

Harvard Negotiation and Mediation Clinical Program

# REACHING THE ZONE OF ONLINE COLLABORATION

**Recommendations on the Development of  
Anti-Harassment Tools and Behavioral Dispute  
Resolution Systems for Wikimedia**

BY HAN A. LEE & JOSEPH CRUPI  
WITH CLINICAL SUPERVISION BY ANDREW MAMO



This report has been prepared by the Harvard Negotiation & Mediation Clinical Program (“HNMCP”) for the Wikimedia Foundation on a *pro bono* basis. HNMCP makes this report available under a Creative Commons Attribution-ShareAlike license. To view a copy of this license, please visit <https://creativecommons.org/licenses/by-sa/4.0/>. HNMCP assumes no responsibility for the implementation of the recommendations expressed herein or for any derivative works that may be based on this report.

# A C K N O W L E D G M E N T S

TO ALL WHO MADE THIS POSSIBLE . . .

First, we would like to thank the Wikimedia Foundation and the greater contributor community who invested their energy and time to teaching us about the complexities and wonders of Wikimedia as a whole. All of the stewards, administrators, and editors that volunteered their hours to invest in the long-term health of the community are truly the lifeblood of Wikimedia's free knowledge movement.

Second, we would like to extend our deep gratitude to all the volunteers who contribute their knowledge to the world at large, to the veteran users who serve as community mentors to incoming Wikimedians, and to new users who devote their passion and patience to a movement that is continuing its growth.

To our Professors at the Harvard Law School and fellow dispute resolution system designers at the Harvard Negotiation and Mediation Clinical Program (“HNMCP”), thank you for your continued support and encouragement. Your thoughtful feedback, patient guidance, and insightful advice have given us inspiration for this project and beyond.

Especially you, [Andrew Mamo](#). Thank you for always keeping us in mind. We could not have made it without you.

Sincerely,

**HAN A. LEE**  
**JOSEPH CRUPI**

# TABLE OF CONTENTS

<b>EXECUTIVE SUMMARY</b> .....	3
<b>INTRODUCTION TO THE WORLD OF WIKIMEDIA</b> .....	5
The Community Health Initiative and Harassment .....	7
<b>OUR APPROACH</b> .....	11
Questions Presented.....	11
Methodology .....	11
Challenges .....	13
Maximizing Impact.....	14
<b>FINDINGS</b> .....	16
Systems.....	16
FINDING 1: Users underreport behavioral issues due to a community culture that emphasizes dialogue as well as difficulties with the reporting process. ....	17
FINDING 2: Noticeboard systems do not facilitate effective case management. ....	23
Policies and Guidelines.....	26
FINDING 1: Standards related to blocking and other forms of administrative enforcement are often vague. ....	26
FINDING 2: Standards frequently set a high bar for administrative action.....	28
FINDING 3: Communities typically do not have a systematic method for converting precedent into standards. ....	30
Processes .....	32
FINDING 1: Communities typically do not have clearly delineated processes for building consensus or for determining when consensus has been reached. ....	32
FINDING 2: Users who operate unilaterally within the consensus-building process may exert more influence on discussion outcomes than users who focus on building consensus and will not take action in the absence of consensus.....	33
FINDING 3: Extended community discussions tend to follow an unstructured conversational pattern that is difficult to navigate.....	34
Discussion processes on systems .....	34
Discussion processes on policies and guidelines.....	37
Discussion processes on behavioral incidents .....	37
Cross-Project Interaction .....	38
<b>RECOMMENDATIONS</b> .....	39

Systems .....	39
Policies and Guidelines .....	41
Processes .....	43
Cross-Project Interaction .....	46
<b>APPENDICES</b> .....	<b>48</b>

# EXECUTIVE SUMMARY

In Fall 2017, the Wikimedia Foundation engaged the Harvard Negotiation and Mediation Clinical Program (“HNMCP”) to provide recommendations for management of reported misconduct across Wikimedia platforms as part of its three-year Community Health Initiative aimed at reducing harassment within Wikimedia projects.

HNMCP collected information through observations of Wikimedia communities, interviews with experts and stakeholders, an online focus group, and academic research. Based on this information, HNMCP provided findings and recommendations related to systems for addressing behavioral issues, community policies and guidelines, and community processes as well as the potential for cross-project interaction.

## FINDINGS



Users underreport behavioral issues on Wikimedia projects due to a community culture that emphasizes dialogue as well as difficulties with the reporting process. These problems are compounded by a noticeboard system which does not facilitate effective case management.

## RECOMMENDATIONS



### SYSTEMS

1. Centralize and increase the visibility of the reporting system; and
2. Replace the noticeboard system with a single centralized evaluation system.

### POLICIES & GUIDELINES

Standards related to administrative enforcement are often vague, and communities typically do not have a systematic method for converting precedent into standards. Additionally, standards frequently set a high bar for administrative action. Together, these factors contribute to an environment in which policies and guidelines are not rigorously enforced.

1. Supplement policies with more specific guidelines informed by community practice; and
2. Better integrate continue to build connections among policies, and guidelines, with and administrator decisions.

## PROCESSES

Communities typically do not have clearly delineated processes for building consensus or for determining when consensus has been reached. Accordingly, users who operate unilaterally within the consensus process may exert significant influence on discussion outcomes. Also, community discussions also tend to follow an unstructured conversational pattern that is difficult to navigate.

1. Offer guidelines for consensus decision making;
2. Organize complex discussions using argument mapping techniques; and
3. Offer guidelines to encourage behaviors that promote effective collaborative discussions.

## CROSS PROJECT INTERACTION

Current participation in cross-project coordination efforts is limited, and significant barriers prevent immediate implementation of more robust cross-project efforts at resolving behavioral issues.

1. Ensure that ideas relevant to smaller projects are widely disseminated; and
2. Standardize systems and policies across communities.

# INTRODUCTION TO THE WORLD OF WIKIMEDIA

The Wikimedia Foundation “is a nonprofit charitable organization dedicated to encouraging the growth, development and distribution of free, multilingual, educational content, and to providing the full content of these wiki-based projects to the public free of charge. The Wikimedia Foundation operates some of the largest collaboratively edited reference projects in the world, including Wikipedia, a top-ten internet property.”<sup>1</sup> The Wikimedia Foundation is part of a broader access to knowledge movement which aims to provide and protect public sources of knowledge. Wikimedia hosts a variety of projects, such as Wikipedia, Wiktionary, and Wikisource. These projects often exist in a variety of languages, and each language-project combination operates independently.

Each project is home to a contributor community. The term “community” may refer to the entire user base of Wikimedia across projects or may refer to a specific user base which contributes primarily within the intersection of a specific project and language (e.g., Portuguese Wikipedia). A contributor community includes administrators with varying levels of privileges (volunteers with local privileges, including the ability to block users and protect pages), editors (volunteers who contribute content to Wikimedia projects), and readers. Communities operate autonomously, creating their own content, policies, and processes. Communities are free to draw on Wikimedia’s values and policies or to refer to content generated by prominent projects, such as English Wikipedia. Communities generally prefer to make decisions through a deliberative consensus-based process, and the Wikimedia Foundation generally does not intervene in local initiatives.

Community initiatives are often born from a request for comment (“RFC”), which is a single user’s call to action based on that user’s idea or a perceived common observation. The members of the community to which the RFC is posed then reply to the initial post until the body of the discussion and observed community consensus is summarized by a self-volunteered user. This summary then proceeds to be further discussed and refined until yet another summary is proffered. This cycle continues for an indeterminate amount of time. The exact process used to determine consensus is unclear: interviews of long-term users suggest that it relies on common sense. Also unclear are processes used to determine when discussions will be closed, what conclusion is supported by the community’s consensus, and how to make a transition to implementation.

Although the Wikimedia Foundation funds and hosts the above-mentioned projects, the Wikimedia Foundation generally respects the autonomy of the contributor communities. Due to

---

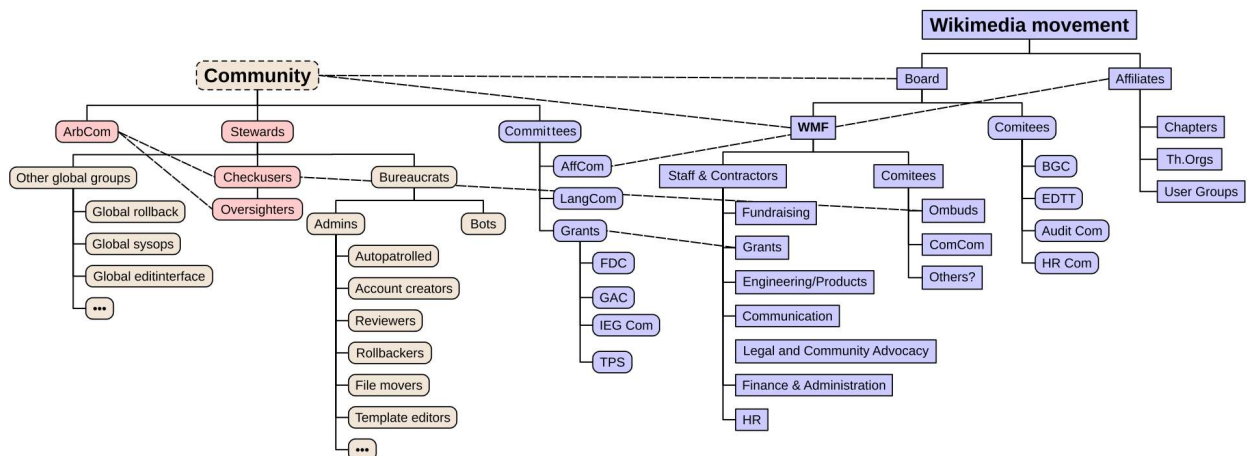
<sup>1</sup> Wikimedia Foundation, <https://wikimediafoundation.org/wiki/Home> (last visited Nov. 26, 2017).



the strong emphasis on free contribution of knowledge by a volunteer community, the contributor community is generally operationally independent from the Wikimedia Foundation. In fact, the Wikimedia Foundation tends to look to community consensus in order to establish pan-Wikimedia policies and values from the bottom up, and contributor communities do not look favorably upon pure top-down decisions and initiatives.

Wikimedia provides a unique platform for interaction between and among the Wikimedia Foundation, the contributor community as a whole, and specific communities. Furthermore, despite a theoretical hierarchy—the Foundation at the top as the owner of the platform, then the Stewards as global administrators, then each project's local administrators and editors at the bottom of the structure—the fact that all users are openly accessible to other users creates a flatter structure in practice, especially in regards to case management for user-reported issues. As a result, administrators generally pick up cases on the basis of their levels of activity and the timeframe of their ongoing work rather than on the basis of purported rank—unless, of course, the case requires access to user rights or powers directly related to a certain position. Indeed, more accessible administrators receive more dispute resolution requests from users.

Below is a chart of Wikimedia's current organizational structure.<sup>2</sup> It does not include the general contributor population, including new, autoconfirmed, or extended autoconfirmed volunteer contributor accounts.



<sup>2</sup> *Wikimedia Power Structure*, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Wikimedia\\_power\\_structure#/media/File:Wikimedia\\_organizational\\_and\\_user\\_rights\\_hierarchy.svg](https://meta.wikimedia.org/wiki/Wikimedia_power_structure#/media/File:Wikimedia_organizational_and_user_rights_hierarchy.svg) (last visited Nov. 26, 2017).

## THE COMMUNITY HEALTH INITIATIVE AND HARASSMENT

The Wikimedia community has acknowledged community health as one of its top concerns.<sup>3</sup> Community health refers to robust and sustained constructive engagement of volunteer contributors in a community.<sup>4</sup> Volunteer contributors include users who have new, autoconfirmed, or extended autoconfirmed accounts. The well-being of community members ultimately contributes to the quality of content and the cohesiveness of the project as a whole.<sup>5</sup> User disputes can be broadly categorized into content, behavioral, and mixed disputes. Content disputes involve debates regarding the substance of articles, such as issues concerning the legitimacy of a source or grammatical corrections.<sup>6</sup> Wikimedia platforms generally have well-developed approaches to resolving pure content disputes. Behavioral disputes involve users engaging in behaviors that violate community norms or project policies, such as vandalism or harassment.<sup>7</sup> Mixed disputes arise when content disputes are associated with the individual editor's personhood and the dispute becomes a behavioral issue.<sup>8</sup>

Behavioral disputes that devolve further to become harassment pose a significant problem for Wikimedia communities. Abusive or disruptive behavior appears to be more widespread on larger projects. According to a survey conducted in early 2017, 73% of Wikipedia volunteers surveyed reported that they had been harassed or bullied on Wikipedia in the previous 12 months.<sup>9</sup> On most smaller projects, less than 10% of users surveyed reported being harassed or

---

<sup>3</sup> *Strategy/Wikimedia movement/2017/Direction*, Wikimedia: Meta-Wiki,

[https://meta.wikimedia.org/wiki/Strategy/Wikimedia\\_movement/2017/Direction](https://meta.wikimedia.org/wiki/Strategy/Wikimedia_movement/2017/Direction) (last visited Nov. 26, 2017)

<sup>4</sup> As determined through interviews with employees from Wikimedia Foundation.

<sup>5</sup> As determined through interviews across stakeholder groups, including the Wikimedia Foundation, stewards, and local administrators.

