**2020-02-28 Cornell LD4P2 Meeting notes**

Date: 28 Feb 2020  
Attendees: Tim, Huda, Jason, John, Simeon  
Regrets: Lynette, Steven

### Agenda & Notes

#### Review actions from **2020-02-14 Cornell LD4P2 Meeting notes**

- **Huda Khan** to discuss with Astrid and David possible collaboration with U Chicago over usability (and maybe others in DOG team)
- **2020-02-28 Chicago plan is to look at author browse and they will engage with DAG on this**
- **Lynette Raylis QA performance**
  - 2020-02-14 Dave has made progress this week, is moving to have all data in index tailored to search in order to avoid SPARQL queries at search time, results are in a blob of RDF. Working on CERL first, expect to get this our soon. Will then try MeSH and OCLC FAST, then LC.
  - 2020-02-21 CERL was deployed with the new index strategy but no before and after to compare. However, this is small so we need to wait for LC or such to get a sense of possible improvement
  - 2020-02-25 New authorities have been brought online all the way through to Sinopia. These include CERL (searching person, corporate, imprint, or all of these) and Ligatus. Additionally, MeSH has been updated to include extended context and support for searching by subject or publication type.
- **Adam Smith** to investigate cost and any issues with setting up a D&A Beta system to allow broader testing of some discovery ideas from this work
- **Simeon Warner** to ask him
- **John Skiles Skinner** to continue discussion with Hathi trust about an API or access to their index
  - 2020-01-31 There is investigation but not sure whether it will result in something we can use
  - 2020-02-14 Some more discussions with Hathi and suggestions was to use current search with debug facility that includes things like facet values in machine readable form (requires either 1) a user account for testing, or 2) to use IP access for our dev machine but there is some issue of fixed external IP for our dev VMs)
  - 2020-02-21 HathiTrust sent over the XML version for one of the queries that John had tried for zero results. This would be the same xml they may be able to open up for us by allowing the IP address of my dev vm to access the URL that would result in XML. Huda set up a controller that parses the xml and returns json with the list of subject heading strings and set of search results being returned (this is the xml version of the search results page that includes subject facet values). John said he should be able to incorporate the subjects and perhaps the results into the zero search results page.
  - 2020-02-22 HathiTrust have allowed institutional accounts to add a query parameter to get XML output, may also provide IP based access for prototypes. Have already made demo with a mock-up of access
- **Huda Khan Tim Worrall John Skiles Skinner** to finish up **lessons learned** from BAM!
- **2020-2021 progress on the BAM lessons learned doc and will aim for finishing that today...**
- **2020-02-28 · Huda Khan** to copy lessons learned into the main wiki and check scripts into github
- **Huda Khan** to collect everyone involved next week week to plan the discovery session for 3/3, review plan on 2/28
- **Meeting was had. Notes. Slides**
- **Simeon Warner** to ask about timing of the April 21 partner meeting, wonder whether 3:30pm end would be possible?
  - Yes, meeting will run 9am–3:30pm, minivan will depart ASAP after from somewhere near LC

