

2020-10-09 Cornell LD4P3 Meeting notes

Date: 09 Oct 2020

Attendees: Tim, Lynette, Steven, Huda, Simeon

Regrets: Jason, Greg

Actions from 2020-10-02 Cornell LD4P3 Meeting notes

- [Jason Kovari](#) and [Steven Folsom](#) to explore where/how/if a discogs id might be added to MARC that would then be used in place of dynamic search to find a matching record (and coordinate with Tracey re. best practice). Would be used to override dynamic search
 - 2020-09-25 - Tracey/Laura/Jason have been discussing and have proposed location (024 7_ \$2 discogs). Steven has been doing some work on catalog numbers (aka publisher id) in music cataloging
 - Steven will create wiki page for this project and document cataloging decision [Using Discogs to Enhance Discovery](#).
 - There are currently no \$2 discogs in the catalog so Frances is coding for the future, they will be added. DONE
- [Steven Folsom](#) & [Tim Worrall](#) to follow up on this discussion with Frances about how we surface the id type in the Solr index in a way that will be easy/efficient to use
 - Discussion was had, and Frances created this issue: <https://culibrary.atlassian.net/browse/DISCOVERYACCESS-6490>.
- [Simeon Warner](#) to find out who we might talk to about librarian instruction at CUL re. possible use cases
 - 2020-10-09 Have asked Dorianne
- [Simeon Warner](#) to follow up about getting redirect from Id4l.org to wiki gateway page and to decommission drupal site
 - 2020-09-25 In progress, <https://culibrary.atlassian.net/browse/DLITSYS-3983>, redirect in place but still need to decommission Drupal site
 - 2020-10-02 Greg has asked Jim to decommission the Drupal site
 - 2020-10-09 All DONE, Drupal site gone
- [Huda Khan](#) and [Tim Worrall](#) work to continue to explore the dashboard design and requirements
 - 2020-09-25 in progress, will discuss with Astrid & Tim next week
 - 2020-10-02: Kickoff design meeting [happened](#) [Huda Khan Tim Worrall](#) & astrid will look at different aspects more and regroup in 2 weeks (Meeting set for 10/14)
- [Greg Delisle](#) to explore how to get the container version that Lynette has on her laptop up on AWS. Question about how to handle environment variables outside of the checked in docker container so that users can config
 - 2020-09-25 Greg is working on that (right now)
 - 2020-10-02 Greg can get it running on ECS but we can't see it directly of from behind load balancer, still working through the details
 - 2020-10-09 On laptop Lynette had an entypoint script which worked fine to allocate assets and connect to mysql etc., this doesn't work on AWS. Lynette is trying to move these processes into the Docker file. Works for bundling but not for precompile, has some input from Justin, Greg will look into the DB connection. Some concern that we are having to make many changes for AWS, will the new version work on laptop or other VM? (changes so far do work on laptop too)
- Toward a demo and discussion with D&A:
 - [Huda Khan](#) & [Tim Worrall](#) - Make sure the live demo meets the Discogs API terms of service in terms of linking and ack. – Had discussion with Melissa about how/where we show acks and licensing in general for our Blacklight. Could then link to this from all pages and perhaps a specific link to the details on this page. A bit more to do
 - 2020-10-09 - Code changes DONE, Tim thinking about whether we should remove not about wikidata which isn't part of the current production plan. Example: ld4p3-web.library.cornell.edu/catalog/6055667
 - [Tim Worrall](#) Create mockup pages for the cases of 1) a completely blended presentation, and 2) a presentation where the Discogs data is separated out
 - 2020-10-09 - DONE, went straightforwardly
 - [Huda Khan](#) to discuss with Adam having a meeting with User Reps sometime after Oct 5
 - October 16th potential date/time: 3:30pm need to send list of names to add to this meeting to Meredith - names added
 - [Tim Worrall](#) to look into how we might display some language pointing out what data is not used in search. For example, what if a user sees the record label in Discogs data and then tries searching on that (which they can't)? Per Tracey suggestion
 - Partly contingent upon the presentation approach selected. Think about it more after 2020-10-16 meeting

Agenda

Discovery (WP3)

- <https://github.com/LD4P/discovery/projects/2> for issues etc.
- Preliminary draft of a discovery plan- intended to get feedback: https://docs.google.com/document/d/1zKYW7FQVVNvyd0XijW0qWznX9PC3jbmOE6Kz_yyGpjs/edit?usp=sharing
- Strand 1: production piece (i.e.: picking features we've already worked on to then push those into production): discogs, autosuggest. Feature branches in the code to work on features in isolation. Still need to discuss discogs with Tracey (post- 9/2). Preliminary pass with discogs features using latest production (from a few days ago) successful (tried on dev vm and will try on ld4p3 demo using continuous integration later)
 - Steps:
 - Discussion with Tracey re. use case and benefit – DONE
 - LTS engagement re. metadata - DONE
 - Production requirements and functionality – [Production decision points](#)
 - Demo and discussion with D&A User Reps and dev team – See actions above from last week
 - Have branch with discogs code on dev VM and demo server from LD4P fork, but still issues with the assets folder under Rails not showing current version (these are generally copied in during the Rails deploy process)
 - Update 10/9/2020: Manually cleared out the assets folder on the two rails servers and ran Jenkins deploy more than once and saw only one copy of the files, so hopefully this is fine now. Will keep an eye on the assets folders.
 - How do we run the demo meeting next week? Simeon intro mention demo and mockups of alternatives, turn over to Steven to mention catalog override and then demo from Tim