<sup>6</sup> *Wikipedia: Dispute Resolution*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Dispute\\_resolution#Resolving\\_content\\_disputes](https://en.wikipedia.org/wiki/Wikipedia:Dispute_resolution#Resolving_content_disputes) (last visited Nov. 26, 2017).

<sup>7</sup> *Commons: Administrators' noticeboard/User problems*, Wikimedia Commons, [https://commons.wikimedia.org/wiki/Commons:Administrators%27\\_noticeboard/User\\_problems](https://commons.wikimedia.org/wiki/Commons:Administrators%27_noticeboard/User_problems) (last visited Nov. 26, 2017).

<sup>8</sup> Interviews with stewards and a member of the Arbitration Committee on English Wikipedia clearly distinguished content and behavioral disputes to highlight the more concerning and troublesome nature of a dispute involving racial, gender, religious, or ethnic identities.

<sup>9</sup> *Community Engagement Insights/2016-17 Report/Audience 1/Goal2 Question 126.01*, Wikimedia Meta-Wiki,

[https://meta.wikimedia.org/wiki/Community\\_Engagement\\_Insights/2016-17\\_Report/Audience\\_1/Goal2](https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2016-17_Report/Audience_1/Goal2) (last visited Nov. 26, 2017).

bullied in the previous 12 months.<sup>10</sup> Harassment has a significant negative impact on community health. Users who encounter behavioral issues frequently disengage with Wikimedia projects, at least temporarily. 49% of Wikimedia community members surveyed indicated they avoided Wikimedia projects at least one day in the past month because they felt unsafe or uncomfortable.<sup>11</sup>

There is general agreement among users that current systems for addressing user issues are deficient. Users have expressed that responses to behavioral issues are frequently inadequate.<sup>12</sup> An analysis of English Wikipedia's incidents noticeboard conducted by the Wikimedia Foundation in 2017 found that of 3093 reported cases in the past 12 months, only 1745 had been resolved.<sup>13</sup> Users have expressed broad support for improving existing systems, particularly through better reporting and evaluation tools. Specifically, 84% of 300 users surveyed requested better reporting tools, 77% requested better noticeboards, and 75% requested better wiki policies.<sup>14</sup>

The Wikimedia Foundation's primary purpose in improving approaches to behavioral issues is to improve community health, and the Foundation is currently engaged in an ongoing, three-year Community Health Initiative precisely aimed at addressing such behavioral issues on Wikimedia projects. As part of the Community Health Initiative, Wikimedia's Anti-Harassment Tools team is currently developing several ideas,<sup>15</sup> many of which originated in various Wikimedia

---

<sup>10</sup> Community Engagement Insights/2016-17 Report/Audience 1/Goal2, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Community\\_Engagement\\_Insights/2016-17\\_Report/Audience\\_1/Goal2](https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2016-17_Report/Audience_1/Goal2) (last visited Nov. 26, 2017).

<sup>11</sup> Community Engagement Insights/2016-17 Report/Audience 1/Goal2 Question 115.03, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Community\\_Engagement\\_Insights/2016-17\\_Report/Community\\_Engagement](https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2016-17_Report/Community_Engagement) (last visited Nov. 26, 2017).

<sup>12</sup> Community Engagement Insights/2016-17 Report/Audience 1/Goal2, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Community\\_Engagement\\_Insights/2016-17\\_Report/Audience\\_1/Goal2](https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2016-17_Report/Audience_1/Goal2) (last visited Nov. 26, 2017) ("Of people who reported harassment or attacks to functionaries (bureaucrats, stewards, arbitrators) in the past six months, 54% indicated the response was not at all useful."; "Of people who reported harassment or attacks to other volunteers in the past six months, 48% indicated the response was not at all useful."; "Of the people who reported harassment or attacks to Wikimedia Foundation staff in the past six months, 77% indicated the response was not at all useful."; "Of the people who reported harassment or attacks to chapter or affiliate staff in the past six months, 75% indicated the response was not at all useful").

<sup>13</sup> ANI data: Generate existing reports for 12 full months, Phabricator, <https://phabricator.wikimedia.org/T178536#3801428> (last visited Dec. 4, 2017).

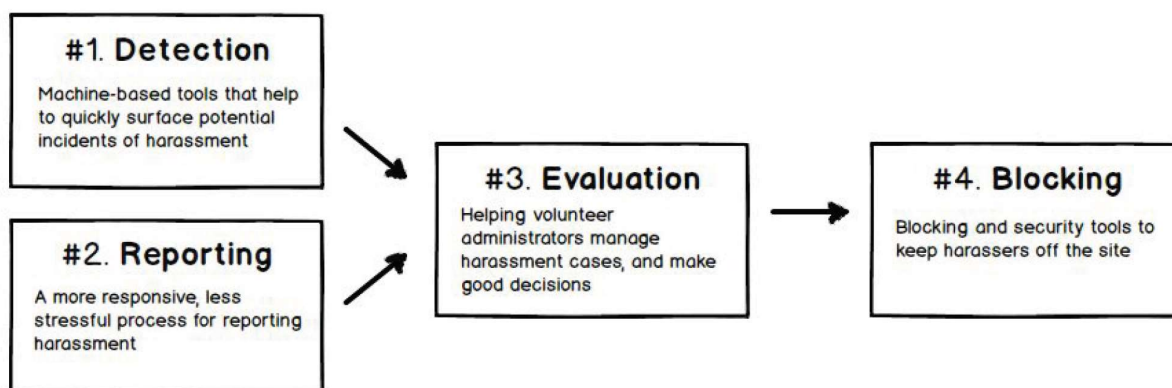
<sup>14</sup> Community Engagement Insights/2016-17 Report/Audience 1/Goal2, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Community\\_Engagement\\_Insights/2016-17\\_Report/Audience\\_1/Goal2](https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2016-17_Report/Audience_1/Goal2) (last visited Nov. 26, 2017).

<sup>15</sup> *File:Wikimedia Foundation grant proposal - Anti-Harassment Tools For Wikimedia Projects - 2017.pdf*, Wikipedia: The Free Encyclopedia,

communities, to combat harassment. Wikimedia also conducts community surveys to better define stakeholder interests and areas of concern.<sup>16</sup>

Wikimedia has categorized its anti-harassment work in four steps:<sup>17</sup>

1. Detection - The discovery of behavioral issues. Usually done by volunteer users.
2. Reporting - Notification of behavioral issues. Usually done by volunteer users.
3. Evaluation - Determining whether a behavioral issue has occurred and what should be done about it. Usually done by volunteer administrators.
4. Blocking - The temporary or indefinite removal of a user from the community or limitation on certain user activities. Usually done by volunteer administrators.



---

[https://en.wikipedia.org/w/index.php?title=File:Wikimedia\\_Foundation\\_grant\\_proposal\\_-\\_Anti-Harassment\\_Tools\\_For\\_Wikimedia\\_Projects\\_-\\_2017.pdf&page=7](https://en.wikipedia.org/w/index.php?title=File:Wikimedia_Foundation_grant_proposal_-_Anti-Harassment_Tools_For_Wikimedia_Projects_-_2017.pdf&page=7) (last visited Nov. 26, 2017).

<sup>16</sup> See, e.g., Dispute Resolution on the English Wikipedia (2012): Survey of active contributors to describe perceptions of and interactions with dispute resolution systems on English Wikipedia. Harassment Survey (2015): Cross-project survey in 16 languages aimed at “learning about ways contributors to Wikimedia projects have experienced harassment and difficult engagements with other users.” Community Consultation (2016): Cross-project survey in multiple languages aimed at learning about community strategic priorities and perceptions of challenges to the Wikimedia movement. English Wikipedia Administrator Confidence Survey (2017): Survey of administrators on English Wikipedia to measure “how admins feel about different kinds of conflict specific activities (wiki-hounding, vandalism, harassment, sockpuppetry), how confident they feel spotting, mitigating, and intervening in these case types, and if they feel supported with tools and other resources from the Wikimedia Foundation.” Community Engagement Insights Report (2017): Broad survey with multiple audiences covering a variety of topics.

<sup>17</sup> *File:Wikimedia Foundation grant proposal - Anti-Harassment Tools For Wikimedia Projects - 2017.pdf*, Wikipedia: The Free Encyclopedia,

[https://en.wikipedia.org/w/index.php?title=File:Wikimedia\\_Foundation\\_grant\\_proposal\\_-\\_Anti-Harassment\\_Tools\\_For\\_Wikimedia\\_Projects\\_-\\_2017.pdf&page=7](https://en.wikipedia.org/w/index.php?title=File:Wikimedia_Foundation_grant_proposal_-_Anti-Harassment_Tools_For_Wikimedia_Projects_-_2017.pdf&page=7) (last visited Nov. 26, 2017).

The Wikimedia Foundation initially engaged HNMCP to evaluate the potential of cross-project dispute resolution systems, and HNMCP's focus expanded to include identified areas of maximum impact (see below) within the broader Community Health Initiative. HNMCP's role in this process involved soliciting information from different stakeholders within Wikimedia communities, compiling such information into findings, and making recommendations based on findings.

# OUR APPROACH

The Harvard Negotiation and Mediation Clinical Program ("HNMCP")<sup>18</sup> at Harvard Law School applies long-standing theories in dispute systems design, negotiation, mediation, and facilitation to deliver conflict management recommendations to its clients.<sup>19</sup> As law students with HNMCP, we approached the problem presented from the perspective of conflict management and found the following questions most critical to determine potential practical solutions.

## QUESTIONS PRESENTED

*What role should the Wikimedia Foundation play to properly address behavioral disputes across its projects to achieve Wikimedia's goal of promoting and protecting community health?*

*How can the Wikimedia Foundation assist the community's engagement with discussions regarding behavioral disputes and does the current model maximize implementation of potential solutions given the high priority afforded to issues related to community health?*

## METHODOLOGY

One major consideration for the project was to ground all our findings and recommendations in stakeholder feedback and relevant research. Accordingly, we consulted existing on-wiki community initiatives and discussions, analyzed data sets, conducted interviews across various stakeholder groups, conducted a focus group, consulted experts in relevant academic fields, and undertook independent academic research. We discuss each in turn.

### ***Extant On-wiki Community Initiatives, Discussions, and Relevant Data***

Our research included observation and analysis of systems, policies, and discussions related to community health and incident reporting and evaluation on Wikimedia projects. We focused primarily on English-language communities as well as larger Spanish-language and Portuguese-language communities. We also conducted a survey of prior work related to community health on Wikimedia projects, including efforts by various communities, the Wikimedia Foundation, and external experts.

---

<sup>18</sup> Harvard Negotiations & Mediation Clinical Program, <http://hnmcp.law.harvard.edu/> (last visited Nov. 26, 2017).

<sup>19</sup> HNCMP clients include U.S. based and international private corporations, non-profit organizations, government agencies, and community groups.

## **Interviews**

Over the course of ten weeks, we conducted interviews with ten stakeholders representing various positions, global regions, and affiliated projects. Among those interviewed are representatives from Wikimedia Foundation, stewards, and users with experience in communities with established dispute resolution and incident reporting systems,<sup>20</sup> and users from project and regional communities without well-established systems.

## **Online Focus Group**

We led a limited, 18-day focus group from November 11, 2017 until November 29, 2017, comprised of volunteers who expressed interest and commitment to a collaborative design-thinking exercise geared toward producing rough prototypes for potential implementable solutions to improve management of harassment cases within Wikimedia projects. Volunteers were recruited by publicizing the focus group initiative on administrator and community noticeboards. Ultimately, five volunteers contributed their time, effort, and insight toward this effort.

The focus group was designed with a dual purpose: (1) to pool the Wikimedia community's knowledge to clearly identify an immediately addressable problem within the context of case management for harassment disputes, address the root cause of that problem, and design potential solutions that will directly address a salient need for maximum impact; and (2) to engage the community in a facilitated and structured approach to design thinking for the purpose of gathering qualitative data about prominent behavioral norms in community discussions.

We limited the scope of the focus group and focused on the development of tools to better manage harassment cases because development in this area was determined to be high-value and within HNMCP's expertise.

Instructions and background information for the focus group are provided in Attachment A. Responses from the community are provided in Attachment B. Community discussion of generated ideas is provided in Attachment C.

## **Academic Research**

We consulted academic literature on the ethnography of Wikipedia, case studies on the development of online dispute resolution systems, the use of wikis as platforms to facilitate online collaboration, prominent behavioral norms and values in good faith collaboration and dispute resolution, and theories on facilitation of design-thinking methodology. Additional guidance from Harvard Law School Professors Christopher Bavitz and Scott Westfahl also informed our approach and final recommendations.

---

<sup>20</sup> English Wikipedia has a standing Arbitration Committee that serves as an informal destination for appeals from claims raised in the noticeboards with unsatisfactory resolutions.

## CHALLENGES

### ***The Online Platform***

Some challenges posed by an online medium are the asynchronous nature of online communication, the need to continuously engage volunteer users, and the limitations of text-based communications such as the absence of voice and tone,<sup>21</sup> facial expressions,<sup>22</sup> and body language.<sup>23</sup> These limitations were most prevalent in the focus group where asking follow-up questions and gathering observational data proved more difficult.

### ***Barriers to Participation***

The following barriers narrowed the cross-section of the Wikimedia community that could be engaged for this study:

#### ***Language***

This research was limited to stakeholders who were comfortable communicating in English. Therefore it was unable to engage with the full breadth of the entire Wikimedia community.

