

2018-08-24 Cornell LD4P2 Meeting notes

Date: 24 Aug 2018

Attendees: Tim, Huda, Jason, Lynette, Steven

Regrets: Simeon, Dean

Zoom: <https://cornell.zoom.us/j/448375153>

Agenda & Notes

Review actions from 2018-08-24 Cornell LD4P2 Meeting notes

- [Huda Khan](#) to follow up on Vitro/VitroLib security issue that Randy asked about
 - 2018-08-17 Test was done, Huda has yet to complete analysis. Will post results as information for the VIVO community but not do any work to address.
 - 2018-08-24 Not yet. Have to read report to see if issues are detailed. Will report back to Harvard and VIVO community. Will attempt to do by 8.31
- [Simeon Warner](#) to work with Mary Beth to set up an LD4/D&A discussion. Should include Melissa & Frances. Frances has done something in an innovation week. Goal is to understand areas of work that might happen in LD4 that would align with existing hopes
 - 2018-08-02 Have sent email to MB, Simeon will follow up again
 - 2018-08-24 no status
 - [Steven Folsom](#) : this is his largest work area; can help coordinate if needed
- [Simeon Warner](#) to make page in LD4P2 space for Cornell team notes
 - Done, this page is in [Cornell LD4P2 Team Meetings](#)
- COHORT????
 - These are to be decided in early September, we think; interested to know more about applicants.
- Issues: <https://waffle.io/LD4P/ld4p2-cornell>
 - issue 8 marked done as integration happened; however, need new issue for adding contextual look-up into BFE for GFT. Action item : [Huda Khan](#)
 - issue 11: created on <https://wiki.duraspace.org/x/ch0RBg> ; currently posted under the Cornell Team meeting page; should be moved to: <https://wiki.duraspace.org/x/gBkRBg> . Action item: [Huda Khan](#)

Status updates and planning

- [bibliotek-o.org/](#) is not secure; should we correct?
 - Q for [Simeon Warner](#) : should we prioritize this? if so, who should work on it?
- Authority Lookups for Sandbox (Production QA instance: https://github.com/cul-it/qa_server/issues/)
 - Priorities QA: keeping historical data, logging of use, then finish refactoring, then performance
 - 2018-08-24: nothing to add; [E. Lynette Rayle](#) on ESMIS this week
- Enhanced Discovery
 - How do we move forward with implementation of schema.org? <https://waffle.io/LD4P/ld4p2-cornell/cards/5b6c955eaabc3e001bb0b173>
 - Need discussion with Stanford and to have some brainstorming – if we add extra connections in out data (e.g. a wikipedia link on a catalog record, how will that show up?)
 - When is this being set up? Who is responsible for planning?
 - Is [Simeon Warner](#) setting up this call?
 - Tom mentioned the same issue that we are having; Stanford Blacklight devs are not the same as LD4P devs... similar to Cornell. Both institutions need local discussions
 - During monthly all-hands call, Tom mentioned a breakdown of work, which implied distribution between Cornell and Stanford; can we confirm? We heard him say they are working on schema and knowledge panel... and that Cornell should work on semantic browsing.
 - Can [Simeon Warner](#) confirm work plans?
 - What is the Iowa component of this work area?
 - How much of this will be put into production in local discovery environment?
 - this will be determined based on discussion with D&A;
 - QA can already facilitate some of these enhanced discovery components (e.g.: pulling into data from DBPedia, etc.)
 - We should begin from a position that we WILL implement in production.
 - Question of source data for production discovery environment... will we be adding URIs to MARC data for this specifically?
 - Question of what is contextual data here... and how to facilitate that more broadly than individual entities in DBPedia linked from MARC records / RDF, etc...
 - Question of what is added to the source data? what is indexed in SOLR? what is added on-the-fly?
 - What resource implications need to be identified on the 'general' D&A side to ensure that work in LD4P2 is leveraged?
 - What is the tangible discovery output? Is there a centralized discovery environment of _____ data from the grant? Is it a GH repo with some code? What is promised?
- Prep for Cataloging Sinatra and other 45's (Discogs data)
 - Application Profiles
 - In mid-to-late September, Steven, Tracey and Jason will start planning for Sinatra BFE profile. Hip Hop SHACL was for 78s... Sinatra will be 45s... HipHop used bibliotek-o.
 - Question: whether we want to help out with Art & Rare Materials BF Ontology Extension profile to ensure that others are facilitated...?
 - Discogs data / API
 - Tim pulled data from Discogs
 - Steven assessed data. Some can easily convert some into local resources while others need reconciliation. Some data will be conversion of literals to resources

- Issue: importance of distinguishing between Release and Master in Discogs
- Want API to feed form that can be supervised into conversion... cataloger has API that searches Discogs for a 45. Once selected, data are then put into a form but no RDF yet... since we might then duplicate resources.
- Can we identify resources for the intermediate hand/auto reconciliation steps post-JSON & pre-RDF.
 - OCLC' LD prototype has a tool that searches VIAF, ISNI and one other. Select one and it brings into editing environment and you then claim bits of the conversion and hit save. Diff = only using RDF sources... so creating redundant RDF in different environment rather than non-RDF to RDF
- Building this out would be really beneficial for community...
- Could be integrated into BFE OR a separate form/app
- How do you go from pre-populated field to a QA look-up? What do we show them? Do we do a pre-look-up with a suggested list?
- Piece that may not be in BFE – you have a way to create or edit a record... but nothing where you go out and search for things you can pull in. This includes a lack of connection to OCLC or other datasets. Astrid and Christina are discussing / looking into this. No search box to look up things that already exist... go look them up and pull them in. This could all be the same functionality (OCLC OR Discogs)... where you look up & pull in... but nothing saves or writes RDF until you hit save.
- How does QA fit into this? There are two components ... how to retrieve the data and how to process the data.
 - would need to add a mapping module into QA as a post-process.
 - from the processing side of QA, inside that does a term fetch; get back JSON ... with mapping module that then translates data from JSON into another format
 - QA the server... not QA the Gem supported by Samvera community
- Action item: discuss during 8/31 mtg whether this could/should be prioritized for work
- Action item: watch the space of importing external Works, etc., data into BFE – to see if this links in
- Action item: [Steven Folsom](#) will create a repo for the Sinatra 45s work item... and add the JSON file there; we'll also post app profiles, etc there. IF we prioritize the QA extension work, we'll create a repo/folder for that work and this can go into there...
- Travel and meetings – what meetings should be be present at between now and end June 2019?
 - ISWC (Steven), DCMI (Jason, not on LD4P2 funds), SWIB (Steven, Huda, Lynette, Simeon), CNI (Dean, not on LD4P2 funds), US2TS (March 11-12, Duke), ALA (Harvard doing PCC liaison)
 - Wiki Conference October 18-21 https://wikiconference.org/wiki/2018/Main_Page – not sure whether anyone is available
 - Will there be another European BF summit?
- LD4P2 meeting LC - 15-16 Oct in DC
 - Jason, Steven, Lynette, Huda, Tim, Dean, Simeon is 7 people, budget is \$1500 each – all go ahead provided within budget, pre-travel to dept head and Simeon

Next meetings:

- 8/31