

2021-01-29 Cornell LD4P3 Meeting notes

Date: 29 Jan 2021

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

Regrets: Steven

Actions from [2021-01-22 Cornell LD4P3 Meeting notes](#)

- [Jason Kovari](#) will continue functional requirements efforts, starting with workflow diagramming and outlining environment
 - 2021-01-22: started working on this task; hope to have something to demo/discuss next week

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**
 - Production requirements and functionality – [Production decision points](#)
 - Implementation work
 - 2021-01-22 End of D&A sprint is today, expect push to production next week. Also have credits page to acknowledge OSS and APIs which includes Discogs (JHU have similar page)
 - 2021-01-29: In Production - DONE
 - Assessment: outside of LD4P; consider engaging D&A group with usability testing and/or user feedback. For browse, melissa put in some tracking to monitor user interaction with browse. can we do something similar with Discogs (e.g.: are people clicking the highlight link and/or the link to get to discogs? What have we learned from virtual shelf browse tracking)
 - **ACTION ITEM:** Tim will speak to Melissa about tracking for the Discogs link.
 - **ACTION ITEM:** [Tim Worrall](#) will raise usability testing for D&A queue (don't carry forward in notes as now outside of LD4P)
- **Strand 2: research: how to go from knowledge graph to an index**
 - [Research decision points](#), [Use cases](#)
 - 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
 - 2021-01-22 Demo from Huda
 - Have added "works about" counts
 - Have fixed formatting of items by using bento box display
 - Added display all subjects for map/timeline
 - Discussion of what to do about case with zero work, currently detail page doesn't (can't) show much
 - What is the path of things that should be done before testing?
 - Build out index
 - Use color to differentiate selected vs related in maps/timeline
 - Acknowledge source of images
 - Fix labels on map
 - Fix page title... to "LIBRARY CATALOG" as on other pages
 - Tie features back to testing tasks
 - 2021-01-29: Huda working to get scripts in place to populate index; bringing in period-O info; focused on locations with Wikidata URIs for consistency. Subject headings: script that takes-in components & breaks those out... and parses into timeline info. On Dave's fuseki, 34 distinct temporal terms with labels. Will finish today with actual index. Will break-down the loading to increase load speed
 - **ACTION ITEM:** adding Wikidata URIs for any subject headings and broader/narrower URIs to index (today)
 - Reached out to IRB to ask about testing: if we want to disseminate results as research data, need to do IRB protocol; has a follow-up. Waiting to hear back but will submit protocol if no word. Simeon's interpretation of reply is that we are crossing line into research and the approval will likely be positive. Depending on how we describe what we're doing it either falls under research OR improving a product... but we're essentially doing research to improve a product so yes to IRB review.
 - 2021-01-29: Next piece to go into production? discuss in Discovery on the Ground and then here
 - Tim was looking to reuse some of the ajax calls Huda did for knowledge panel; reworking portions of HTML but focused on getting LoC URI and calling wikidata. making good progress. Considering layout and not duplicating what we already have in catalog. Working on authors/subjects. Wants another week to work on this
 - **ACTION ITEM:** for 2/5, spend time discussing next thread for production stream

Linked-Data Authority Support (WP2)

- [Qa Sinopia Collaboration](#) – Support and evolve QA+cache instance for use with QA
 - 2021-01-22
 - No meeting with Stanford. Met with Steven and Dave to review the [Evaluation of New Indexing Approach project board](#). We moved 6 issues into the Under Exploration column. 5 appear to be related to indexing. 2 need review of the tests to determine

- if they would be expected to pass. 1 required exploration in QA. For the QA one, the issue was there was an actual line break in the triple which is not supported. So the question is why Jena produced a triple with a line break. Is the problem in the source data or should Dave's code change any line break to the proper escape sequence (i.e. `n`)? The plan is to meet for 30 min once a week either before or after the meeting with Stanford to review the board and work our way through the issues so the new indexing scheme can become production
 - 2021-01-29: no Stanford meeting. Still 5 issues; one for GeoNames that does not pass test though target URI is there. Remaining 4 are in Dave's court. Steven checking whether any other authorities have line breaks; Dave considering generalized approach to repair the line break in the triples. Exploring whether problem in Jena Fuseki or if it is an artifact of making this a blob in the indexing process.
- [Cache Containerization Plan](#) - Develop a sustainable solution that others can deploy
 - 2021-01-22
 - Greg has started a new branch for container in ECS. So far ECS will run task, need to work on load balancer config and deal with restart on error/quit. Require ECS service for this
 - Started AWS documentation - in aws templates folder README, need to describe environment file and steps to implement
 - 2021-01-29: lots of documentation and templating work; one critical piece needing completion (getting service that takes all pieces together and keeps container running in service into a template). rest is good. nearing point where we can put in front of others and say 'give it a spin'. Likely finishing up the template of last critical service piece for next week
- [Search API Best Practices for Authoritative Data working group](#)
 - 2021-01-22
 - The group is reviewing the [Supporting cataloger inclusion of external authoritative data](#) document. We will meet on Monday Jan 25 for the final meeting of the first charter. We will also explore topics for the next charter. Potential topics: language processing, change management, moving user stories to practical and specific recommendations, and triple stores and other technologies. The next charter will likely start in March to give folks some time off between and me to prepare for the selected topic.
 - 2021-01-29: met and did informal voting in group for next topics; also has survey (posted in Slack) that has descriptions of each area, potential areas and video discussion in meeting to ensure sense of the potential topics. Will then send to broader group to get sense of what broader community wants as next steps. 4 areas:
 - 1. change management (source data to consumers), 2nd (tie) among group
 - 2. language processing (anything to do with language),
 - 3. moving forward from what has been done (user stories to specific recommendations),-- 1st among group
 - 4. Linked data specific approaches and what implications does it have for our data being in LD, what are the tools and what can be leveraged (producer side)? 2nd (tie) among group
 - 2021-01-29: plus delta done. plus saying everything well organized and a delta saying that everything was really well organized. steering of conversation based on what was brought to meeting ahead of time. Concerns expressed around whether time will be allocated to adopt recommendations
 - 2021-01-29: work done on supporting cataloging inclusion (categorized version of user stories); meeting one more time next Monday to work through rest.
 - 2021-01-29: got to a good place with good output!