#### ***Visibility***

This research was limited to stakeholders who were able to interact with the Meta-wiki research page. A sample group comprised of these stakeholders is potentially unrepresentative as certain language-project communities of particular concern related to harassment issues may be unaware of the existence of the Wikimedia Foundation or Meta-wiki research pages in general.

#### ***Scope of Wikimedia Community***

The Wikipedia platform alone currently hosts 299 languages.<sup>24</sup> Across all Wikimedia platforms there are 916 wiki communities across the globe.<sup>25</sup>

---

<sup>21</sup> Kate Garklavs, *Voice and Tone as Information Architecture*, 42 *Bulletin of the Association for Information Science and Technology* (June 2016), at 19, 22.

<sup>22</sup> Samantha Baggott, Romina Palermo, and Mark A. Williams, *Involuntary Facial Expression Processing: Extracting information from Two Simultaneously Presented Faces*, *PLoS One* (2011). See also Marie L. Smith & Christina Merlusca, *How task shapes the use of information during facial expression categorizations*, 14 *Emotion* (June 2014) at 478, 487.

<sup>23</sup> A. Milton Jenkins & Randall D. Johnson, *What the Information Analyst Should Know about Body Language*, 1 *MIS Quarterly* (Sep. 1977) at 33, 47.

<sup>24</sup> *Wikipedia*, *Wikipedia: The Free Encyclopedia*, <https://en.wikipedia.org/wiki/Wikipedia> (last visited Nov. 26, 2017).



Combined with the limited time available for community members,<sup>26</sup> stakeholder engagement necessarily was limited to a small segment of the Wikimedia community.

### ***The Nature of HNMCP Involvement***

HNMCP projects are limited to a single academic semester, spanning approximately 12 weeks. This length of time limits the range of available tools and options to conduct a more comprehensive qualitative study of Wikimedia and the contributor community.

## **MAXIMIZING IMPACT**

Taking into account the relative importance of various facets of the problem, the adequacy of current approaches to addressing various facets of the problem, the Wikimedia Foundation's criteria for prioritization,<sup>27</sup> our relevant expertise, and possible synergies with the ongoing work of the Wikimedia Anti-Harassment Tools Team, we identified four issue areas where our team could add significant value:

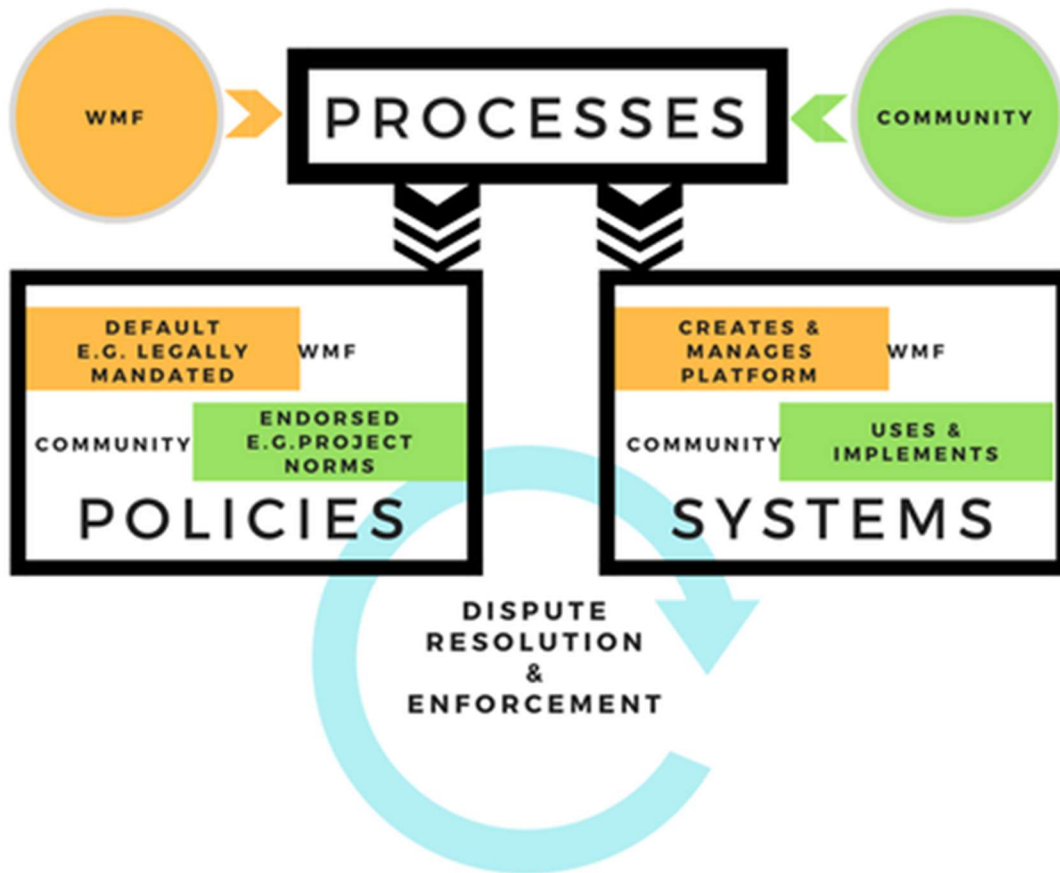
1. Integrated reporting and evaluation **systems**
2. Community administrative **policies**
3. Community discussion and decision-making **processes**
4. **Cross-project** interaction

---

<sup>25</sup> *List of Wikimedia wikis*, Wikimedia Foundation, <https://wikimediafoundation.org/wiki/Special:SiteMatrix> (last visited Nov. 26, 2017).

<sup>26</sup> A steward aptly noted that participation on any Wikimedia platform is at most a hobby.

<sup>27</sup> The Wikimedia Foundation identified four criteria for prioritizing anti-harassment work as the following: 1) Readiness: Ideas that are “designed, defined, and ready for development”; 2) Value: Ideas that “will provide the most value to our users”; 3) Feasibility: Ideas that can be “accomplish[ed] given our time frame and developer capacity”; and 4) Support: Ideas that have “received support from the users who participate in the current workflows” and that “have momentum from people currently affected by harassment.” For more details, see *Community Health Initiative*, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Community\\_health\\_initiative#Prioritization](https://meta.wikimedia.org/wiki/Community_health_initiative#Prioritization) (last visited Nov. 25, 2017).



While significant attention has been devoted to formal dispute resolution bodies and complex delegations of responsibilities, the vast majority of behavioral issues appear to be addressed (or unaddressed) by the community and local administrators. Therefore, our findings and recommendations concentrate primarily on this level of incident resolution.

# FINDINGS

Our research indicates that deficiencies in existing mechanisms and practices likely contribute to the problem of widespread user issues which harm community health. Such deficiencies exist in **systems** for reporting and evaluating incidents involving user behavior and in **policies** which guide administrative action. These problems appear to be at least partially symptomatic of broader shortcomings in community discussion and decision-making **processes**. This section summarizes our findings related to existing systems, policies, and community processes, and concludes with a brief assessment of current cross-project collaboration.

## SYSTEMS

Current systems for reporting, managing, and evaluating user incidents do not appear to be effective at preventing harassment. Disputes that reach administrators are often resolved adequately, but many disputes that could benefit from administrator intervention do not get such attention.

As indicated previously, harassment remains widespread on Wikimedia projects, and users are broadly dissatisfied with responses to reported incidents.<sup>28</sup> However, several participants in our focus group and individuals we interviewed expressed the opinion that report evaluation systems such as noticeboards function fairly well.<sup>29</sup> When explaining the reasons for this opinion, these individuals often stated that administrators typically make good decisions. Notably, this explanation covers only cases which are (1) actually reported and (2) receive adequate administrator attention. Furthermore, this explanation reflects the assessment of only experienced members of the community who are actively engaged in anti-harassment and dispute resolution. Therefore, it is likely that the evaluation aspect of reporting and evaluation systems functions well, at least from the perspective of experienced users. The most acute problems with such systems appear to be related to deficiencies in reporting and deficiencies in case management.

---

<sup>28</sup> Community Engagement Insights/2016-17 Report/Audience 1/Goal2, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Community\\_Engagement\\_Insights/2016-17\\_Report/Audience\\_1/Goal2](https://meta.wikimedia.org/wiki/Community_Engagement_Insights/2016-17_Report/Audience_1/Goal2) (last visited Nov. 26, 2017).

One focus group participant pointed out that data on user dissatisfaction with incident evaluation systems was likely negatively skewed due to the fact that adversarial cases usually involve some losers. However, data on user dissatisfaction with systems typically reflects the experience of reporting users, and many cases do not involve adversarial proceedings between two or more users who are acting in good faith.

<sup>29</sup> Information provided in steward interviews and focus group. See Attachments B and C.

**FINDING 1: USERS UNDERREPORT BEHAVIORAL ISSUES DUE TO A COMMUNITY CULTURE THAT EMPHASIZES DIALOGUE AS WELL AS DIFFICULTIES WITH THE REPORTING PROCESS.**

On many projects, users typically report behavioral incidents on an administrator noticeboard. Larger projects often have several noticeboards, each of which addresses a particular incident category. A comprehensive list of such noticeboards is often listed under the community portal or on a central dashboard.<sup>30</sup> Once a user reports an incident, the incident is typically evaluated by administrators and other volunteers within the community. An administrator may then decide to take an administrative action, such as blocking an offending user.

Since unreported behavioral issues are likely to go unaddressed, consistent reporting is key to reducing harassment and improving community health. A reporting system which does not allow for incidents to be reported in a consistent manner may allow disruptive users to continue such behavior unchecked and may contribute to a victim's sense of powerlessness, possibly leading to user disengagement. Therefore, the inconsistency of reporting within the current system threatens community interests in building and maintaining a safe, productive community environment.

***Community norms appear to decrease the likelihood that a user will report an instance of harassment.***

Wikimedia communities generally embrace openness and dialogue. Users are expected to attempt to resolve disputes through discussion before turning to more formal systems. While these norms are constructive for encouraging collaboration and joint problem solving, they also likely deter users from reporting behavioral issues. Administrative action and formal dispute processes may be in tension with discussion norms. Accordingly, communities implicitly – and sometimes explicitly – discourage the use of such formal processes. For example, an essay on opening a dispute on Wikipedia's administrators' noticeboard/ incidents ("ANI" or "AN/I") begins with the following advice: "Don't. Just... don't. Taking a dispute to ANI is like going to war. War has no victors, only survivors."<sup>31</sup> Submitting a report is often perceived as an adversarial action, and reporting users may be subject to rather harsh treatment. As one user notes, "ANI is not a victim-advocate office. You'll not be treated gently, you'll not be coddled, and you won't be given any special treatment. You and the other editor are on equal footing until consensus has developed. You have the best chance of having a desirable outcome if you can take a little heat."<sup>32</sup> In such an environment, community members often avoid

---

<sup>30</sup> See for example, *Wikipedia: Dispute resolution noticeboard*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Dispute\\_resolution\\_noticeboard](https://en.wikipedia.org/wiki/Wikipedia:Dispute_resolution_noticeboard) (last visited Nov. 25, 2017).

<sup>31</sup> *Wikipedia: ANI advice*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:ANI\\_advice](https://en.wikipedia.org/wiki/Wikipedia:ANI_advice) (last visited Dec.2, 2017).

<sup>32</sup> *Wikipedia: ANI advice*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:ANI\\_advice](https://en.wikipedia.org/wiki/Wikipedia:ANI_advice) (last visited Dec.2, 2017).

reporting and instead respond to abusive behavior in one of three ways that are more in line with community culture: warning, confronting, or not responding.<sup>33</sup>

<b>RESPONSE TYPE 1</b>	
<b>WARNING</b>	
<p>Some users might warn an offending user that their actions violate community policy. This tactic is in line with community norms around discussion, and it avoids unnecessary escalation. Stronger action may then be taken if the offending user continues their abusive behavior. This approach is likely effective in the context of abuse on a single page. However, if the offending user continues their abusive behavior on another page, there will likely be no easily accessible record of prior warnings. Therefore, the community may be slow to identify a pattern of abusive behavior.</p>	
<p>Example 1 "Maleficent"</p>	<p><i>I don't even recall [Maleficent] even turning into a bat or a porcupine, or a bear, or a beetle, or a mountain lion, or a tiger, or even a wolf [...] the only animal that she ever transformed into that I recall is the Dragon. so can someone explain why she is listed under these? – Alice (IP user)<sup>34</sup></i></p> <p style="padding-left: 40px;"><i>Uh, because some dumbass wants to act like a fucking idiot who deserves to get his throaght [sic] slit. I'll change it. – Bob (IP user)</i></p> <p style="padding-left: 40px;"><i>Whoa, watch it. <a href="#">No personal attacks!</a> – Charlie<sup>35</sup></i></p>
<p>Example 2: "US elections"</p>	<p><i>While I agree the popular vote figures should be in the opening paragraph I also think you are a total fucking idiot for thinking the US elections are anything other than Democratic. – Dave (IP user)</i></p> <p style="padding-left: 40px;"><i>A reading of <a href="#">WP:CIVIL</a> is most highly recommended... – Erin (IP user)<sup>36</sup></i></p>

<sup>33</sup> Data from observations of talk page discussions in which one user referred to another user as a "fucking idiot." Terms selected due to frequency of use and consistency of use solely in the context of abusive behavior. See Google search results for: [site:https://en.wikipedia.org/wiki/Talk:fucking\\_idiot](https://en.wikipedia.org/wiki/Talk:fucking_idiot).

