2018-12-07 Cornell LD4P2 Meeting notes

Date: 07 Dec 2018

Attendees: Jason, Steven, Lynette, Huda, Simeon

Regrets: Tim

Agenda & Notes

Review actions from 2018-11-16 Cornell LD4P2 Meeting notes

- Steven FolsomLC Names is in QA without going thru context workflow. Steven Folsomwill work with Huda Khan to set up.
  - 2018-12-06 - LC have updated LCDGT dumps but have not yet updated the dumps for LC Names (with or without RWO). Steven will follow up again. It seems that qtong announced addition of RWOs to LCSH but not to LCNAF which is the one we really want
- Jason Kovari Starting work on getting records to Casalini, will need to work out how to stage data we get back from Casalini
  - 2018-11-02 - CUL-IT and LTS are working on this - pull all our bibliographic records is cumbersome to process; on the queue for Pete Hoyt and Gary Branch
  - 2018-11-16 - think that this is being worked on by Pete/Gary
  - 2018-12-06 - Done, sent to Casalini. Don't know schedule for getting anything back.
- Simeon Warner - set up homework to brainstorm D&A ideas on Google Doc, then set up longer session together to flesh out. Need to scan the large set of D&A issues: https://culibrary.atlassian.net/secure/RapidBoard.jspa?rapidView=19&projectKey=DISCOVERYACCESS&view=planning.nodetail
  - Simeon still making no progress on brainstorming homework
  - Jason will take over and start on this, consider input from other LD4P2 partners too
- Issues: https://waffle.io/LD4P/ld4p2-cornell

Status updates and planning

- SWIB debrief
  - "Good stuff", day 2 >> day 1
  - Interesting vocabulary talks: Osma (Finland) (abstract) doing alignment with LCSH; also Lucy McKenna (Ireland) (abstract) doing alignment with detailed provenance information. In Sinopia we expect to have some alignment with sameAs but do we always have a local entity?
  - SemSpect viz tool: http://www.semspect.de/ - seems to be aimed more at data owners than end users, could be useful sort of exploration for daa cleaning
  - Steven/Huda LD4 talk - question about links to authorities opening in iframe vs window; a number of people interested in form specifications (e.g. Martin Schultz day 2 talk)
  - Wikidata workshop was good (link to draft curriculum), interesting tools: one (wdtaxonomy) to see what properties and classes there are, in part to address question of "what is the taxonomy?", different communities often pick and describe the properties that they will use but this isn't reflected in Wikidata itself
- WikiCite debrief
  - Michelle's WikiCite 2018 Takeaways
  - Full program with links to recordings of presentations and notes from sessions: https://meta.wikimedia.org/wiki/WikiCite_2018/Program
  - Day 2 discussions of possible scope for wikicite interesting: is it ensuring all citations are actual citations, use of wikicite to build curated collections (ala Bambaata work), or trying to have all citations in the world (see moonshot etherpad)? Still not at all clear what Wikicite is or will be. Wikidata has just 40k works at present, very small
- Authority Lookups for Sinopia (Production QA instance: https://github.com/ld4p/qa_server/issues/)
  - QA Priorities: logging of use (not done yet Issue #28), then finish refactoring (no recent work), then performance (some reporting added) and/or new authorities
    - How to implement and check accuracy tests supplied by catalogers? Process will rely on Lynette copying test
  - Lookups for LC NAF: experimenting with regular lc call but with Ajax requests to return context. Have RWO and sameAs coming back in search result-like display and next step is to do additional Ajax calls for rwo info (3xx fields), to generate prototype that can be used in testing. Huda is going to change prototype to work against QA for first search, but will then make other AJAX calls to LC to experiment with context data
    - Huda Khan and Astrid user tests & plans...
    - Overview/summary of findings and results
  - Wikidata as an authority
    - Christine and Steven discussed where to start with wikidata for authorities and think using Events is the bast place to start. Steven has created a repo https://github.com/LD4P/Wikidata – Steven thinking about Events and Christine/Marc thinking about Works. From initial exploration it seems that data from wikidata events might be more useful as extra information to display rather than as the basis for query; will likely need to look for several related properties as data not consistent
  - Lynette to follow up with YAML for https://github.com/LD4P/qa_server/issues/13 and https://github.com/LD4P/qa_server/issues/26, then Steven will shepherd the process with the person requesting. Also with Christine and Mark around Wikidata
  - What would it mean to "productionize" QA & Dave's servers for use beyond the LD4P2 cohort? Are we recording everything in the LD4P2 github repository? How would we containerize all components or perhaps replicate Dave's server on different infrastructure? How might we support higher availability and/or easy local replication and customization? – ACTION - Simeon to mention to Michelle and perhaps start a list of topics for our next partner discussion/meetings
- Prep for Cataloging Sinatra and other 45's (Discogs data, https://github.com/ld4p/qa_server/issues?q=is%3Aissue+is%3Aopen+label%3ADiscogs)
  - Steven making some small changes in the annotations of the discogs JSON, Tim tracking this in the repository
  - Tim completed issue 34 and has begun work on issue 35.
    - ACTION - Steven Folsomto follow up with Tim Worrallabout status of Discogs work and how it fits with expectations, E. Lynette Rayleto follow up about whether we can publish Discogs lookup as part of QA because we think others might be interested in using it
- Enhanced Discovery

- Travel and meetings
    - LD4P2 Cornell Meeting Attendances

Next meetings:

- Jason or Steven to lead 2018-12-14, Jason to lead 2018-12-21 & 2019-01-04
- Simeon out for 2018-12-14, 2018-12-21, 2019-01-04
- Tim out 2019-01-04