

# 2020-06-26 Cornell LD4P2 Meeting notes

Date: 26 Jun 2020

Attendees: Huda, Lynette, John, Tim, Steven

Regrets:

## Agenda & Notes

### Review actions from

- [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
  - 2020-06-26 Five usability tests completed, usability report forthcoming
  - Steven: Report outs in metadata group included discussion of usability tests. People who use online search a lot (search engines etc.) seemed to understand auto-suggest functionality quite well.
- [Huda Khan](#) - To make slides and such to go through indexing approach, compared with client-side only, lessons learned, etc.
  - 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
  - 2020-06-26 Slides to be completed early next week as part of documentation for project

### 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 - [Slides with tasks and examples](#)
  - Huda: Wrap up work. Working on some remaining bugs, hoping to address by Tuesday (and if not, will document what remains to be done). [Lessons learned](#) draft started. Also brief indexing script [documentation](#) (for how to run the script). WHAM! demo video segments to be done by: indexing (Huda - Friday), controller front-end (Tim - Monday), packaging (John - Monday).
  - 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))
  - Last week: 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
  - This week's update:
    - Accuracy testing: [2020-06-24 Accuracy Testing Focused on Capitalization](#) Observations: Stemming is not happening. Exact matches aren't found when the preferred label is language tagged. Dave is looking into turning on stemming and exploring how to handle labels with language tags.
    - Performance: Dave is reindexing LOC to use the new indexing that keeps the extended context in the index.
    - SF - cleaned up the Priorities spreadsheet ([https://docs.google.com/spreadsheets/d/1UhesnSRVHJNt8zFVjh6VKg4hTOzMpSj1AXX\\_jRvmY6l/edit?usp=sharing](https://docs.google.com/spreadsheets/d/1UhesnSRVHJNt8zFVjh6VKg4hTOzMpSj1AXX_jRvmY6l/edit?usp=sharing)), moved closed issues to the complete section at the bottom. It confirms we are currently working on the right issues.
- Linked Data API Working Group
  - Last week: Starting first Monday in June and then every other week for 4 months, logistics put into place, documents:
    - [Current Approaches to Providing Search APIs](#)
    - [Needs of Consumers and Applications](#)
    - [Meeting notes](#)
    - [List of terminology](#) with definitions
  - This week: Had second meeting and discussed current approaches and defined potential use cases.
    - [Current Approaches to Providing Search APIs](#)
    - [Use Cases](#)
    - [Meeting notes](#) (recording: [youtube](#))
  - First use cases added around looking up authorities: keyword or left-anchored search. Filtering also mentioned as a desired feature. Other use cases included caching, batch processing both automatically and supervised (e.g. reconciliation).
  - Issue raised that perhaps we should first define a model that would be supported by shared APIs even if this data stored locally in a different data model.
    - E.g. If requesting information about a person entity, data is returned using the same Person ontology regardless of data source /authority.
  - Discussions will continue around use cases and how the working group will address them.
- LD4P3 Planning
  - 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)
    - Possible relationship to entity management work in FOLIO
    - How much full LD vs linky-MARC?