<sup>34</sup> All usernames are anonymized throughout.

<sup>35</sup> *Talk:Maleficent*, Wikipedia: The Free Encyclopedia, <https://en.wikipedia.org/wiki/Talk%3AMaleficent> (last visited Nov. 26, 2017).

<sup>36</sup> *Talk: United States presidential election, 2000/Archive 4*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Talk%3AUnited\\_States\\_presidential\\_election,\\_2000%2FArchive\\_4](https://en.wikipedia.org/wiki/Talk%3AUnited_States_presidential_election,_2000%2FArchive_4) (last visited Nov. 26, 2017).

## RESPONSE TYPE 2 CONFRONTING

Other users may confront the offending user. Often, abusive behavior occurs in the context of a heated argument. In these scenarios, other users frequently continue and sometimes escalate the discussion. Users rarely escalate a discussion by answering abusive behavior with an explicitly abusive response. Rather, users tend to engage while technically observing community policies, by pointing out weaknesses in the abusive user's argument or cleverly indicating that the abusive user is not to be taken seriously. While frequently effective at shutting down abusive users, engaging abusive behavior without reporting also increases the difficulty of identifying repeated abusive behavior.

Example 1  
"Dudeette"

*dudeette means that a guy mentaly in love with his butt this is not correct u fucking idiot lr is my belief that dudette is a dying or dead term.Dude is more commonly used as a unisex term.I am editing the page to reflect that.Atleast partially. – Frank*

*You need to provide a [source](#) for that please. – Grace*

Example 2:  
"Both Sides of the Story"

*[Judy], you are a fucken idiot – [...] if the article is one-sided that is because it is a fact that the middle eastern animals embarked on a spate of disproportionate revenge attacks - the article does document what happened - it is just that the factual 'both sides of the story' to whcih you refer is not to your liking [...] – Heidi*

*Quite illustrative that you support this fable without offering any logical reason as to why it should be supported, let alone any concurring account that would suggest it isn't complete bullshit. [...] Fancy buying the Sydney Harbour Bridge on ebay while you're at it? – Judy*

<sup>37</sup>

---

<sup>37</sup> *Talk:2005 Cronulla riots/Archive 4*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Talk%3A2005\\_Cronulla\\_riots%2FArchive\\_4?oldformat=true](https://en.wikipedia.org/wiki/Talk%3A2005_Cronulla_riots%2FArchive_4?oldformat=true) (last visited Dec. 2, 2017).

## RESPONSE TYPE 3 NOT RESPONDING

Other users might not respond to the offending user. In situations where abusive behavior is not met with a response, other users may have made a deliberate choice to not engage, had insufficient interest to engage, or simply failed to notice the abusive behavior. While ignoring abusive behavior may be an effective response to shutting down abusive behavior on a particular page, it similarly fails to contribute to the detection of problematic users.

Example 1  
“Two Years  
Later”

*Because no source is given (bad author!), I cannot do more than speculate, but I very, very strongly encourage you to verify this quote, replace it, or remove it entirely. – Mallory (IP user)*

*You are a complete fucking idiot. The quote is word-for-word from the official translation of the Standard Edition of Freud's writings. See the Five Lectures on Psycho-Analysis(1909); Lecture Three. Moron. – Oscar*

It does not appear that a report was filed for this incident. Oscar had already been reported for abusive behavior in the previous year.<sup>38</sup> Two years later, Oscar was indefinitely blocked for abusive behavior.<sup>39</sup>

Some of the examples above highlight the frequent presence of IP users (unidentified users who are not registered in the community) as both perpetrators and targets of behavioral abuse. This factor adds another layer of complexity. IP users are less likely than registered users to be well integrated in the community, and therefore less likely to report abusive behavior. These users are also harder to sanction, and other users may be less likely to report abusive behavior by IP users because they believe IP users are comparatively unlikely to become a major problem or they believe possible sanctions against IP users would be ineffective.

**Users may have difficulty identifying the correct forum for filing a report in noticeboard systems.**

If a user does decide to report an incident, identifying proper locations for reporting user behavior issues requires navigating multiple complex pages. New users in particular may be unfamiliar with how to report incidents. In some cases, the complexity of identifying

---

<sup>38</sup> *Wikipedia: Administrators' noticeboard/IncidentArchive365*, Wikipedia: The Free Encyclopedia, [https://wikivisually.com/wiki/Wikipedia:Administrators'\\_noticeboard/IncidentArchive365](https://wikivisually.com/wiki/Wikipedia:Administrators'_noticeboard/IncidentArchive365) (last visited Dec. 2, 2017).

<sup>39</sup> *User talk:Pfistermeister*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/User\\_talk:Pfistermeister#ANI\\_Notification](https://en.wikipedia.org/wiki/User_talk:Pfistermeister#ANI_Notification) (last visited Dec. 2, 2017).

the correct noticeboard may deter users from reporting incidents, thereby failing to meet their interest in having accessible dispute resolution mechanisms. In other cases, the complexity of identifying the correct noticeboard may lead users to report incidents on an incorrect noticeboard. Such misplaced reporting may make it more difficult for administrators to identify and address issues efficiently. On larger projects with narrowly defined noticeboard topics, incidents involving multiple categories of issues may be difficult to classify.<sup>40</sup> This difficulty may frustrate users seeking to report incidents. This difficulty may also prompt users to report incidents in multiple locations, decreasing administrator efficiency.

---

<sup>40</sup> *Wikipedia talk:Dispute resolution noticeboard/Archive 1*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia\\_talk:Dispute\\_resolution\\_noticeboard/Archive\\_1](https://en.wikipedia.org/wiki/Wikipedia_talk:Dispute_resolution_noticeboard/Archive_1) (last visited Dec. 2, 2017).



## EXAMPLE

### HELP WITH DISPUTE ON POPE FRANCIS TALK PAGE

In April 2013, English Wikipedia user Pat posted a report on Administrator Noticeboard: Incidents (AN/I) alleging that another user had made inappropriate deletions and comments related to the relationship of popes with the Jewish community.<sup>41</sup> Another user pointed out that since the incident appeared to be more of a content dispute than a behavioral incident, it should be addressed on the Dispute Resolution Noticeboard. Pat then posted the report in Talk: Dispute Resolution, which is intended for discussion of dispute resolution systems.<sup>42</sup> Pat appears to have recognized the mistake and posted again, this time on Talk: Dispute Resolution Noticeboard.<sup>43</sup> After some discussion of the report by users, another user noted that the report was misplaced: "This is the correct place for discussing when to file a DRN case when there is an open ANI case. It is not the right place to discuss whether or not (former) Pope Benedict is/was a Nazi." Finally, Pat found the actual Dispute Resolution Noticeboard.<sup>44</sup> However, shortly after several editors made opening comments on the merits of the dispute, another user closed the discussion on the Dispute Resolution Noticeboard with the following explanation: "AN/I opened by filing editor has not been closed. Please re-file this request after the discussion at that venue has closed."

Thankfully, this story has a happy ending. Pat returned to the original article talk page and agreed to abide by community consensus. Pat even acknowledged that they had let their temper get the better of them and apologized to other users on the talk page.

Pat's experience illustrates the possible difficulty of identifying the correct reporting location as well as the procedural barriers to accessing these systems. While Pat remained engaged throughout the process, it is likely that many users, particularly new users, might become frustrated by the difficulty of identifying the correct reporting location and filing a proper report.

---

<sup>41</sup> *Wikipedia:Administrators' noticeboard/IncidentArchive792*, Wikipedia: The Free Encyclopedia, [http://www.wikiomni.com/pages/Wikipedia:Administrators%27\\_noticeboard/IncidentArchive792](http://www.wikiomni.com/pages/Wikipedia:Administrators%27_noticeboard/IncidentArchive792) (last visited Dec. 2, 2017).

<sup>42</sup> *Wikipedia talk:Dispute resolution/Archive 4*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia\\_talk:Dispute\\_resolution/Archive\\_4#Help\\_with\\_dispute\\_on\\_Pope\\_Francis\\_talk\\_page](https://en.wikipedia.org/wiki/Wikipedia_talk:Dispute_resolution/Archive_4#Help_with_dispute_on_Pope_Francis_talk_page) (last visited Dec. 2, 2017).

<sup>43</sup> *Wikipedia talk:Dispute resolution noticeboard/Archive 12*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia\\_talk:Dispute\\_resolution\\_noticeboard/Archive\\_12#Help\\_with\\_dispute\\_re\\_Pope\\_Francis\\_page](https://en.wikipedia.org/wiki/Wikipedia_talk:Dispute_resolution_noticeboard/Archive_12#Help_with_dispute_re_Pope_Francis_page) (last visited Dec. 2, 2017).

<sup>44</sup> *Wikipedia:Dispute resolution noticeboard/Archive 68*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Dispute\\_resolution\\_noticeboard/Archive\\_68](https://en.wikipedia.org/wiki/Wikipedia:Dispute_resolution_noticeboard/Archive_68) (last visited Dec. 2, 2017).

**Once the proper location is identified, users may face additional procedural and technical barriers to submitting a proper report and substantiating evidence.**

Additionally, users may face technical barriers to providing the evidence necessary to evaluate cases. As noted by the Wikimedia Foundation's Anti-Harassment Tools team, "Currently the burden of proof is on the victim to prove their own innocence and the harasser's fault."<sup>45</sup> To meet this burden of proof, users are often required to provide diffs as evidence of abusive or disruptive behavior. A diff is "a permanent link to a single edit or set of edits, showing the difference between the before and after versions."<sup>46</sup> While creating a diff need not be difficult,<sup>47</sup> it is also not an intuitive process, and it may present a challenge for inexperienced users. It is possible that some users may choose to not report behavioral issues due to this technical barrier. The Anti-Harassment Tools team is currently developing tools to reduce technical barriers to reporting.

**FINDING 2: NOTICEBOARD SYSTEMS DO NOT FACILITATE EFFECTIVE CASE MANAGEMENT.**

Even if administrators generally make good decisions, a system is likely to be ineffective if administrators frequently fail to address reported cases. When cases go unaddressed, disruptive activities are more likely to continue, threatening community interests in a safe, productive community environment. According to an analysis of English Wikipedia's AN/I by the Wikimedia Foundation, approximately 44% are either closed or automatically archived without being resolved.<sup>48</sup> While it is possible that this shortcoming is due to a lack of volunteers, this explanation seems unlikely; the English Wikipedia community is generally active on AN/I. According to the Foundation's analysis, 322 administrators participated on AN/I in the past 12 months. If not a lack of volunteers, it is possible the issue is that some cases simply do not merit attention. While this may explain some cases which appear to fall through the cracks, it seems unlikely that close to 44% of reported cases do not merit attention. Another possible explanation is that the method of managing cases on AN/I is suboptimal. This explanation appears to be supported by the statements of administrators responding to the Wikimedia Foundation's Administrator Confidence Survey. According to one administrator, "AN/I is useless. Cases are auto-archived without being closed."<sup>49</sup> Another administrator added, "There is a tendency

---

<sup>45</sup> *Wikipedia:Community health initiative on English Wikipedia*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Community\\_health\\_initiative\\_on\\_English\\_Wikipedia](https://en.wikipedia.org/wiki/Wikipedia:Community_health_initiative_on_English_Wikipedia) (last visited Dec. 2, 2017).

<sup>46</sup> *Wikipedia:Complete diff and link guide*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Complete\\_diff\\_and\\_link\\_guide](https://en.wikipedia.org/wiki/Wikipedia:Complete_diff_and_link_guide) (last visited Dec. 2, 2017).

<sup>47</sup> *Wikipedia:Simplest diff guide*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Simplest\\_diff\\_guide](https://en.wikipedia.org/wiki/Wikipedia:Simplest_diff_guide) (last visited Dec. 2, 2017).

<sup>48</sup> *ANI data: Generate existing reports for 12 full months*, Phabricator, <https://phabricator.wikimedia.org/T178536#3801428> (last visited Dec. 4, 2017).

<sup>49</sup> *Wikipedia:Community health initiative on English Wikipedia/Administrator confidence survey/Results*, Wikipedia: The Free Encyclopedia,

towards mob rule with different outcomes happening depending on who joins the discussion.”<sup>50</sup> Indeed, noticeboards possess several shortcomings which may decrease administrator efficiency and cause cases to go unaddressed.

- First, when an incident is reported on a noticeboard, there is no assurance that an administrator will review and address the incidents in a given amount of time. This problem is particularly acute in cases involving difficult judgment calls, such as many cases which involve possible harassment. For both convenience and practical necessity, many noticeboards are designed to automatically archive reports after a given amount of time. Therefore, on some noticeboards, reports that are not addressed expediently are not likely to be addressed at all.
- Second, on large projects with multiple noticeboards, it is difficult to identify which noticeboards require urgent attention and which noticeboards may be generally underserved. Dashboards which identify backlogs are helpful. However, such tools only address problems after they occur. Furthermore, on noticeboards where users do not always close discussions or mark issues as resolved, it may be impossible to determine whether the noticeboard is underserved or whether administrators have decided to not address certain reports.
- Third, noticeboards provide no indication whether an administrator is monitoring the discussion on a certain report or whether an issue would benefit from the attention of several administrators. As a result, administrators may duplicate efforts on lengthy yet easy to address reports, while potentially difficult issues may receive insufficient input from administrators and other members of the community. Even in situations where an administrator is clearly monitoring a relatively easy case and no other administrative attention is necessary, other administrators may have to read through an extensive discussion in order to make this determination.
- Fourth, on most projects, there is little formal recourse for continuing to pursue reports that have been inadequately or incorrectly addressed by administrators. While arbitration committees on larger projects may sometimes play this role, this approach is likely not universally applicable. Portuguese Wikipedia and Spanish Wikipedia both decided to eliminate their arbitration committees, and for most smaller projects, establishing an arbitration committee is not a practical solution. Even on projects with arbitration committees, these committees are able to hear relatively few cases. Therefore, once a

---

[https://en.wikipedia.org/wiki/Wikipedia:Community\\_health\\_initiative\\_on\\_English\\_Wikipedia/Administrator\\_confidence\\_survey/Results#Comments\\_about\\_policy.2C\\_reporting.2C\\_harassment.2C\\_community\\_culture](https://en.wikipedia.org/wiki/Wikipedia:Community_health_initiative_on_English_Wikipedia/Administrator_confidence_survey/Results#Comments_about_policy.2C_reporting.2C_harassment.2C_community_culture) (last visited Dec. 4, 2017).