Developing Cornell's functional requirements in order to move toward linked data

- C.f. Stanford functional requirements document: https://docs.google.com/document/d/18H6zYGwKuG3SZqm9Q_cxkZThcdmBjknE6HdtQ-RRzk/edit#heading=h.4fu64x8jzm6e
- What does success look like? And then how do we get there?
- Miro board (diagramming): https://miro.com/app/board/o9J_lfXUJ8=
- Notes space: <https://docs.google.com/document/d/1TVPBFak7DkfjBptKI-pCMWQnOaiWHB0XCHswiB3Fr9g/edit?usp=sharing>

Other Topics

- OCLC Linked Data / Entities Advisory Group
 - Request for UI and API testing from Jan 25
 - Lynette has Cornell key (a WSKEY) for testing
 - Call discussed seeding of data. Data for person includes VIAF and other sources; place includes geonames. Steven, Huda, Jason and Lynette signed up for user testing
 - 2021-01-29: User testing happening now
- PCC - Sinopia collaboration
 - 2021-01-15 Looking at support required to engage PCC libraries (documentation, mentoring program, etc.)
- PCC Task Group on Non-RDA Entities
 - 2021-01-15 PCC reviewed proposal but no decisions made yet, looking at description wrt cataloger use, discussion will continue
- Default branch name - WAIT until we can use github tools January 2021
 - Lynette is signed up to be a beta tester for github. Target for public release is end January
 - Github have provided the change button, Lynette has provided feedback as a beta tester. There is an issue that after the change to `main` someone can repush the `master` branch, has suggest a deny list of branch names. Links in comment will get forwarded to `main` unless there is a new branch with the old name
 - For forked repos there will be a message
 - What will it take to make the change to `main` for LD4P repos?
 - Lynette has documented [HERE](#)
 - Lynette will rename for qa_server, work with Greg on qa_server container; document on LD4P wiki
 - For our blacklight fork we should wait until the D&A team change
 - 2021-01-29: no updates. working group for Samvera meeting next week... Samvera is discussing moving away from GH as a whole. QA server investigation – quite a few forks off QA server so not sure it is the best to start - Lynette will reach out to the fork owners and document the aspects. Tech side is not challenging, community side is.

Upcoming meetings

- <https://kula.uvic.ca/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"
 - 2021-01-15 Huda will start a google doc.
 - Folder [Link, CFP + Brainstorming](#)
 - Brainstormed a bit and put together a [draft](#)
 - 2021-01-29: Draft (linked above); Claire has pointers in brainstorming doc and Huda started a narrative thread. Work done to make discovery better, here are examples and then questions about how bringing in data sources relates to how we perceive catalog data curator... and whether that prevents us from taking advantage of making connections more visible. Is this a project update OR a scholarly article that goes into what we think of library data and the space. Consider it an article! ADD COMMENTS BEFORE END OF DAY. Huda will then submit (deadline Sunday)
- [code4lib](#) - virtual next year
 - Steven/Tim/Tracey submitted something about Discogs work
 - Lynette proposed a talk on QA and API WG work
 - Voting happening now
 - Expecting to attend: Huda, Steven, Lynette - Simeon to follow up with discretionary spending and arranging payment
 - 2021-01-22 Simeon to actually do this, registration is open
 - 2021-01-29 Spending approved
- Lynette doing a QA presentation at Samvera partner call in June
- Mann All-Staff : Huda (Dashboard), Tim (Discogs) and Lynette (QA) presenting next Wednesday at Mann All-Hands meeting (2/7)

Next Meeting(s), anyone out?:

- 2/5: all here