Applied Linked Data 2016-11-15

Time: 10:00 AM PST / 1:00 PM EST
Google Hangout: https://plus.google.com/hangouts/_/gv3gejk2am4xggnpntydvh3aa
Backup Phone Line If Hangout Doesn't Work: 712-775-7035 (Access Code: 960009)

Etherpad Group Notes: https://etherpad.wikimedia.org/p/Hydra-LDP-20161115

Attendees:
- Steven Anderson
- Lynette Rayle
- Anna Headley
- Corey Harper
- Zach Schoenberger
- Eric O’Hanlon

Agenda:
1. Brief Introductions and previous group background
2. Cache refreshing for ActiveTriples
   a. Initial Exploration and Community Feedback at... https://github.com/ActiveTriples/linked-data-fragments/issues/31
   b. Presentation on Caching Refresh discussion... https://docs.google.com/presentation/d/1k4fDu9906X24yaJxG5FGniBbNOD6phbP9hgPEHyLXMk/edit?usp=sharing
      i. What to do if configuration specifies expiration_period should be based on URI header and the URI header does not contain expiration info?
         1. Could use configured default if it is missing. Could treat it as 0 and immediately refresh. Needs further exploration. See challenges in the slides.
         2. Can expiration period be configured for a sub-authority?
            1. Initial intention was host of the URI for a term. e.g. aims.fao.org, id.worldcat.org, id.loc.gov
            2. May be desirable for there to be a different expiration_period for each sub-authority. e.g. id.loc.gov/authorities/names, id.loc.gov/authorities/subjects,
   iii. Does the cache refresh do a refresh of LDF cache AND application’s solr?
      1. Desirable. Side car was explored to update solr. Work hasn’t been done.
      2. Challenge: Update to a solr doc with OCR brings back the OCR text, which greatly increases the overhead of a solr document update. How to partially update the solr doc?
   iv. Can a report be generated listing refreshed terms for the periodic background cache refresh process?
   v. What repositories are in use now with LDF?
      1. Marmotta was tried but had resource issues as the number of resources increased.
      2. BlazeGraph is in use at BPL to hold LoC full download
      3. Exploration of Fuseki, but likely to look at BlazeGraph
   vi. When getting query results from LDF, is there a way to rank results?
      1. Not implemented by default. Ranking is not part of the LDF specification which is a pattern match on the $subject, $predicate, or $object values.
      2. Discussed using solr as a ranking index for queries, but there was concern that this would increase the complexity too much.
   vii. There was a very brief discussion of changes that are in progress for Questioning Authority to support access to linked data authorities. This included the potential to access LDF cache using this same procedure. This led us to talk about Metadata Enrichment UI work by Steven Anderson at BPL which can be used to add a lookup for Curation Concerns fields that provides more context for selecting a term (e.g. broader, narrower, same as)
      c. Related work in Questioning Authority to access linked data authorities... https://github.com/ld4l-labs/questioning_authority/tree/linked_data (README)
3. Metadata Enrichment UI
   a. Steven demoed UI for Metadata Enrichment which uses data in Linked Data Fragments-BlazeGraph implementation to populate a property field
   b. Hydra Connect presentation for Applied Linked Data workshop... https://docs.google.com/presentation/d/1B2awnzicV9-1CeSqJE5aRBie6LJC_xUWrADKG16Dh8U/edit#slide=id.g35f391192_00
4. Group structure? (Not Covered - ran out of time)
   a. Are we a Working Group or an Interest Group?
   b. Under Hydra Metadata Interest Group as a subgroup or a top level group?

Next Meeting: Nov 29 at 10pst/1est (and again Dec 13)