Qualitative understanding of marginalized researchers reasons to contribute or not to the Wikimedia projects

Paz Bernaldo[1], Yo Yehudi[1]
[1] OLS - Open Life Science Limited, UK

Abstract

Wikimedia projects' are known as spaces where accidental or intentional biases might result in unwelcoming environments, especially for groups considered marginalized or excluded. Destructive criticism, for example, has been documented to drive away women as Wikimedia contributors. It is, therefore, likely that other marginalized communities could be similarly affected [1,2,3]. As Wikimedia says, "Participation in the Wikimedia global movement has been easier for some users and not others" [4].

We propose to conduct a qualitative investigation into specific actions, behaviors, and other factors that have helped or hindered marginalized scientists' efforts to contribute to Wikimedia projects. Specifically, we would focus on environmental justice researchers, such as those working on or studying Sacrifice Zones [5]. If there are no such researchers with experience contributing to Wikimedia, we will further expand our scope to other researchers from marginalized communities.

Introduction

This project aims to understand why scientists from marginalized communities working on environmental justice issues either stay or leave the Wikimedia community. We would be looking for factors, including actions or behaviors, they encountered that influenced their decision to contribute, keep contributing, start contributing and stop contributing.

The reason to focus on this specific group of researchers relates to the nature of the topic. Environmental justice movements often include or explicitly aim to include scientific evidence as part of their narratives. Sacrifice zones (as pollution hot spots) present a series of intersectionalities that make them unappealing for public media coverage. Having good (diverse, reliable, up-to-date) knowledge available for journalists, policymakers, other related movements, and other researchers is highly needed.

An in-depth qualitative understanding of the reasons marginalized scientists or researchers decide to stay or leave the Wikimedia community is needed to tackle the problem of unwelcoming environments. This is the **first expected outcome** of this proposal, and it relates to Wikimedia's explicit desire to partner with communities to understand why participation in the global Wikimedia movement has been so difficult for some users or groups of users [6].

A second expected outcome of the analysis is a list of new research questions worth investigating by the Wikimedia community.

A third key outcome is to coach a group of social scientists, already part of Wikimedia, in the reflexive approach used in this study, so they could tackle the new research questions. This aligns with Wikimedia's work on building foundational infrastructure, specifically "new methods and guidance for mission-aligned research" [6].

It is therefore expected the entire Wikimedia community could benefit from gaining the proposed understanding.

Start date: 1st of July 2024
End date: 30th May 2025.

Related work

This project will be run by OLS, as part of their dedication to creating a high-quality evidence corpus for effective open research. As well as OLS's history of running open community training, senior researcher Yo Yehudi has completed research into open community sustainability indicators [8], and barriers to open data sharing [9], and is a prominent open infrastructure advocate [10]. Researcher Paz Bernaldo (PI), thanks to projects like Vuela (vuela.cc), contributes relevant experience of working closely with marginalized groups and researchers, specifically trying to find ways in which local problems can anchor purposeful and horizontal collaborative processes between these groups.

The results of this research are intended to shed in-depth light on the problem, and therefore allow for actionable and sensible measures to be tested by the Wikimedia community.

Methods

In-depth semi-structured one-to-one online interviews will be conducted and the data analyzed using a qualitative analysis approach called *Reflexive Thematic Analysis*, developed by scholars Braun and Clarke [7]. Importantly, the study will be open. Anonymized data (transcripts) and results will be open. But also the whole analytical process would be equally open: how and why the researcher arrives at those results and not others, following RTA. This reflexive approach includes the researcher's positionality but goes much further by explaining the assumptions underpinning the study's design and analysis.

By using RTA, the subjectivity of the PI is used as an analytical resource, not as a bias that needs to be controlled. In the final repository future researchers will be able to find the data, results, and list of themes, but also all the justified analytical decisions.

Expected outputs

First output (Deliverable 1): open-access scientific publication, documenting the results and process.

Second output (D2): shorter and nonacademic document summarizing key insights for decision-makers.

Fourth output (D3): Online discussion meeting with all researchers involved in the process (interviewees will be also invited).

Risks

1. Not finding enough environmental justice researchers who have contributed to Wikimedia.
2. Not finding enough Wikimedia social scientists interested in RTA and investigating the resulting questions.
3. The PI not being able to continue due to unforeseen circumstances. OLS would find a qualitative researcher as replacement.

Community impact plan

This research by its nature will look at researchers outside Wikimedia. This research will provide contextual questions for future researchers to answer. Equally important is the coaching of researchers in using RTA.

This study will present the process and results to interviewees and their communities of practice (likely local activists, community organizers, local authorities, etc.).

Evaluation

We see three ways to evaluate this work.

1. Concise explanation of results: enablers and barriers to contributing to Wikimedia and ways to overcome the barriers and strengthen enablers.
2. Complete deliverables D1, D2, D3.
3. Social scientists group coached and happy to try further research.

Budget

- 38,000.00 USD - salaries.
- 5,000.00 USD - ethics waiver, event facilitation, transcription support, and APC for the academic article.
- 6,500.00 USD overhead at 15%.

Prior contributions

- Vuela Projects events by Paz Bernaldo to the broader scientific community. These involved open hardware-making workshops with more than 100 people: from academic researchers to unhoused immigrants.
- OLS events to the broader scientific community.
- The OLS impact research (undergoing) is being conducted by Paz Bernaldo, using RTA.
- Yo Yehudi has run a year-long survey and GitHub API-based indicator study, looking into the sustainability of online communities, and is familiar with qualitative and quantitative analysis techniques, and pressures on collaborative open communities [8].

References

1. Zheng, X., Chen, J., Yan, E., & Ni, C. (2023). Gender and country biases in Wikipedia citations to scholarly publications. Journal of the Association for Information Science and Technology, 74(2), 219-233. https://doi.org/10.1002/asi.24723
2. Gunawardena, S. D. G. S. D., Devine, P., Beaumont, I., Piper Garden, L., & Murphy-Hill, E. (2022). Destructive Criticism in Software Code Review Impacts Inclusion. Proceedings of the ACM on Human-Computer Interaction, CSCW2(292), 1-29. https://doi.org/10.1145/3555183
3. Smirnov, I., Oprea, C., & Strohmaier, M. (2023). Toxic comments reduce the activity of volunteer editors on Wikipedia. ArXiv. /abs/2304.13568
4. Fost, D., Stinson, A., & Miranda, L. (2019, October 4). Research – Wikimedia Foundation. Wikimedia Foundation. Retrieved December 15, 2023, from https://wikimediafoundation.org/research/
5. Alvarez, C. H., & Evans, C. R. (2021). Intersectional environmental justice and population health inequalities: A novel approach. Social Science & Medicine, 269, 113559. https://doi.org/10.1016/j.socscimed.2020.113559
6. Programs - Conduct Foundational Work. (n.d.). Wikimedia Research. Retrieved December 15, 2023, from https://research.wikimedia.org/foundational.html
7. Braun, V., & Clarke, V. (2022). Thematic Analysis: A Practical Guide (V. Braun, Ed.). SAGE.
8. Yehudi, Y., Goble, C., & Jay, C. (2023). Sustainability indicators in an open online community. ArXiv. /abs/2309.12120
9. Yehudi, Y., Goble, C., & Jay, C. (2023). Sustainability indicators in an open online community. ArXiv. /abs/2309.12120
10. Knowles, R., Mateen, B. A., & Yehudi, Y. (2021). We need to talk about the lack of investment in digital research infrastructure. Nature Computational Science, 1(3), 169-171. https://doi.org/10.1038/s43588-021-00048-5