<sup>50</sup> *Wikipedia:Community health initiative on English Wikipedia/Administrator confidence survey/Results*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Community\\_health\\_initiative\\_on\\_English\\_Wikipedia/Administrator\\_confidence\\_survey/Results#Comments\\_about\\_policy.2C\\_reporting.2C\\_harassment.2C\\_community\\_culture](https://en.wikipedia.org/wiki/Wikipedia:Community_health_initiative_on_English_Wikipedia/Administrator_confidence_survey/Results#Comments_about_policy.2C_reporting.2C_harassment.2C_community_culture) (last visited Dec. 4, 2017).

case is closed or archived it may be difficult to request further review without violating community policies or norms.

- Fifth, most noticeboard formats do not allow for the collection and analysis of potentially useful user history information. For administrators evaluating reports, it would likely be helpful to view whether a user has filed other reports and whether a user has been the subject of other reports, both within that project and across Wikimedia communities. However, most noticeboards do not have clearly defined input fields that would allow for the easy collection of such data. As a result, administrators and other community members may lack the information necessary to address the report, increasing the likelihood that the report will go unaddressed.

***Deficiencies in reporting and case management systems may create incentives for users that decrease the efficiency of the system as a whole.***

The complexity and ineffectiveness of noticeboards may incentivize users to develop inefficient alternatives to the noticeboard system. For example, instead of posting a report on a noticeboard, a community member may reach out to administrators directly. While targeting specific administrators may sometimes provide more immediate attention, the selected administrators may not have the time, background knowledge, or desire to address the incident. Users are more likely to reach out to an administrator with whom they are familiar or an administrator who has been active recently. Often, this may not be the administrator who is best suited to address the incident, and quite frequently, selected administrators are already among the most heavily occupied administrators in the community. Moreover, by reaching out to an administrator with whom a reporting user is familiar, the reporting user may create a real or apparent conflict of interest for the administrator.<sup>51</sup> This practice may result in less accountability for disruptive users, creating a less safe and productive community environment.

The complexity of the noticeboard system may also incentivize users to develop inefficient uses of the system. On larger projects with multiple noticeboards, users may report incidents on multiple forums in an effort to achieve a more favorable resolution, even when they are already aware of the correct forum location.<sup>52</sup> Such forum shopping wastes administrator time, as multiple administrators may have to address the same issue on different noticeboards. Conflicting decisions on different noticeboards may frustrate users and undermine the legitimacy of the administrator evaluation process. Forum shopping may also constitute a form of harassment in itself by forcing a user to repeatedly defend legitimate actions. These inefficiencies undermine community interests in efficient administrative action on important matters as well as community interests in maintaining the legitimacy of administrator action.

---

<sup>51</sup> Information provided in steward interviews and focus group. See Attachments B and C.

<sup>52</sup> For an example of alleged forum shopping, see: *Wikipedia talk:Noticeboards*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia\\_talk:Noticeboards#Editor\\_modifying\\_closed\\_discussion.2C\\_when\\_clearly\\_states\\_not\\_to](https://en.wikipedia.org/wiki/Wikipedia_talk:Noticeboards#Editor_modifying_closed_discussion.2C_when_clearly_states_not_to) (last visited Dec. 2, 2017).

## POLICIES AND GUIDELINES

Communities often create standards which define proper procedures and proper conduct. Unlike most pages which may be edited by any user, most pages describing official community standards may only be changed by community consensus. These standards form the basis for administrative action, and members of a community often cite to them during discussions on behavioral incidents.

English Wikipedia recognizes three types of standards: policies, guidelines and essays.<sup>53</sup>

- “Policies have wide acceptance among editors and describe standards that all users should normally follow.”
- “Guidelines are sets of best practices that are supported by consensus. Editors should attempt to follow guidelines, though they are best treated with common sense, and occasional exceptions may apply.”
- “Essays are the opinion or advice of an editor or group of editors for which widespread consensus has not been established. They do not speak for the entire community and may be created and written without approval.”

This section will focus primarily on policies and guidelines related to administrative action and abusive or disruptive behavior.

### **FINDING 1: STANDARDS RELATED TO BLOCKING AND OTHER FORMS OF ADMINISTRATIVE ENFORCEMENT ARE OFTEN VAGUE.**

While communities generally do have policies and guidelines on administrative enforcement, these standards are frequently vague. Some communities do have more developed standards which identify categories of actions that may result in blocking. However, these standards frequently do not clearly define what specific actions would result in blocking. Additionally, blocking policies often fail to distinguish between traditional blocks and less drastic and less frequently used measures such as topic blocks and interaction blocks. Policy vagueness and community norms typically grant administrators significant discretion.

---

<sup>53</sup> *Wikipedia:Policies and guidelines*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Policies\\_and\\_guidelines#Role](https://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines#Role) (last visited Dec. 2, 2017).

## EXAMPLE

### ENGLISH WIKTIONARY'S BLOCKING POLICY

English Wiktionary's blocking policy is deliberately short and vague, so much so that it offers clarification that the policy provided is in fact the actual policy and not merely a summary.

*This policy page consists of two sections: (a) policy, the statement of binding policy; (b) explanation, a non-binding explanation of the policy, and guidelines showing how the policy is usually applied. The only binding section is "policy". The section "policy" is not merely the policy in a nutshell but rather the complete statement of the policy itself.*

#### POLICY

*The blocking policy itself is as follows:*

- 1. The block tool should only be used to prevent edits that will, directly or indirectly, hinder or harm the progress of the English Wiktionary.*
- 2. It should not be used unless less drastic means of stopping these edits are, by the assessment of the blocking administrator, highly unlikely to succeed.*

#### EXPLANATION

*There is deliberately no hard-and-fast rule about what is considered to hinder or harm our progress. Clear examples of such behaviour include:*

- Deliberately harming our content by deleting useful things or adding useless content or pages.*
- Persistently wasting other editors' time by making many edits that have to be undone, cleaned up, or otherwise modified to make them correct.*

*Causing our editors distress by directly insulting them or by being continually impolite towards them.<sup>54</sup>*

This lack of policy detail is not inherently problematic. Some administrative discretion is likely beneficial. It is impractical to codify appropriate responses to every situation, and administrative discretion allows concise, easily understood standards to be applied to a wide range of foreseeable and unforeseeable scenarios.

However, in communities that face acute problems with abusive behavior and that lack systems for resolving disputes consistently, broad administrative discretion may lead to inconsistent decisions, reduced trust in administrators, and extensive time spent evaluating individual cases. This appears to be the case on Portuguese Wikipedia. While

---

<sup>54</sup> *Wiktionary:Blocking policy*, Wiktionary: The Free Dictionary, [https://en.wiktionary.org/wiki/Wiktionary:Blocking\\_policy](https://en.wiktionary.org/wiki/Wiktionary:Blocking_policy) (last visited Dec. 2, 2017).

Portuguese Wikipedia policies do identify recommended block lengths for various categories of actions, such as abuse of public space and improper conduct, these categories leave much discretion to administrators on defining what constitutes a violation.<sup>55</sup> Moreover, pages that might provide more specific guidance are not only similarly vague, they also lack the status of policy: the Disruptive Behavior page is a recommendation,<sup>56</sup> the Civility page is a proposal,<sup>57</sup> and the Harassment page is an essay.<sup>58</sup> One probable result of the administrative discretion allowed on Portuguese Wikipedia is that members of the community view administrators as independent decision makers exercising *their own discretion* rather than individuals enforcing *community consensus*. In practice, this dynamic does appear to influence administrative decisions on behavioral incidents. Administrators are frequently hesitant to take action against abusive behavior by established users due to fear of personal harassment and the possibility of losing standing in the community.<sup>59</sup>

## **FINDING 2: STANDARDS FREQUENTLY SET A HIGH BAR FOR ADMINISTRATIVE ACTION.**

Policies and guidelines often counsel administrators to tread carefully when considering blocks, and on many projects, there is a presumption against blocking. English Wikivoyage, for example, classifies user blocks under the heading “Last Resort.”<sup>60</sup> English Wikipedia’s blocking policy offers another paradigm case. “As a rule of thumb, when in doubt, do not block; instead, consult other administrators for advice. After placing a potentially controversial block, it is a good idea to make a note of the block at the administrators’ incidents noticeboard for peer review.”<sup>61</sup>

---

<sup>55</sup> *Wikipédia:Política de bloqueio*, Wikipédia: A Enciclopédia Livre, [https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Pol%C3%ADtica\\_de\\_bloqueio](https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Pol%C3%ADtica_de_bloqueio) (last visited Dec. 2, 2017).

<sup>56</sup> *Wikipédia:Comportamento desestabilizador*, Wikipédia: A Enciclopédia Livre, [https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Comportamento\\_desestabilizador#Quebra\\_repetida\\_de\\_normas\\_comportamentais](https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Comportamento_desestabilizador#Quebra_repetida_de_normas_comportamentais) (last visited Dec. 2, 2017).

<sup>57</sup> *Wikipédia:Civilidade*, Wikipédia: A Enciclopédia Livre, <https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Civilidade#Exemplos> (last visited Dec. 2, 2017).

<sup>58</sup> *Wikipédia:Assédio moral*, Wikipédia: A Enciclopédia Livre, [https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Ass%C3%A9dio\\_moral#Bloqueio\\_por\\_ass.C3.A9dio](https://pt.wikipedia.org/wiki/Wikip%C3%A9dia:Ass%C3%A9dio_moral#Bloqueio_por_ass.C3.A9dio) (last visited Dec. 2, 2017).

<sup>59</sup> Information provided in steward interviews and focus group. See Attachments B and C.

<sup>60</sup> *Wikivoyage:How to handle unwanted edits*, Wikivoyage, [https://en.wikivoyage.org/wiki/Wikivoyage:How\\_to\\_handle\\_unwanted\\_edits#Last\\_resorts](https://en.wikivoyage.org/wiki/Wikivoyage:How_to_handle_unwanted_edits#Last_resorts) (last visited Dec. 2, 2017).

<sup>61</sup> *Wikipedia:Blocking policy*, Wikipedia: The Free Encyclopedia [https://en.wikipedia.org/wiki/Wikipedia:Blocking\\_policy#Common\\_rationales\\_for\\_blocks](https://en.wikipedia.org/wiki/Wikipedia:Blocking_policy#Common_rationales_for_blocks) (last visited Dec. 2, 2017).

Policies offer various rationales for this caution. Some policies note that newcomers may be less familiar with community policies. For example, English Wikibook's blocking policy states, "Administrators should take special care when dealing with new users. Beginning editors are often unfamiliar with Wikibooks policy and convention, and so their behavior may initially appear to be disruptive. Responding to these new users with excessive force can discourage them from editing in the future."<sup>62</sup> Other policies, such as English Wiktionary's blocking policy, seem wary of disrespecting established members of the community. "It is rare, but occasionally there will be a seasoned contributor, even an administrator, who is causing trouble; such cases must be handled with diplomacy. It is not acceptable to block a whitelisted user or an administrator unless they already know they will be blocked for their actions. In most cases they will not know they will be blocked unless they have received an explicit warning or are deliberately and maliciously ignoring current practice."<sup>63</sup>

In practice, administrators appear to follow this guidance by taking action only in the most straightforward cases. As one administrator noted, "Too often the person perpetrating harassment is adhering to the letter of the law so admins are reluctant to sanction them because what they're [doing] is not over the bright line."<sup>64</sup>

The high standard for administrative action adopted by most projects represents a tradeoff. On one hand, administrators are less likely to frustrate and possibly drive away users who are the subject of reports. In this way, the standard protects community interests in retaining certain users – particularly newcomers who may need time to acculturate and experienced users who contribute significantly to the substance of the projects. On the other hand, users experiencing harassment may have difficulty in getting an administrator to intervene in a timely and effective manner. Here, a lack of action may frustrate and drive away reporting users and also permit abusive and disruptive behavior to continue. There are reasons communities may or may not want to make this tradeoff, perhaps based on the level of harassment within the community. However, it is unclear whether communities typically make deliberate, informed choices about this tradeoff or whether communities simply tend to follow models that have been adopted on other projects.

---

<sup>62</sup> *Wikibooks:Blocking policy*, Wikibooks, [https://en.wikibooks.org/wiki/Wikibooks:Blocking\\_policy](https://en.wikibooks.org/wiki/Wikibooks:Blocking_policy) (last visited Dec. 2, 2017).

