

Pronouns+ in Wikidata: Wikimedia outreach meeting

Date/Time: Wednesday, February 1, 2023, 9:30AM PT/12:30PM ET ([Time zone converter](#))

Duration: 1.5hr

Connection information

Join Zoom Meeting

<https://utoronto.zoom.us/j/87971846222>

Meeting ID: 879 7184 6222

Passcode: [drinkwater](#)

One tap mobile

+16475580588,,87971846222#,,,,*8340593472# Canada

+17789072071,,87971846222#,,,,*8340593472# Canada

Dial by your location

+1 647 558 0588 Canada

+1 778 907 2071 Canada

+1 438 809 7799 Canada

+1 587 328 1099 Canada

+1 613 209 3054 Canada

+1 647 374 4685 Canada

+1 786 635 1003 US (Miami)

+1 971 247 1195 US (Portland)

+1 206 337 9723 US (Seattle)

+1 213 338 8477 US (Los Angeles)

+1 253 215 8782 US (Tacoma)

+1 267 831 0333 US (Philadelphia)

+1 301 715 8592 US (Washington DC)

+1 312 626 6799 US (Chicago)

+1 346 248 7799 US (Houston)

+1 470 250 9358 US (Atlanta)

+1 470 381 2552 US (Atlanta)

+1 602 753 0140 US (Phoenix)

+1 646 518 9805 US (New York)

+1 646 558 8656 US (New York)

+1 651 372 8299 US (Minnesota)

+1 669 219 2599 US (San Jose)

+1 669 900 6833 US (San Jose)

+1 720 928 9299 US (Denver)

Meeting ID: 879 7184 6222

Passcode: 8340593472

Find your local number: <https://utoronto.zoom.us/j/kx0SdSA5t>

Housekeeping

- This is a gender and sexuality affirming space. We are not here to debate the existence and validity of identities. Please help us keep our discussion constructive!
- To contribute a question or comment, please raise your hand to talk or type “duck” 🦆 in chat. We will be taking turns to talk, and the facilitator will keep a running list of names. If someone has spoken many times, the facilitator may reorder the list to prioritise folks who have spoken less.
- We have received requests to record the meeting. We will record, but will review and assess the recorded video post-meeting to figure out how and what parts we will share. We will do this to accommodate comments shared that may be more personal, closer-to-heart, or otherwise less appropriate for public release.
- We would greatly appreciate your help with note-taking! Volunteers?

Attendance

Please include your name, institution, and Wikidata username (all optional!)

- Crystal Yragui, University of Washington, Clements.UWLib
- Kat Koziar, University of California, Riverside, koziarke
- Allison Bailund, San Diego State, Sandrileine
- Sasha Frizzell (Binghamton University)
- Sara Thomas, Wikimedia UK, user:lirazelf / user:Sara Thomas (WMUK)
- Owen Blacker, WM LGBT+ User Group, User:OwenBlacker
- Jessika Davis, Semantic Lab at Pratt | User:DatamuseAttitude
- Karen Hwang, Semantic Lab at Pratt & AAAC |
- Kalliopi Mathios, Teachers College, CU | user: kind_data
- Peter Stephen, Library and Archives Canada
- Arielle Rodriguez, University of Washington, Rodriguez.UW
- Raphael Sasayama, Getty Research Institute
- Steve Baskauf, Vanderbilt University Heard Libraries, baskaufs
- Amy Furness, Art Gallery of Ontario, artchivist1
- Greta Suiter, Ohio University
- Neil M Frau-Cortes, Broward County Public library, dhneil
- Sarah Jones, University of Nevada Las Vegas

- Jere Odell, IUPUI, User:Jaireeodell
- Abby Dover, Northwestern University
- Karly Wildenhaus, New York Public Library, user:infopetal
- Michaela Blanc, Wikimedian in Residence at Perez Art Museum Miami | User: MichaelaBlanc
- Alexa Vallejo, Philadelphia Museum of Art
- AJ Robinson, Washington University in St. Louis, User: LastChapter
- Meg Wacha, City University of New York, Wikimedia NYC
- Lane Rasberry, University of Virginia, Wikimedia LGBT+
- Stephanie Fell, University of Nevada, Las Vegas
- Florence Devouard
- Matthew Haugen, Columbia University Libraries, User: Matthewhaugen
- Tasha Bergson-Michelson, Castilleja School (Palo Alto, CA, gr 6-12) User: Oughtta Be Otters
- Lihong Zhu, Washington State University
- Ahava Cohen, National Library of Israel, User: AhavaCohen
- Peggy Griesinger, University of Notre Dame

Acknowledgments

- Clair Kronk! Life keeps her busy and she is not here today, but so much of this is her brainchild and this work would not exist without her :)
- Wikidata community members who contributed when we were crowdsourcing resources/concerns/recommendations
 - Special shoutout to LD4-Wikidata Affinity Group Call participants from February and July 2022 calls

