

# 2020-06-05 Cornell LD4P2 Meeting notes

Date: 05 Jun 2020

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

Regrets:

## Agenda & Notes

### Review actions from [2020-05-29 Cornell LD4P2 Meeting notes](#)

- [Simeon Warner](#) - Need to update Zooms for future meetings
  - DONE 2020-05-29 do they show for everyone? – except Simeon!
- [Steven Folsom](#) - To work on prioritizing pending authority requests pending in the QA Issues
  - 2020-06-05 Still to do. Steven will reach out to the requesting folks to check and do the prioritization, then give guidance to Dave if there should be change in his work order

### Status updates and planning

- Enhanced Discovery - **WHAM!** (see also <https://wiki.duraspace.org/x/sJI7Bg> and <https://github.com/LD4P/discovery/projects/1>)
  - See: [Organizing doc](#) and [Pseudonym](#) thought. Updates also on running notes [page](#)
  - Updates from Dev chat meeting
  - Usability - [Document](#) with testing plans. Some initial brainstorming. Plan on sharing tasks (perhaps with related screenshots) with Kevin to get some feedback. Will remind regarding putting our usability work on the agenda of the Usability Working Group for next week (week of 2020-06-08).
  - Indexing - Huda worked on trying out the index-heavy approach for handling pseudonym/see also matching and display on the dev VM. Employed a two-phase approach (i.e. index populated first with information and then updated in a second pass.) Tim tried out the front-end with the dev VM index and appears to work. Currently, populating main suggestion index with pseudonym info (there were only ~100 there before), and after checking status of that, will run the index approach script. Once the main suggest index is setup, Tim can try out the front-end code.
  - Stemming/phonetic matching: Checked with Frances and these are not present in our catalog search
  - John added packaging code here: <https://github.com/jskinne3/nectarguide>. ACTION - Should move the code into the LD4P github organization
  - Tim had to update the front-end controller code b/c some of the content of the fields changed and the result set for a query has changed (i.e. what matches/doesn't). Just restarted work on rspec tests for expected results
  - By next week, hope to have usability work scheduled. Will forward tasks draft to group and to Kevin/Usability Working group today.
  - Next week: ACTION: Huda to make slides and such to go through indexing approach, compared with client-side only, lessons learned, etc.
  - Next steps: Usability, Updating suggest index with index-only approach scripts, Finishing up RSpec tests (depend on suggest index updates being completed), Packaging.
- Authority Lookups for Sinopia (Lookup infrastructure: [https://github.com/LD4P/qa\\_server/projects/2](https://github.com/LD4P/qa_server/projects/2), Authority requests: [https://github.com/LD4P/qa\\_server/projects/1](https://github.com/LD4P/qa_server/projects/1))
  - QA performance
    - 2020-05-22 Dave is working through LoC authorities with new caching scheme, expect to hear about now
  - QA accuracy
    - 2020-05-29
      - Pretest results (see [image in QA/Sinopia weekly notes](#)). Summary 60 tests run with 26 failing. Of the 26, 4 were close to passing with the position off by 1 or 2. The other 22, the expected subject URI wasn't in the results at all. Geography data is especially bad at providing unexpected results regardless of whether it is coming from the cache or direct from the authority. 10 or the 22 failing tests were searching for a place.
      - Dave has updated the indexing for all LOC, AGROVOC, NALT, Getty, dbpedia, MeSH, RDA, CERL, and Ligatus. These are ready for a post test run to see what impact it has on the test results.
    - 2020-06-05
      - Organized issues related to accuracy groups into a [project board](#). Separated issues tagged as Cache Indexing into those that are related to new authorities and those that have examples of queries that are requesting accuracy tests or describe queries that did not return results as expected.
      - Used [Issue #201](#) to create new accuracy tests specifically for queries that have an expected result that has diacritics.
      - Ran accuracy tests for LOC\_NAMES\_RWO:person. [Comparison of Results for current production indexing and proposed indexing scheme](#). Clearly shows that the proposed indexing scheme is not an improvement overall
  - Authorities
- Linked Data API Working Group
  - Starting first Monday in June and then every other week for 4 months, logistics put into place
  - 2020-05-29 - Added initial working group documents...
    - [Current Approaches to Providing Search APIs](#)
    - [Needs of Consumers and Applications](#)
  - 2020-06-05
    - First meeting is next Monday
- LD4P3 Planning
  - Stanford functional requirements document: [https://docs.google.com/document/d/18H6zYGwKuG3SZqm9Q\\_cxkZThcdmBjknE6HdtQ-RRzk/edit#heading=h.4fu64x8jzm6e](https://docs.google.com/document/d/18H6zYGwKuG3SZqm9Q_cxkZThcdmBjknE6HdtQ-RRzk/edit#heading=h.4fu64x8jzm6e) —> ACTION - Discuss when Jason and Steven present
    - Possible relationship to entity management work in FOLIO
    - How much full LD vs linky-MARC?
    - CUL LTS working group is hoping to add URIs to MARC before FOLIO migration as a substitute for some of the heading management we do in Voyager in other ways. Perhaps won't be done ahead of switch but as part of the data migration. Considering how to update/maintain going forward

- 2020-06-05 Agree that doing some analysis/write-up for the Cornell context would be good, ACTION for July
  - Greg has been working on setting up a copy of Cornell Blacklight
- Meetings (see [LD4P2 Cornell Meeting Attendances](#))
  - [LD4 Conference 2020](#)
    - Expecting presenters to be contacted soon
  - [SWIB20 online](#)
    - Deadline for proposals 13 July 2020
  - rdfs:seeAlso [Conferences Related to Linked Data in Libraries](#)
- Next meetings
  - ...