

2019-02-22 Cornell LD4P2 Meeting notes

Date: 22 Feb 2019

Attendees: Jason, Steven, Tim, Huda, Simeon

Regrets: Lynette

Agenda & Notes

Review actions from 2019-02-15 Cornell LD4P2 Meeting notes

- **Steven Folsom** keep track of cataloger perspective on work to support LC Names in QA with context
 - 2019-01-11 - Dave had added LCNAF to D.A.V.E. and Steven is working on the context needed for the search lookup UI.
 - 2019-02-01 - Action yesterday. Working with Dave on whether that spreadsheet has all the right information.. may be resolved today?!
 - 2019-02-08 - Answer to previous maybe is "no". Need input from **David Eichmann** in order to resolve question of whether spreadsheet has all the right information (particularly about need for prefLabel or not, language tags or not) **Steven Folsom** will write up new issue and **E. Lynette Rayle** will review; **Steven Folsom** will then follow up with Dave
 - 2019-02-22 - Tools call this week was Steven, Lynette and Dave. Went through the spreadsheet and **David Eichmann** is going to go through the top 3 issues (https://github.com/LD4P/qa_server/issues/53 https://github.com/LD4P/qa_server/issues/50 https://github.com/LD4P/qa_server/issues/47 – no updates on these since Tuesday)
- **E. Lynette Rayle** priority is completion of https://github.com/LD4P/qa_server/issues/27
 - 2019-01-25: significant progress. refactored code showing context; fails on certain LD paths and doesn't expand out URIs (e.g.: if do broader, will only give URI, not URI + label) - has not yet attempted to correct.
 - 2019-02-15 have 1 PR outstanding, will then have context with expansion if requested in config. Will then call refactor done and make major QA release
 - https://github.com/LD4P/qa_server/issues/27 closed, refactor complete
- **E. Lynette Rayle** LDPATH gem is doing load for graph every time it tries to access something (SLOW)... may need to fork LDPATH gem to make more efficient since will not need to reload graph every time (for our use case)
 - Still issue with LDPATH gem, working Chris Beer to resolve
- **E. Lynette Rayle** (from discussion of LC Names in QA with context) 1) investigate & document algorithm for selection of languages with label and /or no label. In cases like AgroVoc where languages are handled well then we just want to document the (good) algorithm; but in other cases where language information is wrong we need to be able to turn that off (need more config for this)
 - 2019-02-22 no update
- **Huda Khan** & **Steven Folsom** to go through spreadsheet <https://docs.google.com/spreadsheets/d/1gklgTQUNd9n3xTf5kdFYK51SRG6TeMRH678N1FtnQ0/edit?usp=sharing> assessing difficulty and "linked data wow" factor (BfB)
 - 2019-02-08 no update, will schedule "bang for buck" (BfB) meeting next week
 - 2019-02-15 had meeting to go through spreadsheet, expect another meeting to fill out BfB/wow column. On **Steven Folsom**'s plate decide whether another meeting with just **Huda Khan** or a larger group
 - 2019-02-22 Discussion of whether we are doing anything that will impact our source (MARC) data or whether this work is just additions at the discovery layer (that may come from other sources such as SHARE-VDE clusters) – agreement that we are interested in experiments that will merge any useful data at the indexing stage. If we wanted to move any of this to production then we'd need to have conversations about linking in MARC but that could be done later / if / when. **Huda Khan** and **Steven Folsom** will meet next week to come up with proposals with an assessment of tractability and wow factor.
- **Huda Khan** is working on whether usability work on catalogs has been done internally as well as externally to CUL. Code4Lib journal and another has some details
 - 2019-01-18: Have been looking at usability work and literature. Found some older material on Blacklight tests (when facets new), not so much after that. Some log analysis on SearchWorks. Also looking at linked data browser tools. Interesting dissertation from a few years ago on LD exploratory search. See [discovery folder](#), literature exploration [NOTES DOCUMENT](#), [initial discussions](#) with Tim regarding knowledge panels etc. , and also Stanford [NOTES FROM LAST SUMMER](#)
 - 2019-02-15: Astrid is doing more user sessions in the next week or so. Huda will then review. Also discovery discussion group had a [meeting](#). For reference, [knowledge panel slides](#) with screenshots/examples plus questions
 - 2019-02-22 **Huda Khan** will schedule a DOG (discovery on the ground) meeting with goal of working toward plans/proposals for implementation (DIG) (flow DAG DOG DIG DUG)
- **E. Lynette Rayle** will work with Dave on Video Game Genre Forms request https://github.com/LD4P/qa_server/issues/42 – pending work from Dave
 - 2019-01-18 Still in Dave's camp
 - 2019-02-08 no progress. In **David Eichmann** camp until data is added to DAVÉ suggest that we should add these to the Tuesday meeting. ACTION **E. Lynette Rayle** to add to Tuesday meeting agenda
 - 2019-02-22 See first issue, other items were prioritized for **David Eichmann** so circle back to this later
- **Jason Kovari** will discuss in LTS whether there are URIs from Share-VDE data that are useful and might be added to catalog
 - 2019-02-22 ON HOLD until we have the real (with clusters) Cornell data back from Casalini (expected early March). Then have discussion in LTS – people will need to see / explore data before forming opinions on utility
- **Simeon Warner** re: Enhanced Discovery work, will talk with Jessie about Stanford plans and then talk to budget & finance
 - 2019-02-22 Still exploring remote work vs contract work, will need budget change approval
- Issues: [Infrastructure Authority Requests](#) <https://waffle.io/LD4P/ld4p2-cornell>

Status updates and planning

- Cohort engagement
- Discovery developer search
- Authority Lookups for Sinopia (Production QA instance: https://github.com/ld4p/qa_server/issues/)
 - Questions (not expressed in https://github.com/LD4P/qa_server/issues/25) of how we index SHARE-VDE data. Doesn't seem feasible to do this differently for different institution's data but we don't have a good understanding of the plan to support copy cataloging. Casalini are providing institutions their own data but we expect to index for lookup the entire set of BF data

- We don't know whether two MARC BIB records that are "the same" (e.g. same OCLC number) from different institutions will end up becoming one or two BF Instances?
- Prep for Cataloging Sinatra and other 45's (Discogs data, https://github.com/ld4p/qa_server/issues?q=is%3Aissue+is%3Aopen+label%3ADiscogs)
 - Tim working on Rspec tests for QA
- Sinatra Profile Work
- Enhanced Discovery
- Travel and meetings (see [LD4P2 Cornell Meeting Attendances](#))
 - LD4 conference and pre-meeting with Harvard folks
 - Expect to send: Jason, Steven, Huda, Lynette, Simeon
 - Simeon has reserved minivan for Wed 8 through Sun 12 May
 - Jason to discuss whether to spend day or half-day with Harvard folks on Thursday (topics: Wikidata, reconciliation)
 - Everyone work on agenda document.
 - LD4P June 2019 Meeting: Agenda formation: <https://wiki.duraspace.org/x/p4B7Bg>
- Next meetings: