

# 2021-01-22 Cornell LD4P3 Meeting notes

Date: 22 Jan 2021

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

Regrets:

## Actions from [2021-01-15 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](https://docs.google.com/document/d/1zKYW7FQVVNvyd0XjjW0qWznX9PC3jbmOE6Kz_yygPjs/edit?usp=sharing)
- **Strand 1: production piece**
  - Production requirements and functionality – [Production decision points](#)
  - Implementation work
    - 2021-01-15 D&A spring has started, Tim has moved Discogs code into codebase. Has removed call via QA to go directly to Discogs. Now in integration environment. Need to sort out ack/disclaimer with user rep feedback, perhaps technical information link
    - 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)
- **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\\_YGeimL1OjNiTdkinWb36r3Gg/edit](https://docs.google.com/document/d/1PgQi3xobsPhr9DUHU_YGeimL1OjNiTdkinWb36r3Gg/edit)
    - 2021-01-15 Demo from Huda
      - Will spend next week finishing up prototype (add show all on timeline, etc.)
      - Question of whether it is best to have "details" link to follow broader/narrower, or just link the label. Question of distinguishing search (usual BL link) or go to entity
      - Comparing page for "Biology" in current interface and demo we see issue with handling large numbers or narrower terms (need "more" or similar), and also would be good to include "works about #" for each term
      - Huda brainstorming how to test: [DOCUMENT LINK](#)
      - What is this page called from a user perspective? Perhaps "subject page", "entity page" (maybe user doesn't know of entities), "bobble head" (because you put it on your dashboard)
    - 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
- Discover Affinity Group
  - DAG recordings are going to Drive, problems with ownership and quotas
  - ACTION - Jason & Simeon to talk about LD4 spaces going forward
    - The LD4P2 spaces are the LD4 spaces for now, changing ownership to Michelle should fix problems

### Linked-Data Authority Support (WP2)

- [Qa Sinopia Collaboration](#) – Support and evolve QA+cache instance for use with QA
  - 2021-01-15
    - Not sure when next Stanford work on Sinopia will start
    - Lynette/Dave/Steven will review project board to decide next steps
  - 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

- [Cache Containerization Plan](#) - Develop a sustainable solution that others can deploy
  - 2021-01-15
    - Greg's work now in main branch. Lynette has updated README documentation for running on localhost
    - Still need to document on AWS
    - Then test on generic linux box
    - Then hand off to Dave and Stanford to test setting it up
    - Later need to work out and document a production deployment (w. restart etc.) and try to get feedback from community
  - 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
- [Search API Best Practices for Authoritative Data working group](#)
  - 2021-01-15 E. Lynette Rayle finish up user stories document [Supporting cataloger inclusion of external authoritative data](#) in early January then consider possible new charter for the group
    - Steven sending out survey results to PCC (before lunch/run today)
    - Lynette has requested feedback on user stories document by Jan 20, last meeting of current group on Jan 25
    - Will consider new charter on Jan 25
  - 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.

## 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](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?
- Miro board (diagramming): [https://miro.com/app/board/o9J\\_lfXUUj8=](https://miro.com/app/board/o9J_lfXUUj8=)
- Notes space: <https://docs.google.com/document/d/1TVPBFak7DkfkjBptKI-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
- 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

## 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](#)
- [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
- Lynette doing a QA presentation at Samvera partner call in June

## Next Meeting(s), anyone out?:

- 2021-01-29 FOLIO community governance meeting next week at 10am, reschedule LD4P3 for 9-10am

