2020-10-02 Cornell LD4P3 Meeting notes

Date: 02 Oct 2020

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

Regrets: no regrets!

Actions from 2020-09-25 Cornell LD4P3 Meeting notes

- Simeon Warner to coordinate work on final report with Michelle
  - 2020-09-25 - This is really happening now
  - Additional - Tim Worrallto shepherd D&A PR https://github.com/cul-it/blacklight-cornell/pull/1300 through and then arrange merge back into our fork so that there is a clear Apache 2.0 license on the LD4P2/3 Blacklight work
    - DONE - See https://github.com/LD4P/blacklight-cornell
- Jason Kovari will discuss plan for DIscogs 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
  - 2020-09-25 - Have introduced topic but more discussion required. (Relates a little to discussion of current work to index TOC data from a vendor)
  - 2020-10-02 - JAK sent a note to LTS Directors; no concerns raised. We did start a conversation around LTS' role in external data within Blacklight (or SOLR) and whether there should be a role for LTS but this is side-lined as we do other planning / higher-priority tasks. Can mark this as DONE
- 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
  - 2020-09-25 - Tracey/Laura/Jason have been discussing and have proposed location (024 7_ $2 discogs). Steven has been doing some work on catalog numbers (aka publisher id) in music cataloging
  - Steven will create wiki page for this project and document cataloging decision Using Discogs to Enhance Discovery.
  - ACTION - Steven Folsom & Tim Worrall to follow up on this discussion with Frances about how we surface the id type in the Solr index in a way that will be easy/efficient to use
- Simeon Warnerto find out who we might talk to about librarian instruction at CUL re. possible use cases
  - 2020-10-02 Still to do
- Simeon Warnerto follow up about getting redirect from ld4l.org to wiki gateway page and to decommission drupal site
  - 2020-09-25 In progress, https://culibrary.atlassian.net/browse/DLITSYS-3983, redirect in place but still need to decommission Drupal site
  - 2020-10-02 Greg has asked Jim to decommission the Drupal site
- Huda Khan and Tim Worrall work to continue to explore the dashboard design and requirements
  - 2020-09-25 in progress, will discuss with Astrid & Tim next week
  - 2020-10-02: Kickoff design meeting happened Huda Khan Tim Worrall& astrid will look at different aspects more and regroup in 2 weeks
- Greg Delisleto 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
  - 2020-09-25 Greg is working on that (right now)
  - 2020-10-02 Greg can get it running on ECS but we can't see it directly of from behind load balancer, still working through the details
- Toward a demo and discussion with D&A:
  - Huda Khan & Tim Worrall - Make sure the live demo meets the Discogs API terms of service in terms of linking and ack. – Had discussion with Melissa about how/where we show acks and licensing in general for our Blacklight. Could then link to this from all pages and perhaps a specific link to the details on this page. A bit more to do
  - Tim Worrall Create mockup pages for the cases of 1) a completely blended presentation, and 2) a presentation where the Discogs data is separated out
    - In progress
  - Huda Khan to discuss with Adam having a meeting with User Reps sometime after Oct 5
    - October 16th potential date/time: 9:30 AM. need to send list of names to add to this meeting to Meredith
  - Tim Worrall to look into how we might display some language pointing out what data is not used in search. For example, what if a user sees the record label in Discogs data and then tries searching on that (which they can't)? Per Tracey suggestion
    - Partly contingent upon the presentation approach selected. Think about it more after 2020-10-16 meeting

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 - DONE
    - Production requirements and functionality – Production decision points
    - Demo and discussion with D&A User Reps and dev team – See actions above from last week
      - Have branch with discogs code on dev VM and demo server from LD4P fork, but still issues with the assets folder under Rails not showing current version (these are generally copied in during the Rails deploy process)
- 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
  - Goal of functional prototype for end of year
  - How will we test it? Remote testing would work but will students want/have-time to talk with us? Can we pay them?
- Discovery Affinity Group is looking at design ideas
- Dashboard design meeting kickoff notes - will also try to understand what our data will support or connections to other data sources

- 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
    - 2020-10-02
      - Meeting with ShareVDE this week (09-29). ShareVDE will provide download for CKB and Dave will cache. Dave plans to do a combined search index for all institution data. There was a discussion on change management for CKB. I provided them with documentation on issues related to Change Management.
      - Discussed direct access but that doesn't look possible because of API and performance, so Dave needs to cache
      - Added local authorities for property_type and property_attribute per request from Sinopia dev team. Released on production (per links)
      - Open request to add small authorities from LC. Dave is adding to cache and Lynette will then add config to QA
  - Cache Containerization Plan
    - 2020-10-02 - Greg is working on deploying on AWS.
  - Search API Best Practices for Authoritative Data working group is still working through use cases, this seems very important to take sufficient time
    - 2020-10-02
      - Met last Monday. Continuing work on developer user stories. Wikidata participated in the conversation.
      - 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?
        - Steven thinks we should go ahead with the cataloger use cases for PCC
        - ACTION - channel communication the Steven when Lynette has written up the use cases and created a survey for cataloger use cases
        - Developer use cases will have wider audience but still want to include PCC
- 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?
  - 2020-09-25 Jason drafted an introduction for LTS staff and no capacity right now in LTS, perhaps discuss again next week as to what next steps will be. Jason and Steven will discuss this week
  - 2020-10-02 Jason will collect information to prime brainstorming from LTS, think about brainstorming then discussions
- Other Topics
  - OCLC Linked Data / Entities Advisory Group
    - OCLC has described minimal viable entities and a number of small vocabs – are these new "standards" that we will have to engage with to ensure alignment?
  - 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
    - 2020-10-02 - Submitted interim report yesterday. Feedback requested before deciding on a few directions and working to stay compatible with LC work. Some discussion of how to align with SHARE-VDE data which uses a rather different profile. Committee is supposed to wrap by the end of 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
    - 2020-09-25 Now have a short list if possibilities and close to decision
    - 2020-10-02 Have definitions of short lists, now creating MARC examples. Interest is working the German national library on alignment – perhaps influence that and then adopt instead of creating a new list

- 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
  - NEXT STEP - Use github tools January 2021
  - 2020-10-02 - As of yesterday, new github repos will have default branch named main. This can be configured by organization or by individual to be a name other than main.
- Upcoming meetings
  - LD4P(2/3) update for CUL-IT all-staff
    - Who/what? - Huda notes conflict with Blacklight summit
  - 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 talked on the panel in the second day
    - Some chat drama about RWOs...
    - Lots of questions about community process (some of which PCC is now dealing with)
  - SWIB

- - - Discovery submission accepted
    - Programme still in progess (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?:

- ...