

# 2020-06-22 Cornell LD4P2 Meeting notes

Date: 22 Jun 2020 – delayed from 2020-06-19

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

Regrets:

## Agenda & Notes

### Review actions from [2020-06-12 Cornell LD4P2 Meeting notes](#)

- [Steven Folsom](#) - To work on prioritizing pending authority requests pending in the QA Issues
  - 2020-06-05 Still to do. Steven to review prioritization, then check in with Dave if there should be change in his work order
  - 2020-06-12 Still to do
  - 2020-06-22 Conclude that everyone is OK with what is happening, DONE
- [John Skiles Skinner](#) - To move the nectarguide code into the LD4P github organization
  - 2020-06-12 Still to do
  - 2020-06-22 DONE
- [Huda Khan](#) - Schedule usability work. Will forward tasks draft to group and to Kevin/Usability Working group today
  - 2020-06-12 Met with Kevin and usability group on Monday. Agreement to do think-aloud and have reps interested in participating, might also get folks from access services. Plan to get tests done by June 25 giving time to create a report
  - 2020-06-22 Three usability tests scheduled Monday + Tuesday
- [Huda Khan](#) - To make slides and such to go through indexing approach, compared with client-side only, lessons learned, etc.
  - 2020-06-12 slides, but not these slides yet
  - 2020-06-22 Started slides [here](#)
    - What about cases with many variants? All are added to the bucket field but only those matching the query are shown as `aka`
    - Why have a combined bucket field rather than searching multiple fields? Provides simple way to search although current implementation searches label and bucket fields to enable matches on the label to be boosted. Also the bucket field has special matching criteria such as start of word, which isn't the case for all fields we might be interested in

### Status updates and planning

- Enhanced Discovery - **WHAM!** (see also <https://wiki.duraspace.org/x/sJl7Bg> 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. Slides with tasks and examples
  - Huda worked in incorporating knowledge panels for people/subject/etc. – demo
    - Note descriptions from wikidata are added to authors for example to help disambiguate. What is the description is very long? Current UI would expand width and then wrap as necessary? They seem mostly to be fairly short but we might want to force truncation for something that is long
  - Tim worked on rspec tests, now completed and pushed into dev
  - John worked on packaging, is moving Tim's code into nectarguide. Have decide not to adopt to use twitter typeahead but instead a more direct move
- 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
    - 2020-06-12 Taking a back seat to accuracy work
  - QA accuracy
    - 2020-06-12
      - [E. Lynette Rayle](#) Worked with Dave on refining the indexing for LCNAF. [Comparison of Results for current production indexing and proposed indexing scheme](#) updated to include intermediate and latest results. At this point, there are 3 out of 24 failing tests for LCNAF. 1) same failure as original due to other legitimate matches ranking higher, 2) special character that is considered an alpha is missing, so it doesn't match, 3) a match moved down in the ranking past the expected position (not sure why)
      - [Steven Folsom](#) I've worked on some tests and moved some issues to completed here [https://github.com/LD4P/qa\\_server/projects/3](https://github.com/LD4P/qa_server/projects/3), where I thought the only hangup was tests. Those issues that remain in the Failing Queries Columns, I've added comments to.
  - Lynette has mostly been focusing on exhibits. Dave is working on performance improvements to eliminate SPARQL quest after matches for LOC. Still have Steven's additional tests, want to see how capitalization tests work and should have same results – hope to get these set up this week
- 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
  - 2020-06-12
    - First meeting was last Monday ([meeting notes](#)). Brainstorming fell into several categories of topics: Caching (and management of cached data), Reconciliation, Accuracy (of results and of selection) (including extended context and ordering of results), API Approach, Data Structure, Scalability.

- Creating a [list of terminology](#) with definitions. Discussions happening in Slack about terminology definitions including a question about the definitions for Reconciliation and Entity Resolution. I used definitions that were used in the first LD4L workshop which defined Reconciliation connecting Things-to-Things and Entity Resolution as connecting Strings-to-Things, but it was suggested that these are actually the same thing.
    - Second meeting today - focus on use cases, so far Lynette has added one
- 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)
    - 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
    - First pass at a draft for the workshop for the LD4 Conference here: [https://github.com/sfolsom/LD4\\_2020\\_Conference\\_Linked\\_Data\\_WorkShop](https://github.com/sfolsom/LD4_2020_Conference_Linked_Data_WorkShop). Comments welcome, but no pressure. I'm still working with the Conference Organizers to firm up workshop slack, email, and office hour communication.
    - Everyone encourages to respond promptly to conference organizers
  - [SWIB20 online](#)
    - Deadline for proposals 13 July 2020
    - Autosuggest / Discovery / Lookup / ...
    - SWIB brainstorming [document](#) for team
  - [rdfs:seeAlso](#) [Conferences Related to Linked Data in Libraries](#)
- Next meetings
  - Organize wrap-up on 30 June