

2020-09-11 Cornell LD4P3 Meeting notes

Date: 11 Sep 2020

Attendees: Greg, Tim, Jason, Huda, Steven, Simeon

Regrets: Lynette

Actions from 2020-08-21 Cornell LD4P3 Meeting notes

- [Simeon Warner](#) to coordinate work on final report with Michelle
 - Due Aug 31, not yet sure about coordination around partner report
 - 2020-09-11 - Still to do
- [Huda Khan](#) to reach out to Tracey (cc. Jason) to review/refine discogs plan before taking to Q&A (after September 2nd)
 - 2020-09-11 - Tim and Huda met with Tracey yesterday. She thinks that bringing in information for display is useful, would be better if some information is brought in for discovery (e.g. publisher). [Notes](#) - Takeaways:
 - Discogs display as it works now (as supplementary information) seems like a beneficial feature
 - Conversation around how supplemental discogs info is tied into discovery would be useful to have. Discovery involves MARC, indexing, etc. so there are questions around whether we are considering a batch update after the main cataloging workflow is completed?
 - Additional language needed around the display so that end users information (in the case of display-only)
 - ==> AGREEMENT - 1) Move ahead with display of additional data; 2) Continue discussion of what it would mean to add this for search (either add to MARC or just in Solr)

Agenda

- Discovery (WP3)
 - <https://github.com/LD4P/discovery/projects/2> for issues etc.
 - Preliminary draft of a discovery plan- intended to get feedback: https://docs.google.com/document/d/1zKYW7FQVVNvyd0XjjW0qWznX9PC3jbmOE6Kz_yygPjs/edit?usp=sharing
 - Which to start with?
 - Discogs - everyone likes this, are we OK showing the data. D&A Reps + Tracey would be a good group to think about this.
 - Autosuggest - need to clarify added value of linked data over what could be done with current indexes, and address updates
 - Open Syllabus - can we rely on that project? are there other (linked) data sources
 - Knowledge Panel -
 - How much does bang-for-buck influence our decision? E.g. discogs would be great for music but that is a relatively small number of items
 - Strand 1: production piece (i.e.: picking features we've already worked on to then push those into production): discogs, autosuggest. Feature branches in the code to work on features in isolation. Still need to discuss discogs with Tracey (post- 9/2). Preliminary pass with discogs features using latest production (from a few days ago) successful (tried on dev vm and will try on ld4p3 demo using continuous integration later)
 - Steps:
 - Discussion with Tracey re. use case and benefit – DONE
 - LTS engagement re. metadata
 - ACTION: Jason will discuss plan for display implementation with LTS directors (note similar to use of external ToC data)
 - ACTION: Jason and Steven to explore where/how/if a discogs id might be added to MARC that would then be used in place of dynamic search to find a matching record (and coordinate with Tracey re. best practice)
 - Production requirements and functionality – [Production decision points](#)
 - ACTION - Huda and Tim to discuss with Adam how work might be coordinated with D&A team
 - Discussion with D&A User Reps and dev team
 - Strand 2: research: how to go from knowledge graph to an index - what decisions are needed. What are the data sources for each (e.g.: how many Cornell faculty in wikidata)? Present: reviewing data sources and questions. Should have more worked out in a week or two. Main areas of concern: browsing and dashboard... and anything we can do to help patrons navigate our collections and how we can highlight an entity... and what does that mean for the index? How do we capture the relevant bits of the graph for an index? Is there a repo for this? Not yet... but can use discovery repo (<https://github.com/ld4p/discovery>) we already have to capture any queries or related work.
 - [Research decision points](#)
 - ACTION: Huda to lead work to consolidate use cases for browsing and dashboard, and refine questions with DOG, look at data sources. Intend to review use cases 2020-09-18
 - Discussion of connection to spotlight resources, our highly curated views
 - Can we get Blacklight fork to not hit the production catalog every time we do a pull request? Not clear way to specify that all pull requests should go to the fork rather than the main branch from which you forked. Huda is investigating with input from Code4Lib inquiry
 - 2020-09-11 so far there appears no way to fix this from within the github UI, will explore command line options
 - Where do we maintain queries and data sources for new discovery work?
 - Previously agreed that we will add data analysis queries in <https://github.com/LD4P/discovery>, nothing to do now but use this when starting on the dashboard
 - LD4P3 demo blacklight site
 - 2020-09-11 - Have got Jenkins to pull from our branch now, some kinks still to work out. Servers are shut down at 6pm to save costs, some issues with Solr not coming up properly at 7am and Greg having to go in a fix
- Linked-Data Authority Support (WP2) - A key element of this work package is a sustainable solution that others can deploy. Questions of budget for deployment. Need to get all code into LD4P repository. What would a good end-product look like both for our maintenance and for others to use
 - [Qa Sinopia Collaboration](#)