Agenda and notes

- Introductions
 - this meeting is video recorded - please be aware as attendees that this is recorded
 - attendees invited to take notes
 - one note taker is Lane Rasberry
 - thanks
 - Clair Kronk - biomedical informatician, and a leader in organizer this project
 - Wikidata community members who responded to calls for support
 - special thanks to the Linked Data 4 Production (LD4) Wikidata group and meetup series

- Crystal Yragui
 - cataloger at University of Washington
- Alex Jung
 - University of Toronto, has done Wikimedian in Residence work at the library
- Goals for today
 - Share past and ongoing work to make gender-related information on Wikidata more precise, accurate, and consensual
 - there are lots of things packed into the "sex and gender" Wikidata property
 - this talk will present on pronouns, but there are many other related issues
 - Hold space for feedback, general discussion, and concerns
 - Seek collaboration on further work
- Background
 - Problem statement
 - Concerns about accuracy and precision of data modelling, implementation of gender for persons on Wikidata from both data and ethical perspectives
 - it is possible to find places where gender is mistaken or harmful
 - there are a series of ethical issues which arise from documenting gender
 - What we're trying to do
 - Address issues documented below through best practices, changes to Wikidata; for example: items, properties, lexemes
 - Implement changes and make sure to clean up legacy data
 - How we've been doing it
 - Bi-weekly meetings, taking on work in phases based on general categories of concern
 - there have just been a few regular attendees of these meetings
 - Student project at University of Washington
 - Arielle Rodriguez is working on Wikidata P21
 - this project is planned for next two academic quarters
 - Communication thus far
 - [Feb 8, 2022: LD4-Wikidata Affinity Group Meeting](#)
 - Key takeaways on community needs:
 - "What not to do"
 - How to remove bad data if needed + ensure it's not added back – anticipating some limits, pending further convo
 - Selecting (or ignoring) sources for identity/demographic info

- Contextualising statements that vary over time or by circumstance
- [July 26, 2022: LD4-Wikidata Affinity Group Meeting](#)
 - Progress report and feedback-gathering
- Past/living documentation
 - Documents:
 - [Gender in Wikidata Priorities](#) (in particular, see [#Areas of concern](#)): There are issues of:
 - consent
 - problems with P21 property (gender and sex are conflated)
 - Given names and labels are challenging because of deadnaming
 - Also bots issue that overreach
 - [Pronouns working doc](#)
 - What not to do
 - Recommended changes to Wikidata
 - Modeling exemplified in LGBTDB
 - LGBTDB is a separate Wikibase instance which hosts certain demographic data:
https://lgbtdb.med.yale.edu/wiki/LGBTdb:About#Names_and_Pronouns
 - Implementation plan
 - [Logic model for application of pronoun sets](#)
 - based on sketch provided in pronouns working doc
 - (remarks from Lane, notetaker)
 - chart raises issues of consent
 - raises issues of documentation of consent
 - raises issues of journalists or experts or some process tracking consent
 - Wikipedia / Wikidata has no obvious existing processes for these things
 - this idea and logic model is coming from the team of 3 researchers
 - the team's idea is to propose this to the Wikidata community for adoption as Wikidata guidelines
 - the team are each Wikidata editors
 - there are no particular Wikidata community groups or other individuals identified as designing this; this is a work in progress from this team
 - to fact-check statements the proposed model is
 - P5102, "nature of statement" --> Hypothesis, P3680 "statement supported by" ---> ...

- Lane commented expectation of "based on heuristic"
<https://www.wikidata.org/wiki/Property:P887>

■

○

- Items for feedback
 - Suggested WD changes – [Best practices](#)
 - What are your thoughts re:
 - Pronoun sets
 - Pronoun sets as items vs. lexemes
 - Pronoun set has part or parts (P527) pronoun (Q36224)
 - Multiple pronoun sets
 - Specify language for each added pronoun set
 - Restriction to humans and fictional characters
 - No start or end dates allowed without references that are direct quotations
 - comments from experienced Wikidata editor:
 - These Wikidata demographic data talks come up frequently and this person expressed that they are stressful to LGBT+ community stakeholders
 - There are many of these talks.
 - The talks put stakeholder communities in conversations with hate-based communities, such as transphobic, discriminatory, and whatever else.
 - When these talks happen, stakeholder communities are called to either expend the resources to advocate for themselves or to have other power brokers make decisions on their behalf by default.
 - The fundamental forces at play are colonialism, power disparities, and external communities making decisions which greatly affect vulnerable communities.
 - Wikidata is itself harmful and does not have Wikimedia community support for many of its routine functions. So for example, a university of research group which complies with Wikidata community norms or Wikimedia research standards does not have the support of the Wikimedia community itself.
 - Wikimedia Enterprise, which is a commercial effort of the Wikimedia Foundation, is controversial in the Wikimedia community who see it as empowering to corporations and disempowering to the Wikimedia community. Demographic data which this project is encouraging is inherently in corporate benefit while not having support of community stakeholders.

