

Wikidata Breakout Sessions Meeting Notes

at June 25, 2019 cohort in-person meeting at the Library of Congress

Meeting agenda, location, etc., here on the LD4P2 wiki: [Cohort Meeting #2, June 25, 2019](#)

<https://stanford.zoom.us/j/578253206>

Session 2

Facilitator: Hilary Thorsen

Note-taker: Marc McGee and Hilary

Recording:

https://stanford.zoom.us/recording/share/tbwe1ZvXxTtIHsFuCTxkpdFgkxW1_lsmk6pPMqmdLOiwlumekTziMw?startTime=1561491191000

Intros

Jeremy - expressed interest to use Wikidata with Sinopia via api; possibility of using an api to use wikidata entities, how it fits with front end

Marc - Wikidata's interactions with sinopia and wikidata description of maps.

Matt - use case around wikidata identifiers and id.loc.gov

Tom - interests transitioning to new ILS FOLIO and exploring vocabularies outside of traditional library vocabs ; secondly digital collections not easily linked to traditional library sources

Interests

Geographic entities

Matt: NAF has a lot geographic headings with 670 with automatic match to Geonames with coordinates, but no identifier; curious if can take coordinates in Wikidata and match with geographic headings

Marc: a lot of point data; interested in printed maps which record bounding box, but property proposal for bounding box wasn't accepted; geoshapes in Wikidata, but Marc hasn't seen where you can pull good coordinates from Wikidata; Who's on First might have that data

Jeremy: was used in Blacklight knowledge panel work

Marc: point in Geonames for neighborhood with text might be difficult with granularity

Matt: significant amount that could be added to Wikidata in terms of geocustering

Marc: interested in mapping bibliographic aspect of maps in Wikidata; not a whole lot of maps described in Wikidata at this point, maybe 7,000--some British Library maps; is it useful to Wikidata community to have property ?

Matt: preliminary mapping to Wikidata from BF work; reconciliation of predicates to find out what we need in Wikidata--less than a dozen that align when mapping a monograph

Marc: Hilary and I talked about the best approach; put up property proposal for cartographic relief; suggesting class structure around predicate and some use cases; Statement of responsibility is a qualifier of subtitle

Wikidata in Sinopia

Marc: can instance for digital map, but wants to add Wikidata entities, so it would be easier to publish out a converted set to Wikidata

Matt: are there identifiers associated with maps

Marc: there isn't unless they have ISBNs or using published bibliography, which wouldn't be ideal; 50%

Matt: recently added BF instances and works to id.loc.gov, so interested in how many maps might be in it

Jeremy: beyond QA to lookup Wikidata use Sinopia as editor to catalog and push out

Marc: using Wikidata for hosting local authorities; would need to be able to get Q#; create on the fly in Wikidata and paste Q#

Hilary: use in cataloging workflow--increase production efficiency

Marc: set up profile that identifies entity type and reconcile entity to make sure it doesn't exist in Wikidata

Wikibase

Matt: Library flavor docker image for wikibase; significant hurdles for implementation; distribution with pre-loaded data (predicates defined);

Marc: BF predicates?

Matt: yes, some existing rdf/wikidata properties e.g. "instance of" + some BIBFRAME Properties

Hilary: Wikidata Affinity Group, concern has been brought up about resources to run some of the tools

Matt: if everyone is using same flavor of implementation; possible some of the data will persist through common mappings

Hilary: pushing data out from wikibase to wikidata?

Matt: no, but could map Properties and generate quickstatements

Marc: is it possible in wikibase--one of the nice things is that it's clean--in Wikibase can you have a set of properties predefined to populate based on type

Matt: automatically populates based on item type? Could do it via API

What would be helpful?

A list of properties associated with specific entity types

-Wikidata model items (https://www.wikidata.org/wiki/Wikidata:Model_items)-
could be discussed at future Affinity Group meeting

-e.g. Shakespeare for playwright

-Query for full list of model items:

https://query.wikidata.org/embed.html#SELECT%20%3Fsubject%20%3FsubjectLabel%20%3FmodellItem%20%3FmodellItemLabel%20WHERE%20%7B%0A%20%20%0A%20%20%3Fsubject%20wdt%3AP5869%20%3FmodellItem%20.%0A%0A%20%20SERVIC%20wikibase%3Alabel%20%7B%20bd%3AserviceParam%20wikibase%3Alanguage%20%22%5BAUTO_LANGUAGE%5D%2Cen%22.%20%7D%0A%7D

-Astronomy entities also have good examples

-Sum of all paintings -

https://www.wikidata.org/wiki/Wikidata:WikiProject_sum_of_all_paintings

-[Historical map properties](#)