- 2020-09-11 - Container for qa_server is up and running. All pages confirmed to work. Need to clean it up and create documentation.
 - Next working group meeting is Monday, Sept 14
 - [Cache Containerization Plan](#) agreed with a prioritization of containers
 - 2020-09-04 - Tried switching to Postgres instead of MySQL. Still having issues. Tried to get Hyrax docker running as a sanity check. It also had problems. Tom said that Docker on Mac has been problematic for folks. He works on Ubuntu. Tried to install VirtualBox VM to be able to run Ubuntu, but the install failed.
 - [Search API Best Practices for Authoritative Data working group](#) is still working through use cases, this seems very important to take sufficient time
 - 2020-09-04
 - Working on developer user stories. Not as clear cut as expected. Some confusion around UI user stories and API user stories. Working on reorganizing to those two sets.
 - Need to get contact information from Steven to share with PCC. Looking at creating a survey that will allow for prioritization by PCC and the working group.
- Developing Cornell's functional requirements in order to move toward linked data
 - C.f. Stanford functional requirements document: https://docs.google.com/document/d/18H6zYGwKuKg3SZqm9Q_cxkZThcdmBjknE6HdtQ-RRzk/edit#heading=h.4fu64x8jzm6e
 - What does success look like? And then how do we get there?
 - 2020-09-11 Jason drafted an introduction for LTS staff and no capacity right now in LTS, postpone at least another week
- Other Topics
 - OCLC Linked Data / Entities Advisory Group
 - Current communication focused on UX
 - PCC Sinopia Profiles Working group
 - Going through large spreadsheet comparing Sinopia profiles with BSR and CSR, close to having a comparison, will be part of report to POCO
 - 2020-09-11 - Fewer meetings recently because of COVID and/or fires. A small number of outstanding questions that might go to POCO, still on target to complete in October
 - PCC Task Group on Non-RDA Entities
 - 2020-09-11 - Preliminary report went to POCO suggesting a small vocabulary of entity types, discussion about how to host /implement
- Default branch name
 - Instructions for creating a [New Repo with Main Branch](#)
 - 2020-09-04: As of Oct 1, on github all new branches will be main unless you set the default in organization: <https://github.com/github/renaming#:-:text=On%20October%201%2C%202020%2C%20any,this%20change%20at%20any%20time>.
- Upcoming meetings
 - DCMI 2020 Virtual: <https://www.dublincore.org/conferences/2020/programme/>
 - Virtual Blacklight Summit in early October. (Main site [link](#) and [document](#) to add lightning talk/breakout session ideas) - Huda is on unofficial committee. Expecting various follow-ons from earlier meeting this year. (Discussions still in progress and, at some point, draft program will be shared with the community).
 - Agenda and registration link to go out soon. Schedule if a few hours each day with opportunity for lightning talks
 - European BIBFRAME Summit; registration now open. Occurring 9/22-9/23 (9-11:30am Eastern each day): <http://www.casalini.it/bfwe2020/> (no cost)
 - SWIB
 - Discovery submission accepted

Next Meeting(s), anyone out?:

- ...