

2020-09-18 Cornell LD4P3 Meeting notes

Date: 18 Sep 2020

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

Regrets:

Actions from 2020-09-11 Cornell LD4P3 Meeting notes

- [Simeon Warner](#) to coordinate work on final report with Michelle
 - 2020-09-18 - Still to do
- [Jason Kovari](#) will discuss plan for display implementation with LTS directors (note similar to use of external ToC data)
 - 2020-09-18 - It is on the queue for LTS directors, expect to have something next week
- [Jason Kovari](#) and [Steven Folsom](#) 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). Would be used to override dynamic search
 - 2020-09-18 - Still to do
- [Huda Khan](#) and [Tim Worrall](#) to discuss with Adam how work might be coordinated with D&A team
 - 2020-09-18 - Still to do
- [Huda Khan](#) 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
 - See [Use cases](#)
 - ACTION - [Simeon Warner](#) to find out who we might talk to about librarian instruction at CUL re. possible use cases

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
 - 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 -
 - Production requirements and functionality – [Production decision points](#)
 - Discussion with D&A User Reps and dev team
 - Next steps: meeting with Adam, develop plan for how to move toward production and arguments about benefits
 - 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](#)
 - [Use cases](#) - first review 2020-09-18
 - Interest in the dashboard idea (full page for entity) that extends on the idea of an embedded knowledge panel
 - ACTION - work to continue to explore the dashboard
 - 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
 - LD4P3 demo blacklight site
 - 2020-09-11 - Have got Jenkins to pull from our branch now, some kinks still to work out.
 - 2020-09-18 - Had some trouble with updates for Solr index, also coordinating with Robbie on disaster recovery process. Hope to have index up and running again today. Don't yet know whether we have fixed the issue with servers being shut down at 6pm to save costs, some issues with Solr not coming up properly at 7am. Will see when indexes are stable
- 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](#)
 - 2020-09-18 - Met with Stanford staff and ShareVDE on how Sinopia will send data to ShareVDE. Next meeting will be in two weeks (09-29) on how to access CKB, whether through cache or direct through ShareVDE API. API structure tentative. Trying to encourage them toward creating a QA compatible API, but even if they are compatible, there are some concerns about whether it will be performant.
 - [Cache Containerization Plan](#) agreed with a prioritization of containers
 - 2020-09-18 - QA service can be brought up through docker. Need to determine the best way to connect authority configs to the service. Need to understand what needs to happen for it to be deployed to AWS in a production environment. Need to write instructions on how other institutions use this.
 - ACTION - [Greg Delisle](#) to explore how to get the container version that Lynette has on her laptop up on AWS. Question about how to handle environment variables outside of the checked in docker container so that users can config
 - [Search API Best Practices for Authoritative Data working group](#) is still working through use cases, this seems very important to take sufficient time
 - 2020-09-18

- Continuing work on developer user stories. Made distinctions between UI development and backend development. Sometimes got off on tangents about how a user story would be implemented and whether it was possible to implement. Tried to redirect back to the central question of whether the user story was desired. We can work on the details of how once the user stories are prioritized.
 - 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. Should the user stories for PCC only be the cataloger user stories?
- Developing Cornell's functional requirements in order to move toward linked data
 - C.f. Stanford functional requirements document: https://docs.google.com/document/d/18H6zYGwKuCg3SZqm9Q_cxkZThcdmBjknE6HdtQ-RRzk/edit#heading=h.4fu64x8jzm6e
 - What does success look like? And then how do we get there?
 - 2020-09-18 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
 - Discussion of comprehensiveness and accuracy, "Minimum Viable Entity". There was a presentation at DCMI
 - PCC Sinopia Profiles Working group
 - 2020-09-18 - Still going through large spreadsheet comparing Sinopia profiles with BSR and CSR, close to having a comparison, will be part of report 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
 - 2020-09-18 - Meeting today. Use multiple types rather than combinatoric explosion
 - Id4l.org website - ACTION - [Simeon Warnert](#) to follow up about getting redirect to wiki gateway page and to decommission drupal site
 - Content is in Wayback machine https://web.archive.org/web/*/ld4l.org , Jason will see how comprehensive
- Default branch name
 - 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.>
 - 2020-09-18 There is a final report from the Samvera group. Seems sensible to wait now until the github tools are available. Expected January 2021. Have changed default for new repos in ld4p to "main" <https://github.com/organizations/LD4P/settings/repository-defaults>
- Upcoming meetings
 - DCMI 2020 Virtual: <https://www.dublincore.org/conferences/2020/programme/>
 - DONE this week
 - Presentations have been recorded to deal with timezones and will eventually be available openly
 - Huda learned that IIF *is* Linked Data, Glen Robson talked about pulling in annotations. Also [this talk](#) was interesting
 - 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](#) have gone out. Will be intro workshop, time for lightning talks and longer sessions. Sessions for plugin discussion including call-number browse
 - European BIBFRAME Summit; registration now open. Occurring 9/22-9/23 (9-11:30am Eastern each day): <http://www.casalini.it/bfwe2020/> (no cost)
 - Steven will be talking on the panel in the second day
 - SWIB
 - Discovery submission accepted
 - Programme still in progress (not up on website yet: <https://swib.org/swib20/programme.html>)
 - <https://kula.journals.publicknowledgeproject.org/index.php/kula/announcement/view/1> . Call for Proposals - Special Issue: "The Metadata Issue: Metadata as Knowledge". Due January 2021 (abstract 300-500 words). Includes "The use of linked open data to facilitate the interaction between metadata and bodies of knowledge" and "Cultural heritage organization (libraries, archives, galleries, and museums) and academic projects that contribute to or leverage open knowledge platforms such as Wikidata"

Next Meeting(s), anyone out?:

- ...