<sup>63</sup> *Wiktionary:Blocking policy*, Wiktionary: The Free Dictionary, [https://en.wiktionary.org/wiki/Wiktionary:Blocking\\_policy](https://en.wiktionary.org/wiki/Wiktionary:Blocking_policy) (last visited Dec. 2, 2017).

<sup>64</sup> *Wikipedia:Community health initiative on English Wikipedia/Administrator confidence survey/Results*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Community\\_health\\_initiative\\_on\\_English\\_Wikipedia/Administrator\\_confidence\\_survey/Results#Comments\\_about\\_policy.2C\\_reporting.2C\\_harassment.2C\\_community\\_culture](https://en.wikipedia.org/wiki/Wikipedia:Community_health_initiative_on_English_Wikipedia/Administrator_confidence_survey/Results#Comments_about_policy.2C_reporting.2C_harassment.2C_community_culture) (last visited Dec. 4, 2017).



### **FINDING 3: COMMUNITIES TYPICALLY DO NOT HAVE A SYSTEMATIC METHOD FOR CONVERTING PRECEDENT INTO STANDARDS.**

Precedent appears to operate as a persuasive though not binding authority on Wikimedia projects.<sup>65</sup> Some communities have made some effort to collect and organize past decisions, including decisions on behavioral incidents, but these efforts occur on a relatively small-scale.<sup>66</sup> Nevertheless, precedent does appear to have a central role in the formation of many standards. For example, instead of describing which offenses merit administrative action, English Wikipedia's blocking policy lists common rationales for blocks with the apparent assumption that administrators will generally act in line with previous community decisions.<sup>67</sup>

The uncertain status of precedent as well as the absence of a system for organizing precedent exacerbate potential problems with administrator discretion. A lack of clear precedent may result in inconsistent decisions, reduced trust in administrators, and greater inefficiency in evaluating individual cases

---

<sup>65</sup> *Wikipedia:Practical process*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Practical\\_process](https://en.wikipedia.org/wiki/Wikipedia:Practical_process) (last visited Dec. 2, 2017);

*Wikipedia:Consensus*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Consensus#Consensus\\_can\\_change](https://en.wikipedia.org/wiki/Wikipedia:Consensus#Consensus_can_change) (last visited Dec. 2, 2017).

<sup>66</sup> See, e.g., *Wikipedia:Precedents*, Wikipedia: The Free Encyclopedia, <https://en.wikipedia.org/wiki/Wikipedia:Precedents> (last visited Dec. 2, 2017); *Wikipedia:Arbitration/Index/Principles*, Wikipedia: The Free Encyclopedia, <https://en.wikipedia.org/wiki/Wikipedia:Arbitration/Index/Principles> (last visited Dec. 2, 2017).

<sup>67</sup> *Wikipedia:Blocking policy*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Blocking\\_policy#Common\\_rationales\\_for\\_blocks](https://en.wikipedia.org/wiki/Wikipedia:Blocking_policy#Common_rationales_for_blocks) (last visited Dec. 2, 2017).

## EXAMPLE OUTING

In 2008, a user posted on English Wikipedia's AN/I accusing user Sybil of "outing" by posting the IP address of a registered user. Various members of the community commented with different opinions. "Unless it's someone with Checkuser abusing their rights, the posting of an IP address is not outing. [...] It's a pretty childish and perhaps an uncivil thing to do, but I don't believe it's Outing whatsoever." Another community member disagreed. "Sure looks to me like a case of harassment and attempted outing, well worthy of a block." An administrator then analyzed the situation, concluding, "Whether or not Sybil's actions constitute a violation of the letter of policy, posting someone's putative IP address where there is no relevance to the discussion at hand is certainly a violation of the spirit of the harassment policy. I have now cautioned Sybil that we take this sort of issue very seriously, and warned that he will be blocked if he does it again."<sup>68</sup>

In 2011, user Trent accused another user of "outing" by posting Trent's IP address. Again, community members commented with differing opinions. "Linking a user to an IP address without there having been a public link made by that user is considered to be 'outing'." However, another community member argued that Trent's use of unregistered editing was to blame. "Trent outed himself. There is no 'outing' issue here." An administrator closed the discussion, noting that "there's nothing to left to be done." At no point did anyone refer back to the 2008 discussion. While the 2008 discussion would not have resolved all of the issues here, it may have provided clarity on aspects of the case. However, because case outcomes are not categorized or routinely noted in policies or guidelines, it is not surprising that contributors here were unaware of the 2008 discussion.<sup>69</sup>

---

<sup>68</sup> *Wikipedia:Administrators' noticeboard/IncidentArchive501*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Administrators%27\\_noticeboard/IncidentArchive501#Requesting\\_block](https://en.wikipedia.org/wiki/Wikipedia:Administrators%27_noticeboard/IncidentArchive501#Requesting_block) (last visited Dec. 2, 2017).

<sup>69</sup> *Wikipedia:Administrators' noticeboard/IncidentArchive684*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Administrators%27\\_noticeboard/IncidentArchive684#Berean\\_Hunter\\_harrassment](https://en.wikipedia.org/wiki/Wikipedia:Administrators%27_noticeboard/IncidentArchive684#Berean_Hunter_harrassment) (last visited Dec. 2, 2017).

## PROCESSES

Wikimedia communities have underlying decision-making and implementation processes which ultimately create systems and policies. As such, difficulties encountered in these processes are partially responsible for deficiencies in systems and policies.

### **FINDING 1: COMMUNITIES TYPICALLY DO NOT HAVE CLEARLY DELINEATED PROCESSES FOR BUILDING CONSENSUS OR FOR DETERMINING WHEN CONSENSUS HAS BEEN REACHED.**

Although Wikimedia communities generally prefer consensus-based decision-making which often operates as a central principle in community discussions, the concept of consensus as applied to these communities is rarely well-defined. The English Wikipedia:Consensus page, for example, offers one sentence on how consensus ought to be determined: "Consensus is ascertained by the quality of the arguments given on the various sides of an issue, as viewed through the lens of Wikipedia policy."<sup>70</sup> This guidance, however, appears to describe an objective strength decision-making criterion, where positions are evaluated based on their quality, rather than a consensus decision-making criterion, where positions are evaluated based on their level of support. Furthermore, although the Wikipedia:Consensus page identifies instances which may require a greater level of consensus relative to typical decisions, the page provides no guidance on how users should determine whether that required level of consensus exists in such circumstances.

The lack of clearly delineated processes and criteria for determining consensus seems to encourage behaviors against stated policies. For example, while English Wikipedia policies advise that polling is not a substitute for discussion-based consensus,<sup>71</sup> in practice consensus frequently appears to be determined by an undefined supermajority: for example, one English Wiktionary policy passed with 75% support<sup>72</sup> while another English Wiktionary policy failed with 63% support.<sup>73</sup> Additionally, despite Wikimedia's general preference for consensus-based decisions, in both cases, a single user unilaterally assumed authority to determine whether there was a consensus. The deciding user cited

---

<sup>70</sup> *Wikipedia: Consensus*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Consensus#Determining\\_consensus](https://en.wikipedia.org/wiki/Wikipedia:Consensus#Determining_consensus) (last visited Dec. 2, 2017).

<sup>71</sup> *Wikipedia:Polling is not a substitute for discussion*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Polling\\_is\\_not\\_a\\_substitute\\_for\\_discussion](https://en.wikipedia.org/wiki/Wikipedia:Polling_is_not_a_substitute_for_discussion) (last visited Dec. 4, 2017).

<sup>72</sup> *Wiktionary:Votes/pl-2010-01/New blocking policy*, Wiktionary: The Free Dictionary, [https://en.wiktionary.org/wiki/Wiktionary:Votes/pl-2010-01/New\\_blocking\\_policy](https://en.wiktionary.org/wiki/Wiktionary:Votes/pl-2010-01/New_blocking_policy) (last visited Dec. 4, 2017).

<sup>73</sup> *Wiktionary:Votes/pl-2009-08/Add en: to English topical categories*, Wiktionary: The Free Dictionary, [https://en.wiktionary.org/wiki/Wiktionary:Votes/pl-2009-08/Add\\_en:\\_to\\_English\\_topical\\_categories](https://en.wiktionary.org/wiki/Wiktionary:Votes/pl-2009-08/Add_en:_to_English_topical_categories) (last visited Dec. 4, 2017).

no information outside the vote count to support the consensus determination yet community members involved in both discussions did not raise issues with the deciding user's determination.

**FINDING 2: USERS WHO OPERATE UNILATERALLY WITHIN THE CONSENSUS-BUILDING PROCESS MAY EXERT MORE INFLUENCE ON DISCUSSION OUTCOMES THAN USERS WHO FOCUS ON BUILDING CONSENSUS AND WILL NOT TAKE ACTION IN THE ABSENCE OF CONSENSUS.**

In practice, the emphasis on consensus-based decision-making tends to entrench the status quo. Since the accepted course of action in the absence of a consensus is generally to do nothing, existing systems, policies, and behaviors tend to continue in the absence of a consensus for change. As described above, it is often difficult to know *ex ante* what is required to achieve consensus on a given issue. Accordingly, users who try to build consensus in a discussion may not have any way of knowing when consensus has been achieved, let alone what action steps to take.

In some cases, however, systems, policies, and behaviors need not be instituted through formal consensus-based decision making in order to be approved by the community. In practice, users who unilaterally take initiative within the consensus process may exert more influence on discussion outcomes than users who strive to respect the spirit of the consensus-building process. For an example, see "Example: Sexual Harassment" on page 37.

### **FINDING 3: EXTENDED COMMUNITY DISCUSSIONS TEND TO FOLLOW AN UNSTRUCTURED CONVERSATIONAL PATTERN THAT IS DIFFICULT TO NAVIGATE.**

Community discussions also suffer from technical problems. Pages which host extended discussions tend to be especially poorly organized and difficult to navigate. Discussions tend to follow an unstructured conversational pattern, and it is typically difficult to quickly identify central arguments or understand how complex arguments interact.

#### **EXAMPLE MP4 VIDEO**

In 2014, the Wikimedia Foundation opened a Request for Comment on Commons regarding whether the Wikimedia Foundation should support the MP4 video file format. Compared to most community discussions, the conversation was well-organized. Summarized arguments for and against the proposal were prominently placed at the top of the page. The community could then vote on a range of options, including various partial support options. Near the bottom of the page, community members could suggest alternative proposals or comment on other relevant issues.

However, it was difficult to identify which arguments were most persuasive to the community without reading the entire conversation to pick out repeated reasons. The necessity of repeating arguments to demonstrate support for a certain rationale added significantly to the length of the conversation. Also, community members frequently broke off into smaller conversations on individual votes, many of which were redundant.

While this conversation method is effective in deciding among multiple options, it is also inefficient. In the present case, this inefficiency appeared to decrease the visibility of potentially important arguments and options. Additionally, both the inefficiency of the process as well as the evaluative mindset of most contributors likely prevented the community from generating and implementing creative solutions to the underlying problem.<sup>74</sup>

These general process problems are pervasive across projects. Furthermore, unique shortcomings evident in discussions on systems, discussions on policies, and discussions on individual incident reports merit particular attention.

#### **DISCUSSION PROCESSES ON SYSTEMS**

System discussions involve designing and implementing systems and system components to address behavioral issues. Wikimedia Foundation and individual communities have

---

<sup>74</sup> *Commons:Requests for comment/MP4 Video*, Wikimedia Commons, [https://commons.wikimedia.org/wiki/Commons:Requests\\_for\\_comment/MP4\\_Video](https://commons.wikimedia.org/wiki/Commons:Requests_for_comment/MP4_Video) (last visited Dec. 4, 2017).

devoted significant effort to initiating and developing anti-harassment tools. Although implementation of technical anti-harassment proposals, such as the development of AbuseFilter improvements and user history tools, have seen gradual and effective progress, implementation of proposals requiring broad community action appears ineffective and generally nonexistent.<sup>75</sup>

The lack of effective or clear processes for transitioning the community from the ideation of solutions to the implementation of such solutions may be an underlying cause. The sheer number of community interactions and the asynchronous nature of online discussions, combined with the lack of clear guidelines about engaging in collaborative community efforts, foster unending discussions.

Much of the observed community behavior within these discussions runs counter to recommendations in collaboration,<sup>76</sup> design thinking, and dispute resolution theories. First, the number of collaborators is amorphous and constantly changing.<sup>77</sup> Users who initiated the discussion may not necessarily continue to engage as the discussion evolves.<sup>78</sup> As a result, discussions may be dominated by a more involved group of users before reaching a conclusion. Second, discussions are not structured to allow for thorough research about the needs of the broader community.<sup>79</sup> Initiatives are often met with a flood of suggested solutions despite not having identified the underlying causes of the problem. The solutions risk addressing symptoms rather than underlying causes, which may in turn result in wasted resources for minimal impact. Third, these solutions are often immediately evaluated, risking tangential discussions<sup>80</sup> and discouraging potential contributors who are met with sudden harsh criticism.<sup>81</sup>

---

<sup>75</sup> See, for example, proposals to establish a global arbitration committee:

*Requests for comment/Global arbitration and dispute resolution*, Wikimedia Meta-Wiki, [https://meta.wikimedia.org/wiki/Requests\\_for\\_comment/Global\\_arbitration\\_and\\_dispute\\_resolution](https://meta.wikimedia.org/wiki/Requests_for_comment/Global_arbitration_and_dispute_resolution) (last visited Dec. 4, 2017); *Global arbitration committee*, Wikimedia Meta-Wiki [https://meta.wikimedia.org/wiki/Global\\_arbitration\\_committee](https://meta.wikimedia.org/wiki/Global_arbitration_committee) (last visited Dec. 4, 2017).