- Community member expressed great appreciation for this research, and because this all is the closest to an ethical process for data curation, but also concern that it can still be harmful to be done.
- Response from Lane
 - I support what that user is saying as normal and common Wikimedia community perspectives. Takeaway: all research and activity in this space causes great harms to vulnerable communities.
- Response from Alex:
 - There are benefits to this research and activity.
 - There are individuals who benefit from this kind of demographic profiling.
 - There are ways to do this activity ethically and helpfully.
 - There are ways to fix the problems that Wikidata is causing.
 - No one is seriously having the conversation about shutting down Wikidata so I suppose we have to find a way to accept it and adapt it for the better.
- Ahava
 - I am with the Israeli Wikipedia community and for our languages, Hebrew and Arabic, second-person pronouns matter a lot
 - This talk did not include addressing pronouns in those languages, and what applies in English does not inform how things work in those languages
 - Wikidata is multilingual and English only policy cannot be acceptable
 - There are grammatical gender considerations beyond pronouns
 - Scholia is a tool which can balance gender and representation off-wiki
 - I would like to see inclusion of pronoun policies for languages outside of English
- Alex:
 - Our model has only considered English language to this point.
- Crystal:
 - We will likely need several language/cultural sections in a shared best practices document.
- (others) agree with Ahava
- Lane
 - You are proposing something new - great, no comment on that
 - Not this team's issue, but I am fairly sure that there is no documentation on how things are now
 - If we lack documentation on how things are now, then difficult to discuss how or why we should change things
- Response from Crystal
 - Right, we have listed different places where we identified documentation of practices

- We thought to start with pronouns to get some clarity on this
- For example, we have identified that there are some Wikidata bots that operate to add gender to biographical items under certain circumstances
- one of those circumstances is interpretation of the gender of given names
- assigning gender based on given name is incorrect with some frequency
- Karly
 - I do Wikidata editing to correct demographic disparities in representation
 - I hear and acknowledge that this data curation causes problems
 - Wikidata properties sometimes have "property constraints" in place which prompt editors to add P21, "sex or gender", in cases where editors would find it debatable as to why it is needed
 - Is there a central discussion place for talking about why constraints are in place? I am unaware of any such discussion
 - Proposal: remove all constraints requiring sex or gender. It is never essential. Why is it in place anyway?
 - consider
 - <https://w.wiki/66pV>
 - query of all properties with that constraint (Will Kent created this, thanks)
 - Comment from Wikidata post-meeting: "I wrote a [query](#) that expands on the one discussed during the call, with the goal of more fully listing constraints that suggest or require that certain statements be accompanied by statements involving *any* properties related to gender. The query result may not yet fully list those cases since the properties may not yet [be annotated](#) completely. —[BlaueBlüte](#) (talk) 00:07, 2 February 2023 (UTC)"
- Proposed logic for the application of pronoun sets – [Flowchart](#)
- Prioritisation of areas of concern generally
- Future work
 - Next steps (on our end)
 - Setting up a community task force, considering a list and a WikiProject
 - Pronouns are one example of something that applies to everyone – while this work supports the LGBTQ+ community, fundamentally broader in scope than “an LGBT issue”
 - Looking to strategise on implementation/rollout (see under “how you can help”)
 - How you can help – You can email us (cec23@uw.edu; jya.jung@utoronto.ca) or **add notes below**
 - Input on community setup
 - Context–under LGBT+ umbrella? Part of something else?
 - Audience is much broader–considering bot activity on P21 as example
 - List? WikiProject?

- Help us identify community needs we missed
 -
- Pronouns rollout/implementation
 - Some immediate items:
 - Setting up pronoun sets
 - Relating pronoun sets to Wikidata lexemes and reconciling any issues
 - Doing this work in more languages
 - Recruit language expertise
 - Activate other Wikidata language communities on these issues more broadly
 - Call for participation – Are you willing to help with this?
 - Have a list – will flag discussion that need community input (same list as above)
- Starting to work on a specific area
 - Use cases – send us, identify/collect conversations; sticky points
 - Flag overreaching bot activity
- Ideas and further conversation
 - The Wikimedia Foundation and the Wikimedia community routinely have conflict on ethical and social issues
 - the Wikimedia Foundation regularly avoids addressing major community concerns, until and unless those concerns include off-wiki support from media or similar
 - (discussion of Framgate) - to review, as one perspective in a complicated space, this was a case where the Wikimedia Foundation applied a punishment to an editor in circumvention of the Wikimedia community's own justice and arbitration process. The problem was the circumvention of community norms.
 - (discussion of other ways to collab with WMF and Wiki community)
 - (mention that those “community norms” are set by a very Privilege-based group of people: very white, very cis, very male, very anglophone, very based in the Global North etc)
 - propose to establish Demographic Undertaking with Care and Kindness (DUCK)
 - this could be a WikiProject, task force, or named project to set guidelines in this space
 -
- Staying in touch (we will not spam you!)
 - Please fill out [this google form](#) (1 min)