Codifying Digital Behavior Around the World:
A Socio-Legal Study of the Wikimedia Universal Code of Conduct

Florian Grisel

University of Oxford

Giovanni De Gregorio

Universidade Católica Portuguesa

Abstract

The research project aims to examine the Wikimedia Universal Code of Conduct (the "UCoC") as a critical step for the governance of online behavior across Wikimedia projects. The research team will investigate the codification process of the UCoC, focusing on its origins, its structure and the narratives guiding its enforcement. Based on empirical data collected through interviews and document analysis, this project will analyze the dynamics surrounding the adoption of the UCoC and debates concerning its enforcement. The results of this research will be disseminated through various channels, including conferences, a final report and an open-access paper published in an international journal.

Keywords: UCoC; content moderation; universal guidelines; Wikimedia Foundation; digital governance

Introduction

The spread of harmful behavior online has raised questions concerning the protection of users' rights and community values. Particularly, hate speech and disinformation have increasingly turned digital environments into hostile spaces which potentially lead users to abandon these spaces. These challenges have encouraged online providers and networks to produce standards and rules on which they can rely when regulating conduct that could affect their communities.

The emergence of the UCoC is a significant development in this regard. The UCoC seeks to "provide a <u>universal</u> baseline of acceptable behavior for the entire movement without tolerance for harassment."[1] This set of guidelines is "being developed in consultation with the Wikimedia community with respect to context, existing local policies, as well as <u>enforcement and conflict resolution structures</u>".[2]

The proposed project aims to examine the codification process of the UCoC and its enforcement guidelines. By examining this process, we hope to explore two key questions. The first question is to examine how the drafters of the UCoC identify *universal* values across Wikimedia projects, despite significant differences among these projects. The second question is to understand how the *inclusive* values of the UCoC interact with the existing "enforcement and conflict resolution structures" of the Wikimedia projects, which are often criticized for being *non-inclusive*.[3]

This project will take place from September 2023 until May 2024, and it will involve two scholars and their research teams based in leading academic institutions in Europe.

Methods

Our methodology is empirical and inductive. Our goal is to give voice to the actors who have prime knowledge of the UCoC, before identifying theoretical frames that provide relevant explanations for the users' experience. We will gather empirical data to address the research questions of this project.

Firstly, the research team will conduct a series of semi-structured interviews with the members of the UCoC drafting committees, revisions committee and coordinating committee, as well as administrators and members of the arbitration committees in two language editions of Wikipedia (English and French).

---

[1] See <https://meta.wikimedia.org/w/index.php?title=Universal_%20Code_of_Conduct&oldid=20931352> (last checked March 21, 2023, emphasis added).

[2] *Ibid.* ((last checked March 21, 2023, emphasis added).

[3] See, *eg*, Zachary McDowell & Matthew A. Vetter, *Wikipedia and the Representation of Reality*, Routledge, 2021. We note that the Dutch arbitration committee referred to the UcoC in one of its recent decisions (see <https://nl.wikipedia.org/wiki/Wikipedia:Arbitragecommissie/Zaken#Deblokkade_Hans_Muller_(3)> (last checked March 21, 2023).

We have selected the latter editions for several reasons: (i) they are, respectively, the first and fifth largest Wikipedias, which will make the results of our research more representative and generalizable, (ii) they allow to contrast the ways in which editors grounded in different cultural environments react towards the UCoC, and (iii) they open up the possibility for the project to expand towards areas of the world that are currently understudied (e.g., Africa) as a subsequent step.

Secondly, the team will track the ongoing discussions and debates within Wikimedia communities relating to the codification of the UCoC and its enforcement in content and discussion pages, and email lists. We will do a content analysis of these pages with the coding of key terms and discussion items. The goal is to explore the ways in which Wikipedians talk about and make sense of the codification efforts initiated by the Wikimedia Foundation.

We note in this regard that "Wikipedians are [...] passionate about the issue of governance by and for their decentralized group of volunteers" and that "[t]he messaging from the Wikimedia Foundation about the new UCoC has been careful to make clear that Wikipedia is not intending to become more centralized or professionalized in the manner of Facebook and Twitter."[4]

We have identified different places in which these discussions take place: (i) the 2020 initial consultations launched by the Trust & Safety Team of the Wikimedia Foundation;[5] (ii) the 2021 local consultations;[6] (iii) the 2021 global consultations;[7] (iv) the various ratification surveys and votes;[8] (v) the

online archives of Wikipedia's newspaper (*The Signpost*);[9] and (vi) the online archives of Wikipedia's email lists such as Wikimedia-I, Wikipedia-I and WikiEN-I.[10]

Conclusions

During the project, the research team will issue two **interim reports** detailing the progress in the research (in November 2023 and February 2024).

At the end of the project (May 2024), the research team will issue a **final report** (around 5,000 words) explaining the main findings of the research. This report will also serve as a first step for the research team to initiate a broader project on the UCoC.

The three reports (interim and final) will be advertised for comments on MetaWiki:Research and the final report will also be used for dissemination activities (e.g., conference submission).

We will subsequently write a **scientific paper** (around 10,000 words) based on the results of this project. We will submit this paper to an international journal specializing in law and social sciences.

The research outputs will be presented at the LSA Annual Meetings (June 2024), Wikimania 2024 (July or August 2024), and the GigaNet Annual Symposium (November 2024).

---

[4] Stephen Harrison, "The Tensions Behind Wikipedia's New Code of Conduct," *Slate*, March 2, 2021.

[5] See <https://meta.wikimedia.org/wiki/Universal_Code_of_Conduct/Initial_2020_Consultations> (last checked March 21, 2023).

[6] See <https://meta.wikimedia.org/wiki/Universal_Code_of_Conduct/2021_consultations/Enforcement> (last checked March 21, 2023).

[7] See <https://meta.wikimedia.org/wiki/Universal_Code_of_Conduct/2021_consultations> (last checked March 21, 2023).

[8] See, eg, <https://meta.wikimedia.org/wiki/Universal_Code_of_Condu

ct/2021_consultations/Ratification_survey> (last checked March 21, 2023).

[9] See <https://en.wikipedia.org/wiki/Wikipedia:Wikipedia_Signpost/Archives/Years> (last checked March 21, 2023).

[10] See, eg, <https://en.wikipedia.org/wiki/Wikipedia:Mailing_lists> (last checked March 21, 2023).

Codifying Digital Behavior Around the World:
A Socio-Legal Study of the Wikimedia Universal Code of Conduct

Florian Grisel

University of Oxford

Giovanni De Gregorio

Universidade Católica Portuguesa