[Slack channel: https://ld4.slack.com]

LD4 Wikidata Affinity Group

2022-10-04
Tuesday, October 4, 2022 at 9am PT / 12pm ET / 16:00 UTC / 6pm CET ([Time zone converter](#))
Facilitator: Kyla Jemison, University of Toronto Libraries, ld4-wikidata-coord@googlegroups.com

Connection Information

- https://utoronto.zoom.us/j/87209417969
  Meeting ID: 872 0941 7969
  Passcode: drinkwater
  One tap mobile
  +15873281099,,87209417969#,,,,*3733301335# Canada
  +16132093054,,87209417969#,,,,*3733301335# Canada

  Dial by your location
       +1 587 328 1099 Canada
       +1 613 209 3054 Canada
       +1 647 374 4685 Canada
       +1 647 558 0588 Canada
       +1 778 907 2071 Canada
       +1 438 809 7799 Canada
  Meeting ID: 872 0941 7969
  Passcode: 3733301335
  Find your local number: https://utoronto.zoom.us/u/kD2UYCNYs

Recording

[Recording, chat, and transcript](#)

Presentation Slides and Other Materials

[Lana's slides](#)

Attendance (if this is your first meeting, please write your name, Wiki* username, institution or affiliation, and interest in Wikidata if you'd care to share)

- Raphael Sasayama, iconindexsymbol, Getty Research Institute
- Argyri Anastasopoulou, Municipal Library of Chania, Creta, Greece
- Eric Willey, Illinois State University
- Christa Strickler, University of Notre Dame
- Kyla Jemison,University of Toronto
- Roman Tashlitskyy, University of Toronto
- Lana Soglasnova, University of Toronto
- Hollis Wittman, Virginia Tech
- Heidy Berthoud, Library of Congress
- [Alli Khalil](),
- Robert Rohrbacher, Stanford University
- Natalia Mykhaylychenko, University of Toronto
- Anna Slawek (University of Toronto)
- Prema Desai (National Agricultural Library)
- Christine Fernsebner Eslao, Harvard Library (User:Fernsebner)
- Gerry Walden, Harvard Library
- Steve Baskauf, Vanderbilt Heard Libraries
- Soo Bae, National Gallery of Art
- Susan Hoover, University of Houston (User: Ɛs Huvər)
- Mary Aycock, Texas State University
- Kirill Tolpygo (UNC-Chapel Hill)
- Cammeron Girvin, Library of Congress
- Joe Cera, Berkeley Law Library
- Bob Kosovsky, New York Public Library | [user:kosboot]()
- Ivan E Calimano (Retired) | user calimano
- Bert Nyabadza, Zimbabwe Library Association (ZimLA) | [user:DeLilGwashoper]()
- Susan Radovsky (Harvard Library) | User: Sradovsk
- Karly Wildenhaus, New York Public Library
- Cliff Landis, AUC Woodruff Library
- Amy Ruskin, Northeastern University
- Pilar Ramirez Amurrio ([dokusare.com]())
- Darnelle Melvin, University of Nevada, Las Vegas
- Mingyan Li, University of Illinois, Chicago
- Z. Blace, Multimedia Institute in Zagreb, Open GLAM Croatia
- Alla Roylane, NYU
- David Bigwood (Lunar and Planetary Institute)
- Patricia Thurston (Yale)
- Amanda Xu (National Agricultural Library)

- Adam Schiff (University of Washington)
- Monika Soler Correa (University of North Carolina)

Agenda

Welcome + announcements

- Lana Soglasnova and Roman Tashlitskyy on adding Slavic names to Wikidata
- Q&A

Announcements

- ○ Next Wikidata Working Hour:
  - ■ Friday, October 14, 2022 at 10:00am PT / 1:00pm ET / 17:00 UTC / 7:00pm CEST (Time zone converter)
- ○ Next Wikidata Affinity Group Call
  - ■ Tuesday, October 18, 2022 at 9am PT / 12pm ET / 16:00 UTC / 6pm CEST (Time zone converter)
  - ■ Gadget, user script, and query sharing session
  - ■ Sign up to demo your favorite gadget, user script or SPARQL query during the call: 📆 Sign up for 10/18/22 LD4 Wikidata Affinity Group Call
- ○ Next Wikibase Working Hour
  - ■ Monday, October 24, 2022 at 11:00am PT / 2:00pm ET / 18:00 UTC / 8:00pm CEST (Time zone converter)

Help us organize the LD4-Wikidata affinity group!