<sup>76</sup> See Lawrence Susskind, *Breaking Robert's Rules*, 22 *Negotiation Journal* (July 2006), at 351, 355 (describing a procedural shift from a parliamentary process to a consensus-based decision making process). See also Lawrence Susskind et al., *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement* (1999).

<sup>77</sup> Lawrence Susskind et al., *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement* (1999), at 169, 197.

<sup>78</sup> *Id.* (emphasizing importance of including the vast and silent majority because the loudest critics do not allow for accurate assessment of the popularity of an idea).

<sup>79</sup> Lawrence Susskind et al., *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement* (1999), at 99, 135. See also *id.* at 241, 272.

<sup>80</sup> "Often, it seemed that people were reacting to the mere suggestion of change itself -- the community had managed itself in a certain way for almost thirty years, and some members

Perhaps the most obvious gap between prevalent behaviors in Wikimedian interactions and recommended behaviors in the above-mentioned literature may be the lack of a facilitator,<sup>82</sup> as evinced by the voluntary nature of proffered discussion summaries. As mentioned previously, the transient nature of user participation in community discussions further complicates focused discussions and steady progress as the ebb and flow of different groups of users can derail conversations and disperse productive energy. An example of wasted energy would be user engagement with naysayers during the brainstorming/ideation phase, where best practice is to suspend evaluation or judgment to prevent defensive behaviors that do not add to the value of the discussion as a whole or generate other ideas. In such cases, the community would greatly benefit from an appointed facilitator who can guide engaged community members through a structured process and create comprehensive summaries that highlight the most relevant points of discussion.

All of the above-mentioned norms of Wikimedia community discussions and consensus-building processes are further exacerbated by the lack of a clear timeframe in community discussions.

---

seemed to feel that any proposed change threatened the well-being of a crucial organization in their lives." *Id.*

<sup>81</sup> Lawrence Susskind et al., *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement* (1999), at 305, 320.

<sup>82</sup> *Id.* at 199, 238.

## **DISCUSSION PROCESSES ON POLICIES AND GUIDELINES**

Policy discussions involve drawing distinctions between acceptable and unacceptable user behavior. Since such discussions are often inherently controversial, communities tend to reach consensus only on broad concepts. Therefore, the consensus-based nature of policy discussions is likely a primary cause of policy vagueness.

### **EXAMPLE**

#### **SEXUAL HARASSMENT —Common sense policies are often not so common**

Even comparatively uncontroversial policy decisions appear difficult to discuss and implement. For example, users involved in an extended English Wikipedia discussion regarding an explicit prohibition against sexual harassment were unable to reach consensus on a specific course of action. A resolution was finally achieved only when a user unilaterally edited the harassment policy to read: "Harassment against any editor because of a belief or perception regarding race, color, national origin, ancestry, gender, religion, religious practice, age, disability, or sexual orientation is not allowed anywhere on Wikipedia." Other users agreed to keep the edit in place with minor changes. However, it was unclear whether the final resolution reflected the consensus of all previous participants in the conversation. Of the dozens of users who participated in the discussion, only seven participated in the poll on retaining the unilateral edits, and only five supported the final resolution. Interestingly, despite its flaws, this discussion was described by participants as being among the more constructive attempts at consensus-based deliberation. As the user who closed the discussion summed up, "In all honesty, and to paint a utopian view, this is one of those days where consensus just works."<sup>83</sup>

## **DISCUSSION PROCESSES ON BEHAVIORAL INCIDENTS**

Discussions related to particular incidents tend to be shorter and involve fewer users than discussions related to systems or policies. On many projects, administrators are generally free to determine a resolution to incident evaluation discussions. However, on some projects, certain incident evaluation scenarios may require a community consensus before action is taken. This consensus requirement for incident evaluation presents an exceptionally high standard for administrative action. So long as a user who violates policy is supported by a few other users, it is unlikely that users will reach a consensus. Therefore, a consensus requirement may effectively allow an influential user to violate policies with no repercussions. On Portuguese Wikipedia, for example, a community

---

<sup>83</sup> *Wikipedia talk:Harassment/Archive 5*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia\\_talk:Harassment/Archive\\_5](https://en.wikipedia.org/wiki/Wikipedia_talk:Harassment/Archive_5) (last visited Dec. 4, 2017).



consensus is often required in order to take administrative action against an established user. These users are often able to solicit support off-wiki, enabling them to evade administrative action. This consensus requirement is likely a partial cause of the unusually high incidence of harassment on Portuguese Wikipedia.

Due to their typical brevity, incident evaluation discussions are less susceptible to organizational problems. However, incident evaluation discussions involving multiple issues, complex problems, or long-winded participants may become lengthy. When such cases do become extended, it is difficult to bring organizational structure due to the adversarial nature of incident evaluation discussions as well as the possibility of high emotions. Lengthy, poorly organized incident evaluation discussions may have immediate consequences, as administrators may be less likely to engage in discussions that are difficult and time-consuming to follow.

## **CROSS-PROJECT INTERACTION**

Cross-project interaction within dispute resolution occurs in two tiers: cross-project learning about dispute resolution systems and policies; and discussions to resolve disputes that run across multiple projects. Accordingly, cross-project interaction among communities is vital for ensuring proper dissemination of ideas, inclusivity in discussions that are important to Wikimedia as a whole, and effective methods to address users who violate behavioral policies on multiple projects.

However, user participation in cross-project coordination efforts is often minimal or nonexistent despite the existence of the Meta-wiki platform. Many communities are not connected to the broader movement on Meta, either due to a lack of awareness, lack of interest, or language barriers.<sup>84</sup> Major systems, policies, tools, and processes are often formed within larger communities and may eventually trickle down to smaller communities, but these smaller communities are rarely involved in the design and development of these initiatives. While smaller communities often do adopt the end products of initiatives on larger communities, adoption is uneven and often slow. There is also little collaboration on cross-project behavioral issues. User blocks are typically administered locally, with no notification provided to other communities.

There are three main barriers to greater cross-project collaboration. First, communities are often unaware of initiatives on other projects or, in some cases, even the existence of other projects. There is often little local information available on connecting with other communities. Second, communities are separated by language. Information on platforms for cross-project discussion, such as Meta, is generally presented in English. Third, communities often develop strong interests in their autonomy and unique community culture. As such, they may prefer to seek solutions to problems within the community.

---

<sup>84</sup> Information provided in steward interviews and focus group. See Attachments B and C.

# RECOMMENDATIONS

## SYSTEMS

### ***Centralize and increase the visibility of the reporting system.***

Communities have a strong interest in ensuring that it is simple and straightforward to report incidents of abusive behavior. This objective has already achieved significant support within various communities. Indeed, the Anti-Harassment Tools team plans to create a new reporting system which “doesn't place the burden of proof on or further alienate victims of harassment.”<sup>85</sup>

A reporting system may achieve this goal by prioritizing **visibility** and **functionality**. First, communities should consider placing a reporting link in a highly visible central location. This idea has been proposed previously by users, with the idea sometimes referred to as the “big red button.”<sup>86</sup> Second, communities should consider connecting this link to a multi-step intake form. In this formulation, the form would likely replace noticeboards as the location for reporting incidents. While the precise nature of the form should be determined in consultation with communities, the following fields may be helpful:

- **Classification of the nature of the incident.** Classification would allow administrators to quickly identify and address the types of problems with which they are most comfortable. Classification may serve the function currently served by a multiple noticeboard system.
- **Classification of the importance of the incident.** In our findings, we noted that disruptive or abusive behavior may continue because users do not find certain incidents worth reporting to administrators. Classifying incidents by importance might allow a user to flag an issue without initiating a full incident evaluation and discussion process. Instead, the report could later be used as evidence of repeated problems should the behavior continue. This feature would likely further community interests in quickly identifying users with a pattern of disruptive or abusive behavior.
- **Identification of involved users.** Notifications could then be sent automatically, reducing the burden on reporting users.
- **Assistance with creating diffs or other forms of evidence.** Users unfamiliar with diffs and other features would thus face fewer obstacles to effective reporting.
- **Automatic reporting user identification, or in extreme cases, anonymity.** In some extreme cases, reporting users may fear further harassment. In these cases, communities may wish to consider the possibility of allowing users to report anonymously, possibly with their identity revealed only to administrators.

---

<sup>85</sup> *Wikipedia:Community health initiative*, Wikipedia: The Free Encyclopedia, [https://en.wikipedia.org/wiki/Wikipedia:Community\\_health\\_initiative#Reporting](https://en.wikipedia.org/wiki/Wikipedia:Community_health_initiative#Reporting) (last visited Dec. 4, 2017).

<sup>86</sup> Focus group.

- **Captchas or other methods of user verification to filter out spam.**

A centralized reporting system with a multistep intake form would address many of the deficiencies in the current reporting system, meeting stakeholder interests in simple, clear access to reporting. Users would no longer have to navigate multiple pages to identify the correct reporting location. For incidents involving multiple issues, users could simply select multiple classification options. A single reporting link would also remove the possibility of forum shopping.

Increased visibility of reporting systems would likely lead to a significant increase in reported incidents. While this is likely a positive development insofar as it meets stakeholder interests in simplified access, it may present risks to stakeholder interests in an efficient, well-managed process. First, disruptive users may file many meritless reports. This problem could be addressed by requiring users to complete an additional step certifying the legitimacy of the report and submitting additional information. Communities might also choose to limit reporting to registered users. These steps would likely deter at least some disruptive reporters. This problem might also be addressed by blocking users who submit intentionally disruptive reports. Second, volunteers may not be able to promptly address new incident reports. This problem might be addressed by efforts to streamline evaluation or efforts to recruit additional volunteers. The best approach will likely depend heavily on the specific community context.

### ***Replace the noticeboard system with a single centralized evaluation system.***

Communities have a strong interest in ensuring that reported cases are addressed promptly and adequately and that difficult cases do not fall through the cracks. In order to improve case management practices, communities should consider moving the incident evaluation process from noticeboards to a single centralized location. This centralized location should allow administrators and other volunteers to quickly determine where their efforts are most needed and where the system may not be functioning optimally. While a centralized system could be constructed on a wiki platform, other platforms may provide greater functionality.

A Phabricator-like interface might be one possible platform option. For illustrative purposes, possible aspects of this option will be identified below.

Reported cases might be listed as blocks in three columns. Filed reports would automatically be classified in the first column. When an administrator decided to address a given report, the administrator would move the case from the first column to the second column. Moving a case to the second column would indicate only that an administrator was following the development of the report with the intent to address the report. It would not preclude other administrator or community member activity on the case. When the case was closed by an administrator, the administrator would move the case to the third column. In the third column, community members could continue to discuss the case as well as express support or opposition for the outcome. Community opposition to a case outcome might draw attention from other administrators to review the initial decision. After a certain period of inactivity in the third column, a case would be automatically archived.

In addition to displaying basic information on the case block, cases could be easily classified and sorted. Users could select which categories of case they would like to view by checking boxes in a dropdown menu. Categories of cases would roughly correspond to current noticeboard categories. The default setting would display all cases. In each column, cases could be sort by a range of criteria, including by date reported and by date of last activity. Cases could also be sorted by tags. Communities could develop tags as needed, with possible tags including difficult cases requiring additional community attention and cases possibly requiring attention from bureaucrats or the Wikimedia Foundation.

Such a system would likely address many of the deficiencies in the current noticeboard-based evaluation system, meeting stakeholder interests in the timely resolution of reported cases. Administrators could easily identify cases which require attention and which cases likely do not require attention. By only archiving resolved cases and allowing cases to be sorted oldest to newest, the system would help ensure that all reports would be addressed. While users might still misclassify reports, other users could easily correct classification mistakes by editing a tag. The system also provides for a community-based appeals method that relies on existing deliberative and administrative processes for resolution. The intake forms also collect and organize data which could be used to identify users who have frequently violated community policies as well as users who frequently file meritless reports.

## **POLICIES AND GUIDELINES**

### ***Supplement policies with more specific guidelines informed by community practice.***

While community policy discussions are frequently characterized by substantial disagreement, it is probable that there is more agreement on policies in practice.<sup>87</sup> For example, while community members might disagree whether sexual harassment should be mentioned in community policies, community members are more likely to agree that the community should respond in a certain way to a particular instance of sexual harassment.

By identifying underlying practical consensus where it exists, communities could state with greater specificity how they expect administrators to act in various situations. This development would likely increase the consistency of administrative decisions and bolster community trust in the competence and objectivity of administrators. It would also decrease the amount of recurring interpretive work by administrators. Moreover, the identification of a practical consensus would not threaten community interests in administrative discretion as any guidelines based on a recognized practical consensus need not have the status of policy.

