2020-11-20 Cornell LD4P3 Meeting notes

Date: 20 Nov 2020

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

Regrets: Steven

Actions from 2020-10-30 Cornell LD4P3 Meeting notes

- @ anyone - tell Simeon who Steven's new avatar is
- Jason Kovari will kick off work on functional requirements
  - Jason working on pictures of edit workflows to parallel Stanford work, not much new this week
  - At the all-hands on Tuesday there was discussion of setting up a Miro board to capture workflows from participating institutions
- E. Lynette Rayle to finalize the survey for cataloger user stories for PCC feedback and ranking by working group, check with IRB
  - 2020-11-20 - Survey finalized. Slow response from IRB - got an ack but no detail yet, did training and think that IRB won't apply in this case because the purpose of the survey is merely process improvement
- Tim Worrall will sort out code in LD4P3 repo and look at adding button to highlight Discogs field
  - 2020-10-30 Tim is been working on this. Worked to improve speed by moving code from JavaScript to Rail but that had significant slowdown (~1s). Now doing search in the controller and then doing second call in JavaScript. Image comes back with first call. Thinking about timeout to avoid problems if Discogs response is slow
  - 2020-11-20 Tim did demo of the highlight feature to user reps last week. Decision to highlight both the data and label for Discogs fields which Tim has now address. Yesterday Tim pushed code to LD4P3 repo so ready to go. Expect to be able to move to production in January sprint

Agenda

Discovery (WP3)

- https://github.com/LD4P/discovery/projects/2 for issues etc.
- Draft of a discovery plan: https://docs.google.com/document/d/1zKYW7FQVVNvyd0XjjW0qWznX9PC3jbmOE6Kz_yygPjs/edit?usp=sharing
- **Strand 1: production piece**
  - 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 – DONE
    - Implementation work
      - 2020-11-06 Didn't manage to get demo for User Reps on Nov 5 due to other pressures. Still making some tweaks to the layout and highlighting
      - Question of using QA vs going directly - will use Discogs directly for now to avoid adding QA
      - Will do demo next Friday for User Reps
      - Assuming User Reps approve to move forward, will plan to move code into LD4, add issues as part of D&A to add tests. Expect to deploy as part of January D&A sprint
      - 2020-11-20 Tracey has provided examples of bib records with 024 $2 discogs for Frances and Tim to test out their code. Tim asked a great question about how much data we can expect in the bib from Discogs if the Discogs id is in the bib. This has prompted me to reach out to Tracey to establish local a policy for deciding what data is essential to include in the MARC record in order for it to be able to stand alone, and what can catalogers get away with not copying and pasting (and let Discogs feature pull in automatically). – Frances updated index to include this as `discogs_display` field, Tim will check this field for the id which will then go straight to discogs call, very simple change
      - 2020-11-20 Essentially on-hold with this work until January D&A sprint though not much time before then anyway
- **Strand 2: research: how to go from knowledge graph to an index**
  - Research decision points
  - Use cases - first review 2020-09-18
    - First goal: DASH! dashboard (full page for entity) that extends on the idea of an embedded knowledge panel, aim to have functional prototype for end of year
  - DASH! (Displaying Authorities Seamlessly Here)
    - Dashboard design meeting kickoff notes - will also try to understand what our data will support or connections to other data sources
    - https://docs.google.com/document/d/1PgQi3xobsPhr9DUHU_YGeimL1OjNiiTdkiNWb36r3Gg/edit
    - 2020-11-13
      - https://jamboard.google.com/d/1VHkygzzY-NWpwWEpZ9PqFjrnfPMADfTgjh8gFRi2MoE/viewer?f=5
    - 2020-11-20
      - Huda working on prototype, think that we can test as a team the first week of December (though SWIB next week)
      - Huda will work on IRB paperwork for user testing even though we think we will be exempt again
      - Tim will have time after the ESMIS sprint, work with Huda but focus on person entities

Linked-Data Authority Support (WP2)

- Qa Sinopia Collaboration – Support and evolve QA+cache instance for use with QA
  - 2020-11-20 Lynette will check in with Dave about getting the new indexing scheme out in production. Still a few outliers in the accuracy tests to check on with Dave and Steven. Stanford team aren't working on Sinopia right now
- Cache Containerization Plan - Develop a sustainable solution that others can deploy
  - 2020-11-20 Lynette will pick this up again next week, Greg also in

- [Search API Best Practices for Authoritative Data working group](#)
  - 2020-11-20 Lynette working on user stories - grouping by different levels and working out how they tie to cataloger stories (as opposed to developer)

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?

Other Topics

- OCLC Linked Data / Entities Advisory Group
- PCC Sinopia Profiles Working group
  - 2020-11-20 The PCC Sinopia Profiles Task Group Report was discussed at yesterday's POCO call, and it prompted big questions about where in the PCC will these profiles (or the platform agnostic version of them will live and be maintained). There will be a meeting in December with POCO, LOC NetDev, and Sinopia folks to discuss this.
- PCC Task Group on Non-RDA Entities
  - 2020-11-20 The PCC non-RDA report was finalized and submitted to POCO earlier this week. It will be discussed at a future meeting.
- Steven consulted with Kevin Ford on how publication frequency is modeled in BF. This could lead to an added node... classic RDF math, with a layer of abstraction in order to provision for frequencies that change over time and/or need to be qualified with exceptions to the rule. Steven has reached out to LTS folks to see how we query the data in MARC to understand what useful queries would look like (there are 2 places for this data with some variants in our MARC data, not always up-to-date, shows as Frequency field in Blacklight (perhaps from 310 rather than 321, or 008))
- Default branch name - WAIT until we can use github tools January 2021

Upcoming meetings

- SWIB
  - Programme out https://swib.org/swib20/programme.html , registration free but conflicting with Thanksgiving in part
    - Monday keynote
    - Tuesday workshops
    - Huda moderating bibframe session on Wednesday - Paloma speaking on LD4P2 Rare Books work as part of that
    - Huda speaking on Thanksgiving morning
    - Friday breakout and lightning talks
- https://kula.journals.publicknowledgeproject.org/index.php/kula/announcement/view/1 . Call for Proposals - Special Issue: "The Metadata Issue: Metadata as Knowledge". Due January 31, 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?:

- 2020-11-27 - No meeting, Thanksgiving
- 2020-12-04 - Simeon out, Jason will lead