We're moving to make it easier for more of us to participate in organizing LD4-Wikidata affinity group programming! Here's what you can expect in coming weeks:

- Starting **this month**, we're opening up coordination meetings + channel on Slack
  - ○ Coordination meetings take place once a month!
    - ■ Current schedule: Monday at 1pm PT/4pm ET (Time zone converter)
    - ■ Next call -- save the date: **November 14 at 11am PST/2pm EST** (Zoom link in Slack channel)
  - ○ Join LD4 Slack channel #wikidata-coordination to join the conversation and get further updates!
- Anyone can sign up for various rotational roles and propose group call/working hour topics, either during each organizing meeting or in the days following through the coord Slack
- Runsheets will be available for each role, and role hand-off will be a team effort. Depending on capacity/comfort level, you are welcome to take on a smaller role then grow into bigger roles.
- Here are roles that we've identified as gaps! These are some starting points, but we invite any and all sorts of contribution based on your skillset:
  - ○ note-taking for a call

- ○ documentation -- help refine and build on key pages, and continue to develop/nudge others to add content
  - ○ help communicate/promote programming
  - ○ present during a group call on some topic
  - ○ lead a working hour on some topic
  - ○ help facilitate calls

Weekly summary

Listed in [Wikidata:Status updates](#) and posted to the [mailing list](#) / [on-wiki](#) to your talk page

- Recent summaries
  - ○ [Week before 2022-09-12](#)
  - ○ [Week before 2022-09-19](#)
- Help write [the next summary](#)!

More editing communities

- *[Live Wikidata Editing](#),* a Wikipedia Weekly Network live streaming show hosted by Jan Ainali and Albin Larsson. Wide variety of topics covered, archive of YouTube videos going back to March 2020.
- Telegram Wikidata channels:
  - ○ Join Wikidata channel: https://t.me/joinchat/IeCRo0j5Uag1qR4Tk8Ftsg
  - ○ Join Wikidata Help channel: https://t.me/wikidatachat
  - ○ Join Wikidata Ex Música channel: https://t.me/exmusica
  - ○ Join Wikidata Lexicographical data: https://t.me/joinchat/TwVx898JKJTAajm4

Presentation Notes

Cyrillic Script, Name Authorities and Wikidata
Lana Soglasnova (University of Toronto Libraries) and Roman Tashlitskyy (University of Toronto Libraries) in adding Slavic names to Wikidata, with Kyla Jemison and acknowledgements to Alex Jung

- Lana: Catalogers often rely on Wikidata and Wikipedia in their work, but relevant information is isolated from library systems and the Wikimedia world
- Open bibliographic data from the library world into the Wiki world, as a way to give back
- Started as a learning exercise
  - ○ Example of a Wikipedia article about a village abandoned due to the Chernobyl disaster and a bibliographic record of a book about that village
  - ○ Book was not in references in the Wikipedia article, so they added the citation

- ○ Saw an opportunity for more systematic approach to working between Wikipedia and bibliographic data -- authority work and workflows bridge the divide!
- Kyla: Translating PCC authority records into Wikidata
  - ○ UTL Wikidata Training for Cataloguers, 2020 (thanks Alex!)
  - ○ Revisited training in 2022
  - ○ Participated in the PCC Wikidata pilot, developed training materials and documentation, had Wikidata Fridays within a smaller group
  - ○ One project to take NACO records and translate them into Wikidata schema
    - ■ Manual at first
    - ■ And then a reconciliation process through OpenRefine for working in larger batches of records
    - ■ Developed documentation, teaching and training community
  - ○ Lana and Roman began using the framework established here, though Slavic names/people presented some unique challenges.
- Roman:
  - ○ Wikipedia Editing Event for Ukraine
  - ○ Initiative in spring 2022, virtual edit-a-thon for 4 days
  - ○ Adding links to LC authorities as well as editing various types of Wikipedia articles about Ukraine
- Cataloging Workflows for Original Cataloging & Adding value to Wikipedia
  - ○ Gifts often provide unique content
  - ○ Can use the unique materials to add information to Wikipedia
  - ○ Adding citations to articles about people, creates links on Wikipedia pages to OCLC WorldCat to aid discovery & delivery workflows for readers
- Transforming LC authority records to Wikidata schema
  - ○ Incorporating Wikidata work within their cataloging workflows
  - ○ Difficult to automate due to the complexity of the romanization of Slavic names
  - ○ Examples where surnames end with -sky, -tsky, where names have relations to class status, places, religious factors
  - ○ Example of Volodymyr Zelenskyy and diverse ways of romanization
    - ■ Wikidata includes a wide variety of "also known as" but does not include the romanization per LC transliteration standards
    - ■ Q3874799
    - ■ Common in popular culture and media for spellings to vary across different countries and sources
    - ■ Looking at the official website for the president of Ukraine provides their own "official" spelling of the transliterated name, yet this contradicts LC transliteration rules per the Ukrainian Romanization table (different than the official Ukrainian transliteration table)
    - ■ Complicated enough with a very famous person, but much more so with a person who is less public, or there is less information publicly available about them.
  - ○ Example of Marco Paslavksyi, author of 1 book