- Strand 2: research: how to go from knowledge graph to an index - what decisions are needed. What are the data sources for each (e.g.: how many Cornell faculty in wikidata)? Present: reviewing data sources and questions. Should have more worked out in a week or two. Main areas of concern: browsing and dashboard... and anything we can do to help patrons navigate our collections and how we can highlight an entity... and what does that mean for the index? How do we capture the relevant bits of the graph for an index? Is there a repo for this? Not yet... but can use discovery repo (<https://github.com/ld4p/discovery>) we already have to capture any queries or related work.
 - **Research decision points**
 - **Use cases** - first review 2020-09-18
 - Interest in the dashboard idea (full page for entity) that extends on the idea of an embedded knowledge panel
 - Goal of functional prototype for end of year
 - How will we test it? Remote testing would work but will students want/have-time to talk with us? Can we pay them?
 - Discovery Affinity Group is looking at design ideas
 - 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
- Linked-Data Authority Support (WP2) - A key element of this work package is a sustainable solution that others can deploy. Questions of budget for deployment. Need to get all code into LD4P repository. What would a good end-product look like both for our maintenance and for others to use
 - **Qa Sinopia Collaboration**
 - 2020-10-09
 - Had a meeting with ShareVDE this week. I provided them with the [API doc](#) I put together last year. We discussed issues related to [Change Management](#) more and I offered to have discussions with them and Dave about how best to do this. In related work the Sinopia team provided sample records which SVDE will do a trial ingest with to understand any data changes from round trip. Sinopia team now off work cycle until perhaps Jan 2021. Next meeting with ShareVDE is Tues, Oct 20.
 - Added LC small vocabs to cache and through QA ([example: carriers q=audio](#)). They have requested the ability to get all terms for vocabs that are super short. My plan is to add a config for `{terms: allowed=true | false}`. Most linked data authorities should not have this because there are hundreds of thousands of term, so it will default to false. This requires changes minimally at the QA level and possibly at the qa_server level as well. QA does support fetch all terms for some non-linked data authorities already, so the infrastructure to handle this is already there.
 - Dave has now released most authorities under new index scheme. Lynette is doing a before/after analysis with accuracy tests
 - **Cache Containerization Plan**
 - 2020-10-09 -Update under action items.
 - **Search API Best Practices for Authoritative Data working group**
 - 2020-10-09
 - Working through developer user stories. Next meeting is next Monday.
 - ACTION: Lynette needs to put together survey for cataloger user stories for PCC feedback and ranking by working group.
- 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
 - What does success look like? And then how do we get there?
 - 2020-09-25 Jason drafted an introduction for LTS staff and no capacity right now in LTS, perhaps discuss again next week as to what next steps will be. Jason and Steven will discuss this week
 - 2020-10-02 Jason will collect information to prime brainstorming from LTS, think about brainstorming then discussions
- Other Topics
 - OCLC Linked Data / Entities Advisory Group
 - OCLC has described minimal viable entities and a number of small vocabs – are these new "standards" that we will have to engage with to ensure alignment?
 - PCC Sinopia Profiles Working group
 - 2020-09-18 - Still going through large spreadsheet comparing Sinopia profiles with BSR and CSR, close to having a comparison, will be part of report to POCO, still on target to complete in October
 - 2020-10-02 - Submitted interim report yesterday. Feedback requested before deciding on a few directions and working to stay compatible with LC work. Some discussion of how to align with SHARE-VDE data which uses a rather different profile. Committee is supposed to wrap by the end of October
 - 2020-10-09 Almost done, questions about role of item template which doesn't fall within PCC scope
 - PCC Task Group on Non-RDA Entities
 - 2020-09-11 - Preliminary report went to POCO suggesting a small vocabulary of entity types, discussion about how to host /implement
 - 2020-09-25 Now have a short list of possibilities and close to decision
 - 2020-10-02 Have definitions of short lists, now creating MARC examples. Interest is working the German national library on alignment – perhaps influence that and then adopt instead of creating a new list
 - 2020-10-09 Discussions of how reconcile RDA list with Non-RDA entities. Possible effects on id.loc.gov
- Default branch name - WAIT until we can use github tools January 2021
- Upcoming meetings
 - Virtual Blacklight Summit in early October. (Main site [link](#) and [document](#) to add lightning talk/breakout session ideas) - Huda is on unofficial committee. Expecting various follow-ons from earlier meeting this year. (Discussions still in progress and, at some point, draft program will be shared with the community).
 - [Agenda](#) and [registration link](#) have gone out. Will be intro workshop, time for lightning talks and longer sessions. Sessions for plugin discussion including call-number browse
 - Pre conf workshop Tuesday on blacklight customization (4h on zoom!), continuing with 2h/day
 - Huda gave talk on Blacklight and linked data
 - SWIB
 - Discovery submission accepted
 - Programme still in progress (not up on website yet: <https://swib.org/swib20/programme.html>)

- <https://kula.journals.publicknowledgeproject.org/index.php/kula/announcement/view/1> . Call for Proposals - Special Issue: "The Metadata Issue: Metadata as Knowledge". Due January 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"

Next Meeting(s), anyone out?:

- 2020-10-16 - Steven out