#### Status updates and planning

- Discovery presentation on 3/3: what is agenda? who is speaking? who is announcing to CUL?
  - **Goal is to disseminate our work and then get some feedback about what might be promising**
  - **Simeon to timekeep**
  - **We need to have a way to take good notes – designate a note taker**
  - **View from staff involved with virtual libraries would be interesting, also D&A user reps**
  - **Need to be careful about managing expectations for stage this work is at and what might happen or not going forward**
- **Cataloging Sinatra and other 45’s (Discogs data, [https://github.com/ld4p/qa_server/issues?q=is%3Aissue+is%3Aopen+label%3ADiscogs](https://github.com/ld4p/qa_server/issues?q=is%3Aissue+is%3Aopen+label%3ADiscogs))**
  - **Lookups for place not usable and hence places are not being recorded, relies on work from Dave to fix:** [https://github.com/ld4p/qa_server/issues/248](https://github.com/ld4p/qa_server/issues/248) & [https://github.com/ld4p/qa_server/issues/240](https://github.com/ld4p/qa_server/issues/240)
  - **Have a currently insurmountable issue with nested profiles. When create Work profile with nested Instance profile there isn’t a URI for the Instance (it just gets hung from a bnode). Without a URI the title of the Instance doesn’t get indexed.** The Sinopia team are unable to fix this in the near term.
  - **Cataloging work continues with the above limitations**
  - **2020-02-28 Steven update – I did a bunch of PCC profile and LOC policy related writing/correspondence; met with Huda, Tim, and John to discuss the Discovery Event (happy to help facilitate/notetake/rove on the day of the event); worked with Sinopia team to understand title search and display bugs that have been affecting Sinatra work (Jeremy has created [https://github.com/ld4p/sinopia_editor/issues/2090](https://github.com/ld4p/sinopia_editor/issues/2090) which looks at part of the problem); I still need to clean up the QA/Sinopia priority list to reflect the work completed by Lynnette and Dave.**
- **Enhanced Discoverys (see also [https://wiki.duraspace.org/x/sJI7Bg](https://wiki.duraspace.org/x/sJI7Bg) and [https://github.com/ld4p/discovery/projects/1](https://github.com/ld4p/discovery/projects/1))**
  - **SMASH! (dev to run through 7 Feb, then user testing, video and write-up) – dev complete, video done, Hitchcock homage and cameos still under consideration,** Lessons learned document in process and also set up document for annif use summary
  - **Open meeting March 3, 2-3:30pm in Mann 102 and should Zoom it too**
  - **Will continue on Hathi work...**
  - **How will we decide what to take forward from KAPOW!, BAM! and SMASH!? (or as Tim put it, “what happens in late February?”)**
    - Do discovery session... get feedback
    - Knowledge panels – could we make a component that is easily reusable in any Blacklight? How much are local customizations key?
    - Semantic stuff .. annif ... relationships in data to get relevant semantic links and use of hierarchy in data
    - Call number browse and other virtual browse notions, with semantics/facets?
• Use of linked-data descriptions from Sinopia - what can we do in discovery that is different?
  • Authority Lookups for Sinopia (Lookup infrastructure: https://github.com/LD4P/qa_server/projects/2, Authority requests: https://github.com/LD4P/qa_server/projects/1)
    • When deployment issue solved... will then put out CERL and ligatus and extended context for MeSH along with sub-authorities. Also some refactoring associated with monitoring status page (including fixing a memory leak due to a long-used hash - ruby doesn't reclaim space from deleted entries)
    • Not sure whether Dave has redone the index to avoid SPARQL for MeSH – if it is done then we will have a comparison
    • ACTION - Lynette to ask Tiziana about SHARE-VDE APIs for real-time up-to-date search and for possible engagement in linked data best practices for authoritative data working group
  • Travel and meetings (see LD4P2 Cornell Meeting Attendances)
    • Knowledge Graph Conference (Columbia University), May 6-7, workshops 5/4-5/5
      • Workshops due 1/15 (not doing this), proposals due 2/28, talks 20mins
      • Preliminary brainstorming
      • Huda Khan will lead effort to consider proposal
    • LD4P2 cohort and partner meeting, LC, 20 & 21 April
      • Minivan reserved - leaving Ithaca Sun 19, returning after meeting Tue 21 April
      • Partner meeting on 21 April - https://docs.google.com/document/d/17OS9rsF__N8n9hiGDYuxggdSRS0EXSPBiw2XgZTb4jw/edit#heading=h.jiyum6rb9zes
    • LD4 Conference 2020 at College Station, TX (TAMU) - May 13/14, 2020
      • Proposals went in 2020-01-31 – Discovery, QA/Sinopia, BOF for best practices, Intro to LD workshop
      • Expect hear back soon, registration will open then (first come first served)
  • rdfs:seeAlso Conferences Related to Linked Data in Libraries
• Next meetings:
  • ...