- - - ■ Created an LC authority record, enumerating all of the variants of his name found in that item
        ■ Wikidata Q17683301
        ■ Complexities with diacritics, full, short, and colloquial versions of his name as well as the author having a military codename
        ■ Wikidata item label had a simplified ending, anglicized first and middle name, and "w" to convey the person's path through Poland, another variation of the name not covered in the LC authority record
        ■ Important to do research to avoid creating duplicates, predict what name variants could be out there when searching for the person's Wikidata item -- hard when you don't fully understand their biography
            ● Duplicate records were merged in Wikidata for this person
    ○ Example of Kyiv mayor, Vitali Klitschko where alternative spellings of his name exist, relating to periods of time in his life spent in Germany. Can compare to his current Twitter.
    ○ Automation is not straightforward because cannot match LC transliterations with what may exist in Wikidata
    ○ Also, some frustration where the work is subject to future editing, regardless of the quality of the work you have done to create accurate statements.

- Lana: Cyrillic script challenges for holdings outside countries of origin
    ○ Keep in mind that we work in a globalized environment, where people often move across geographic locations, encountering language variants, and you can also be encountering diaspora and emigre publications
    ○ Slavic names represent possessive suffixes, linguistic diversity of Slavic regions, social adaptation, and the standards created by governments and bibliographic standards
    ○ Slavic Cataloging manual contains several chapters for unique considerations based on country of origin (Westernized Slavic names, Polish, Czech, Slovak, Kyrgyz, et cetera)
    ○ Examples of a character name in Nabokov's *The Gift*, in different languages over time since its original publication in 1937
    ○ Example of cyclist Denis Menchov - different transliterations between LC and Wikidata names
    ○ Example of changes to Cyrillic scripts across continents and times -- Microbiologist's name had spelling variants among Russian publications different from their Germanized official form and then their obituary published in Argentina.
    ○ Risks of creating duplicate items where you don't know the biographical information of people who moved across geographic regions where different official and unofficial name representations.

Q&A

Q: Have you adopted best practices for adding LC spellings to Wikidata items if it is not included in the label and name variants of the Wikidata item? How do you try to think of every possible spelling or try to rely on published sources?
A: Roman -- the published sources can contain name variants, will include them in the LC record if that is so.
Lana -- LC romanization not possible to add as a formal language, though they could add an additional alternative name with LC romanization qualifier. They will also be sure to link to the LC authority record in the Wikidata item.

Q: (Comment about managing transcriptions of names over time)
Lana -- the Russian passport official translation has changed at least four times, so this is something they need to be constantly aware of as well.
Roman -- Zelenskyy issued passport in the 1990s when the rules were different, possibly related to how his name is different over time. The impact of bureaucracy, difficult to change official documents

Q: Is there documentation about the official changes?
Lana shared an article in the chat

Q: Workflow to add info to Wikipedia -- is this a lot of extra work? Do you only do this for special kinds of cataloging?
Lana -- as a supervisor, you need to have the support of your management, and their management is open to these kinds of contributions to Wikipedia and Wikidata. There is an upfront investment in learning to do these things. You have to be selective, pace yourselves because the amount of projects you can do will exceed your capacities. For the authority work processes, they'll scan pages, and save them for later (don't need the book in hand). What can we catalog right now, what can wait for later.
Roman -- not very time consuming. They've been focusing on gift collections, very rare materials. Creating the original records for the books is time consuming, but it's not especially time consuming to add additional links or references to Wikipedia which they believe is valuable for aiding researchers.