meeting Executive Summary](#)
 - [Blacklight LD Working Meeting Notes](#)
- Prep for Cataloging Sinatra and other 45's (Discogs data, https://github.com/ld4p/qa_server/issues?q=is%3Aissue+is%3Aopen+label%3ADiscogs)
 - AWAITING more work in Sinopia to import data. Sinopia work cycle 2 will we hope include the ability to read in RDF back from Trellis. We hope that we can leverage this to import RDF from a lookup in Discogs or ShareVDE. STARTING with items not in Discogs
 - Had kick-off meeting on Wednesday (Oct 2) and will follow up with a review of BIBFRAME and how it will be used for this work, and training around the tooling. Catalogers has accounts with Sinopia etc.. Steven adding Sinatra training material to the Sinopia drive space, immediate feedback is that some understanding of the model is required to be able to think about things like work vs instance title
 - Finding issues with sub-entities being tied to profile for the parent limiting editing flexibility (e.g. instance is tied to the particular work profile)
- Enhanced Discovery (see also <https://wiki.duraspace.org/x/sJI7Bg> and <https://github.com/LD4P/discovery/projects/1>)
 - BAM! (to run through end October)
 - For reference: [Initial draft of BAM plan](#)
 - Did some evaluations on mockups at Stanford last week with freshmen undergrads, some of whom had never yet used the library catalog. Included broader and narrower in subject facets; and offering different browse axes for author, subjects, geographical and timeline
 - In item view, Tim looking at call number browse and expanding work on subject browse started in KPAOW!
 - For timeline, John looking at subject era facet that depends on FAST which has labels and connections
 - Huda looking at querying subject headings, discussed with Dave and running long queries to extract data set. Looking at fairly coarse grained set of top level subject headings as a possible entry point
 - For geographical browse looking at subject areas for regions. In our current catalog the data is quite sparse
- Authority Lookups for Sinopia (Lookup infrastructure: https://github.com/LD4P/qa_server/projects/2, Authority requests: https://github.com/LD4P/qa_server/projects/1)
 - Lynette will keep focus on QA next week
 - RDA registries included as a single authorities with sub-authorities for the different areas, now available in Sinopia. This has been straightforward because of consistent SKOS model and clean data
 - Open question about cohort members who want to use LC versions of RDA vocabularies. Discussion on id.loc.gov email list and Matt Miller asking for feedback on data dumps
 - Performance – analysis on laptop clearly shows that from QA server perspective the http load is the dominant part of the problem with the slowest third of requests taking 2.5-6s for load. The normalization part is next slowest, taking ~0.8s for the slowest third. Need to check analysis is correct and also run on AWS. Will discuss the load portion with Dave. And also do swarm test.
- Travel and meetings (see [LD4P2 Cornell Meeting Attendances](#))
 - Blacklight Summit will be at Duke, 9, 10, 11 October at Duke

- Huda
 - Samvera Connect, week of October 21 (WUStL)
 - Lynette to present on QA (Simeon also going but not on LD4P funds)
 - Fall partner and cohort meeting in DC, November 12/13
 - [Cohort and Partner Meetings#3: November 12-13 2019, Library of Congress](#)
 - Tim will be on D&A sprint that week, everyone else should plan on attending
 - Team bus will leave Ithaca ~2pm Monday, return Thursday morning. We'll stay around Dupont, perhaps Holiday Inn or Darcy next door? Simeon will get the bus
 - 5th International LODLAM SUMMIT at the The Getty Center in Los Angeles. February 3-4, 2020
 - Steven is on the planning committee, Lynette and Simeon to apply
 - Expect to have a "tool challenge" - a competition before the conference
 - LD4 Conference at College Station, TX (TAMU) - May 13/14, 2020
 - Jason is on PC
 - rdfs:seeAlso [Conferences Related to Linked Data in Libraries](#)
- Next meetings:
 - 10/11 Huda @ BL summit