Determining whether an underlying practical consensus exists on a given issues could be accomplished in two ways. First, community members could analyze administrative decisions to identify area of agreement or common interests. By analyzing actual decisions, community

---

<sup>87</sup> See Roger Fisher and William Ury, *Getting to Yes* at 42-57 (1981) (discussing situations in which differing positions may be reconciled by recognition of common interests).

members would likely pick out the types of issues that occur most frequently and are therefore likely the most important to the community. Second, community members could organize a community discussion around a real or constructed case to identify areas of community agreement and disagreement. This approach is likely less time-consuming for individual volunteers, and it incorporates the views of community members outside of administrators. Both approaches have unique strengths and would likely function best if used together. For illustrative purposes, a possible example of this option is as follows:

After reading through many cases of reported harassment, a community member might notice that instances where a user makes a few disparaging remarks about another user's edits are relatively common. In the page Wikipedia talk: Harassment, the community member might identify a particular case where a user made a few disparaging remarks about another user's edits and an administrator determined that a short interaction block was appropriate. The community member might draw out salient features from the case and ask how cases with such features ought to be addressed. Other members of the community would then comment on whether the administrators made an appropriate decision as well as brainstorm possible alternatives. If there is a consensus, that consensus position could be incorporated as a guideline under the harassment policy.

Guidelines constructed using these approaches might take the format of legal restatements, secondary sources which restate accepted practice. Restatements typically offer a concise explanation of the accepted practice. They also often include simplified examples that illustrate applications of the accepted practice in common cases.

***Better integrate policies and guidelines with administrative decisions.***

Incorporating decision examples into policies and guidelines is a constructive step, but it is important that policies and guidelines inform decisions and remain responsive to new developments.<sup>88</sup> To this end, it may be constructive to display relevant policies at the beginning of an incident evaluation discussion. For example, when an incident is categorized by a reporting user, links to or a summary of relevant policies could be provided along with other central information near the top of the discussion. This practice would focus discussion around an objective criterion and possibly remove pressure on individuals to find relevant policy and deter individuals from complicating conversations by copying and pasting large, possibly irrelevant sections of policy into the discussion.

Following the resolution of a report, administrators and other community members should also be able to provide feedback on whether existing policies were adequate to address the issue. This practice would allow the community to identify and address important policy gaps.

---

<sup>88</sup> See Roger Fisher and William Ury, *Getting to Yes* at 42-57 (1981), at 82, 95.

## PROCESSES

### ***Offer guidelines for consensus decision making***

To further community interests in productive discussions with actionable outcomes, it would likely be beneficial to establish standards and processes for consensus decision making prior to a discussion.<sup>89</sup> These standards would most likely take into consideration the percentage of participants in favor, the presence of strong minority views, and the importance of the proposal under discussion. Processes would take into consideration how to summarize proposals, close discussions, and identify action steps. Here, as with policies and guidelines, it would be worth exploring whether widespread agreement exists on what constitutes a consensus in various situations. Discussions aimed at clarifying the definition of a consensus in various contexts could follow a similar methodology, as well, by asking the community to evaluate whether a consensus exists in several example cases.

### ***Organize complex discussions using argument mapping techniques***

In order to keep discussions comprehensible and emphasize the most important arguments, community discussions should utilize argument mapping techniques. One possible system for a discussion on the merits of a single policy option might include the following components:<sup>90</sup>

- A central thesis
- Arguments which directly support or oppose the central thesis
- Nested arguments which directly support or oppose first-level arguments
- Voting options which demonstrate support for or opposition to the thesis and supporting or opposing arguments individual arguments

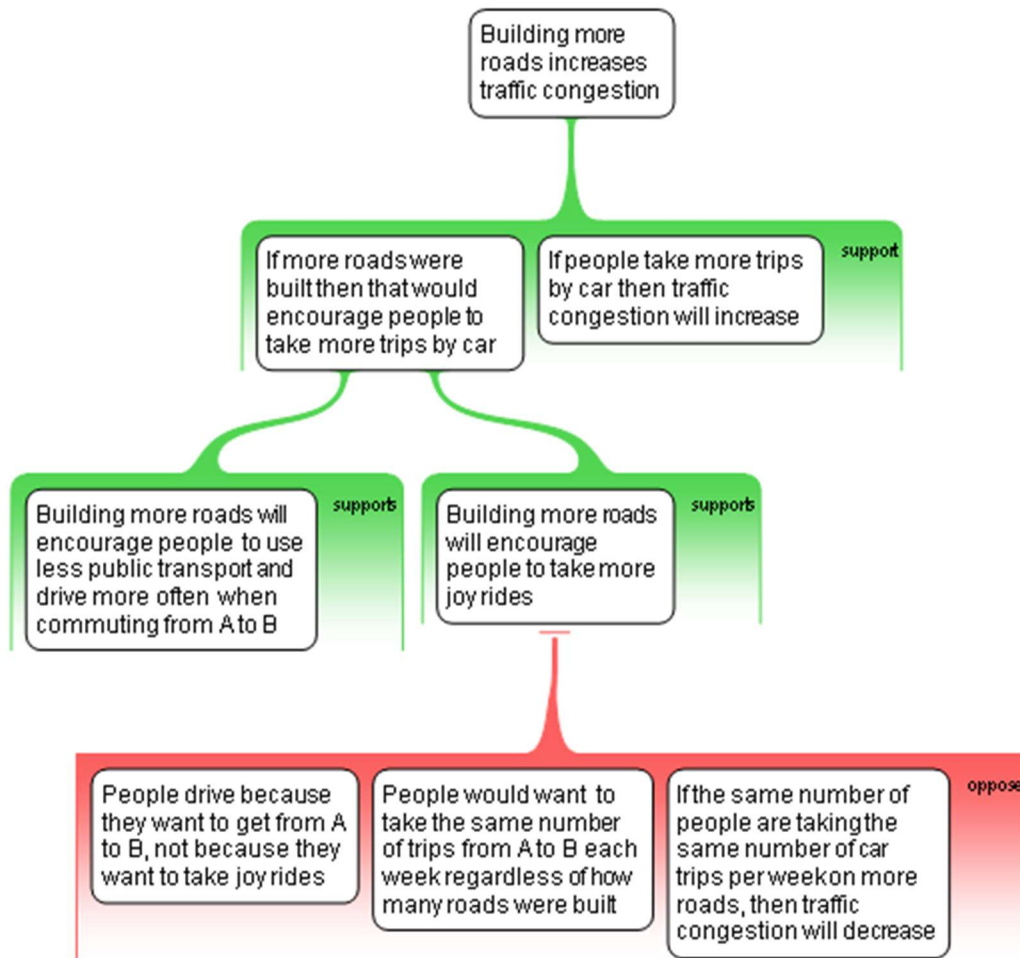
A simple discussion containing a central thesis and nested arguments might conform to the following model.<sup>91</sup>

---

<sup>89</sup> See generally Susskind et al., *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement* (2009).

<sup>90</sup> Peter Sbarski, et al., *Visualizing Argument Structure*. in: Bebis G. et al. (eds) *Advances in Visual Computing*. ISVC 2008. Lecture Notes in Computer Science, vol 5358. See also <https://www.kialo.com/> for an example of such an argument mapping system.

<sup>91</sup> Floris Bex, *Argument Mapping and Storytelling in Criminal Cases*, Legal Information Institute, <https://blog.law.cornell.edu/voxpath/2010/04/16/argument-mapping-and-storytelling-in-criminal-cases/> (last visited Dec. 2, 2017).



By increasing the clarity of arguments, an argument mapping system would likely encourage more user engagement and higher quality user engagement. An argument mapping system would also help identify areas of consensus and reasons for possible disagreement. Furthermore, the premise-conclusion framework implicitly discourages off-topic statements and personal attacks which do not fit naturally into the framework. If users do add unconstructive comments to the discussion, voting options allow these comments to be efficiently relegated to a lower position. This style of argument mapping could also be easily expressed on the standard wiki platform.

**Offer guidelines to encourage behaviors that promote effective collaborative discussions**

Wikis have received positive attention as platforms for collaborative design. James West<sup>92</sup> and Margaret West<sup>93</sup> have recognized the potential of wikis as a platform for online collaboration

---

<sup>92</sup> James West is an associate professor in the Department of Instructional Design and Technology at Western Illinois University. Prior to joining the faculty, he was senior consultant for

and learning. However, they also emphasized the need for planning and effective facilitation to maximize the utility of wikis as a collaborative opportunity.<sup>94</sup>

Collaborative discussions are more than mere group chats. As explored in the focus group,<sup>95</sup> collaboration requires guiding principles and a set of behavioral norms that encourage sharing ideas and thoroughly exploring presented information. In other words, participants must agree to uphold values that create a “highly accessible, shared, participatory environment.”<sup>96</sup> Many of the recommended ground rules for collaborative thinking overlap with recommended guidelines in dispute resolution systems design: some prevalent examples include suspending judgment and encouraging learning conversations.<sup>97</sup>

This recommendation directly addresses issues touched upon in “**Discussion Processes on Systems**”. In a nutshell, due to the nature of online interactions, clear process guidelines for community collaborations become especially important when attempting to generate and refine ideas and carry them through to implementation. Namely, we recommend the following:

1. Elect discussion facilitators — As mentioned throughout, online discussions have unique challenges (e.g. asynchronous nature, the lack of usual information communicated through tone, body language, etc.) that require a lead to guide the conversation. Notably, individuals in this role are not enforcers or instructors but rather encouragers. In other words, they ask follow up questions to access information that may have otherwise

---

Performance Systems Design Corp., where he designed training, distance learning, knowledge management, and curriculum development for companies such as Sears, Unext, and Monsanto. For more information, see James A. West and Margaret L. West, *Using Wikis for Online Collaboration: The Power of the Read-Write Web*, Jossey-Bass A. Wiley Imprint, 1e. 2009.

<sup>93</sup> Margaret West is an instructional technology systems manager in the Center for Innovation in Teaching and Research at Western Illinois University. She was a senior consultant for Performance Systems Design Corp., where she designed distance learning and training programs for companies such as Sears, Diamond Technology, and Motorola. She has over twenty years of experience as an instructional designer and consultant. For more information, see *id.*

<sup>94</sup> James A. West and Margaret L. West, *Using Wikis for Online Collaboration: The Power of the Read-Write Web* (2009).

<sup>95</sup> See Attachments A, B, and C.

<sup>96</sup> James A. West and Margaret L. West, *Using Wikis for Online Collaboration: The Power of the Read-Write Web* (2009), at 22.

<sup>97</sup> See e.g. Nancy Rogers et al., *Designing Systems and Processes for Managing Disputes* at 132-135; Roger Fisher & Daniel Shapiro, *Beyond Reason: Using Emotions As You Negotiate* (2005); and Stone et al., *Difficult Conversations: How to Discuss What Matters Most* (2nd ed. 2010) (discussing recommended behavioral norms and guidelines for dispute systems design and negotiation). See also Gotheif & Seiden, *Lean UX: Designing Great Products with Agile Teams* (2nd ed. 2016) (discussing recommended behavioral norms and guidelines for collaborative design thinking).



been ignored due to subtle personal biases.<sup>98</sup> They may also gently ask quieter participants for more information by offering support and positivity.

2. Have guidelines — Guidelines should be tailored to help participants suspend judgment and to help encourage learning conversations. The focus group features many recommended rules gathered from collaborative design thinking and dispute systems design literature.

3. Set the tone for a collaborative mindset — Participants should be expected to fully engage with guidelines. This may take the form of a simple agreement form to solidify commitment to the process.

The case for implementation for the above-listed recommendations lies in Wikimedia's preference for the development of solutions tailored to individual communities. Promoting behavioral norms that will in turn aid volunteers in staying on the topic of discussion will not only weaken the barrier to entry--the potential sea of relevant or irrelevant text nestled within a discussion topic--it will enable community members to better identify underlying causes of symptoms. Enabling each of Wikimedia's diverse communities to clearly identify their underlying needs will create more impactful solutions in the long run.

A potential limitation to the implementation of this recommendation is Wikimedia's bottom-up policy structure. The contributor community, especially users from larger projects that regularly exercise its autonomy, may be resistant to employing an alternate method of community discussion. Although the sample of interviewed Wikimedians unanimously found the current consensus-building process frustrating, whether the community as a whole may be convinced to adopt recommendations based on consensus and collaboration literature is unclear.

## **CROSS-PROJECT INTERACTION**

It is impractical to establish a cross-project dispute resolution or anti-harassment system from the top down, and the challenges presented by efforts at cross-project coordination would likely prevent the immediate implementation of any such system. Absent improved translation technology, language barriers will likely prohibit sustained collaborations among many communities. Nevertheless, there are steps communities may take to encourage cross-project collaboration which may allow for the organic development of cross-project systems. This process may eventually meet stakeholder interests in a more efficient case management system, the proliferation of best practices, and special help for smaller communities facing acute problems.

First, volunteers should ensure that ideas relevant to smaller projects are disseminated to those projects. All volunteers have a role to play in this process. Users working on development initiatives should consider scalability and the applicability of the initiative beyond the local

---

<sup>98</sup> Lawrence Susskind et al., *The Consensus Building Handbook: A Comprehensive Guide to Reaching Agreement* (1999), at 219.

community. Users interacting on platforms such as Meta should actively reach out to smaller communities regarding the development of systems, tools, policies, processes that they might find particularly helpful. Multilingual users should consider translating information related to important developments on other projects. To the extent possible, users on smaller projects should attempt to engage with the broader Wikimedia community.

Second, communities might consider adopting systems and policies which are identical to those of other communities. System and policy standardization across communities allows for the possibility of cross-project cooperation on matters involving those systems and policies. For example, if all communities in a language group share the same anti-harassment policies and evaluation system, a community experiencing acute problems with harassment could more easily request assistance from less occupied administrators within a different community. Communities could also more efficiently identify and take action to address users who engage in abusive or disruptive behavior across multiple projects.