Hilary: Discovery in maps via Wikidata

Marc: This map shows an entity type that's a mountain, let's retrieve all maps that depict a natural feature; spatial discovery--usually have explicit coordinates--bounding box--don't need to know administrative names--you can just zoom into a map and see a spatial representation; also interested in descriptive practice for historic maps that are note based--make them queryable--might not know to check a map for an image of a monument that might not exist--linked data/Wikidata opens up how we might make use of granular descriptions that can then feed into discovery in interesting ways

Marc: If we turn place of publication into an entity rather than as a text string, ie. Falkland Island war--place of publication of Argentina vs. UK

Matt: you can have both claims in same entity with different qualifiers to represent opposing views

Marc: having the source is important for making those claims

Hilary:Wikidata has more ways of recording and sourcing inaccurate data

Marc: Maps sometimes misstate scale; we have text conventions for representing erroneous scale and data representations; way to say that the stated scale is inaccurate; explored modeling questions within BF--in some ways easier to do in Wikidata

Matt: Can make statements about statements--provenance gets into quads--graph vs. RDF--will that ever become a problem?

Session 1

Facilitator: Hilary Thorsen

Note-taker: Christine Fernsebner Eslao

Recording:

https://stanford.zoom.us/recording/share/tbwe1ZvXxTtIHsFuCTxkpdFgkxW1_lsmk6pPMqmdLOjwLumekTziMw?startTime=1561486242000

Intros

Brittany Washington, Ransom

Paul Burley, Northwestern

Christine Fernsebner Eslao, Harvard

Hilary Thorsen, Stanford

Matt Miller, LC

Nate Traill, LC

Astrid Usong, Stanford

Huda Khan, Cornell

E. Lynette Rayle, Cornell

Burning topics?

None

Current interests in Wikidata

- Brittany: Training model for linked data, within Ransom
- Huda: Lookups and enrichment of metadata; discovery enhancements, search suggestions
- Astrid: Discovery, knowledge panels

- Nate: enriching id.loc.gov, “name authority lite”
- Matt: adding wikidata links to id.loc.gov, connections between systems, wikibase, discovery
- Christine: small projects,
 - focused collections- digitized maps,
 - collection of live music recordings from Boston area, with a focus on bands and venues
 - how does describing events fits into wikidata, or does it?
 - Incunabula, and lots of things.
 - Wikidata as a complement to work in ISNI. Things possible in wikidata that aren’t in ISNI, build connections between them. Some materials not clear that it’s appropriate to describe in wikidata.
 - Interested in wikibase for local authorities.
- Paul: personal interest, describing historical photographs, multilingual infoboxes
- CFE: defining skill sets, get a sense of people’s skill levels, who we connect with LD4; who has done what with each tools
- Lynette: limited exploration into APIs for use in QA lookups
- Hilary: working with Wikimedia Deutschland and best practices document from QA, to see if Wikidata can provide data differently
- CFE: reconciliation API
- Lynette: could be that I have yet to find something that could be a good fit. Search API doesn’t let you put in that kind of limiter. Only takes string query. Can’t pass in entity. Some sub-setting of data, but on property or item.
- Paul: posters in Hausa, which can’t be fully transcribed in MARC; not ethically appealing to present materials in Latin alphabet
- Nate: character limitations are because Voyager is LC’s system of record
- Matt: concern that Wikidata entity creation needs to be tied to library identifiers; if Wikidata entries could be used to generate MARC, that would be great
- Hilary: how do we make it easier to make connections between systems?
- Matt: Need more discussion with PCC, NACO Lite; would be great to package a library-flavored Docker image
- CFE: having more people in Id4 knowing what is possible through quick statements. It’s not going to take a long time to create a lot of data. Pre-packaged queries for how you evaluate stuff librarians care about
- CFE: how to federate wikibases together? Haven’t seen any examples. What is it that we could federate?
- Scott: my caution is that we’re still just learning about the space, but we’ll be talking about it in the next grant. Where does this fit in? Next year? Don’t want to lose momentum we’re building now to develop expertise. Want to make sure we can do what we set out to do.
- CFE: know more about this space than BIBFRAME editor space. UI that everyone seems positive about learning. Comes with built in SPARQL query engine. It’s a good teaching tool. Can see consequences of your actions.

- Huda: what are the advantages of having a wikibase instance? What are the expectations for the data?
 - Hilary: permissions, ontology changes
 - Matt: The only federated wikibase I've seen is the Wikimedia Commons metadata wikibase
 - Hilary: BNF/ABES CFP
- CFE: all data driving authorities- is it meant to be public? Not control, but ethical issues. Features that may not get built out unless we do it. There are community standards that are different. There may be some things important to us (library community vs. wiki community).
 -
- Topics for calls: Useful properties for discovery; useful widgets; ways to efficiently query for discovery/infoboxes
- ; SPARQL queries
- Library flavored Wikibase--what should be part of it?
 - [Etherpad](#) from Stacy Allison-Cassin and Dean Seeman's LD4 [presentation](#)