

Heterogeneous Practices in Collective Governance

Sohyeon Hwang and Aaron Shaw, Northwestern University

1. INTRODUCTION

Self-governing organizations and communities have produced some of the most compelling examples of collective intelligence [Benkler et al. 2015; Woolley et al. 2015]. Within these communities, however, governance practices and procedures such as rule enforcement are more structured than might seem obvious given the absence of formal bureaucratic hierarchy [Forte et al. 2009; Joyce and Butler 2012]. In addition, the scale of many such communities requires rules that are *localized* by sub-communities [Chandrasekharan et al. 2018; Fiesler et al. 2018]. Such localized rules are key to successful self-governance [Ostrom 1990], but have attracted little empirical attention in prior research.

We describe below exploratory work examining the prevalence of localized, heterogeneous governance across communities on a major peer production platform: Wikipedia. Wikipedia remains one of the most striking examples of collective intelligence today [Benkler et al. 2015], particularly in its governance system. In ongoing work, we find that the top five language editions of Wikipedia show high variance in rule pages and organization, despite their operation within the same overarching project, using the same software, and pursuing the same goals of creating high quality encyclopedic content. These preliminary findings illustrate how multi-community analyses of collective governance can provide insight into the prevalence, role, and dynamics of *heterogeneous* practices of the governance that coordinates and manages collective intelligence at scale.

2. RELATED WORK

The idea that distributed, often conflicting, opinions about rules develop into consensus has seen success in peer production projects [Black et al. 2011; Matei and Dobrescu 2010; Osman 2013] as well as in online community content moderation [Lampe and Resnick 2004]. Given this bottom-up process, rule localization provides a fundamental mechanism by which governance operates.

However, conflict about *rules* may undermine regulatory effectiveness [Kiesler et al. 2012], reduce platform accountability [Chandrasekharan et al. 2018], and lend to power plays by more senior, experienced users [Kriplean et al. 2007; Halfaker et al. 2013; Matei and Dobrescu 2010]. This remains an issue for many self-governing communities, within which there may be sub-groups with conflicting localized sets of rules, such as what counts as harassment [Chandrasekharan et al. 2018]. No systematic analysis has considered rule convergence or divergence among Wikipedia’s self-governing language edition communities.

2.1 Collective governance work on Wikipedia

Wikipedia has collectively managed and governed mass participation on a very large scale and over a longer time period than almost any comparable organization [Kriplean et al. 2007; Kittur et al. 2007; Forte et al. 2009; Black et al. 2011]. The design of Wikipedia’s governance as participatory (any contributor can edit the rule pages) and decentralized (any contributor can enforce rules) makes Wikipedia particularly appropriate for investigating how parallel communities each collectively develop and practice rules. Wikipedia continuously deals with governance development and interpretation by contributors [Keegan and Fiesler 2017], creating a gap between rule text and practice.

Prior research has predominantly focused on English Wikipedia. However, there are increasing calls to diversify work, as users exhibit different dynamics across language editions [Hara et al. 2010; Hecht and Gergle 2010]. Moreover, “parallel” communities of Wikipedia — language editions — generate significantly different outputs, both in articles and article content [Hecht and Gergle 2010]. Even when comparing an article to its cross-language-edition equivalents, they are often not organized in the same manner relative to other articles within the respective language edition [Hecht and Gergle 2010].

3. SETTING, DATA, METHODS

We consider a language edition of Wikipedia a distinct sub-community of editors on the broader platform. We look at the five language editions of Wikipedia with the most active editors (more than 5 edits in the last month): English, German, Spanish, French, and Japanese. As the most active language editions, these groups have more people likely to know about and use rules and more developed rule pages. Crucially, a feature called “interlanguage links” (ILLs) [Wikipedia 2020a] connects pages considered equivalent across language editions. Because language editions share several rules, one can compare how different but essentially parallel communities navigate enforcement of the same rules.

Wikipedia “rules” are divided into policies, guidelines, and essays in addition to a set of core principles [Wikipedia 2020b]. As policies and guidelines both have enforceable consequences (e.g. banning) if violated, we consider only them and core principles under the umbrella term “rules”, unless otherwise noted. In German Wikipedia, rules are organized under one category *Richtlinie* instead of the three sub-categories; we count only rule pages linked from the *Wikipedia:Richtlinie* page or *Richtlinie* navigation box, or with the *Richtlinie* navigation box on its page. This is a conservative collection of rule pages on German Wikipedia. We use a Python script to systematically get all relevant rule pages.

We first use the ILLs to highlight divergences in written rules across language editions of Wikipedia, particularly focusing on core shared rules. We later turn to policy invocation — the practice of invoking rules with direct links — on Wikipedia, which allows us to see not only how the enforced norms are rooted in written rules, but also what written rules are being actively enforced. Through the combination of policy invocation and ILLs, we can extract what rules are common across many language edition communities as well as how the communities write about, discuss, and enforce them.

4. HETEROGENEOUS GOVERNANCE PRACTICES ON WIKIPEDIA

4.1 Divergence in rule pages and organization

Across the five language editions, we collected a total of 743 rule pages (Feb 2020). Of those 743 pages, 29.7% were linked to the other four language editions, 14.3% to three, 11.6% to two, 12.8% to one, and 33.0% to none. Table 1 shows the percentage of rule pages in a language edition (row) that have an ILL to a page in one of the other four language editions (column). The highest percentage is 81.5% and the average is roughly 53.1%.

These percentages all drop if we ask whether a rule has an ILL to a *rule* page in one of the other four language editions (bracketed percentages in Table 1). This drop can be attributed to the fact that rule organization, even of shared core rules, across editing communities considerably varies. For example, a core principle “Ignore All Rules” in English, Spanish, and French Wikipedias is not an official rule in Japanese Wikipedia, but one still in the proposal stage. What is considered a “hard” or more important rule in one language edition may be considered a flexible or unimportant one in another. While this variation in organization may be unsurprising, the lack of cross-community overlap in rule pages indicates that there is no clear one-to-one mapping of rules across language editions, even as they share the same goal of peer producing encyclopedic content and operate on the same technical system. Instead, collective governance is highly localized.

	de	en	es	fr	ja
de	-	64.3% [45.2%]	45.2% [26.0%]	41.4% [28.8%]	45.2% [30.1%]
en	31.7% [10.3%]	-	31.1% [17.2%]	34.1% [19.3%]	35.5% [20.8%]
es	58.0% [20.0%]	78.0% [57.0%]	-	65.0% [39.0%]	59.0% [46.0%]
fr	41.7% [16.5%]	72.4% [50.4%]	48.0% [29.9%]	-	55.9% [35.4%]
ja	53.5% [19.6%]	81.5% [60.7%]	59.8% [40.2%]	60.7% [40.2%]	-

Table I. : Percentage of language edition X's (row) rule pages that ILL to lang. ed. Y (column). In brackets, percentage of row X's rule pages that ILL to a *rule page* in Y. All values are reported as of February, 2020.

4.2 Rule practice: policy invocation

Rule pages unique to a language edition could be reasonably argued to be an unsurprising variation due to culturally or geo-spatially-sensitive biases. Thus in ongoing work we are considering a smaller set of rules that are core across language editions: Neutral Point of View (WP:NPOV), No Original Research (WP:NOR), and Verifiability (WP:V). Rather than focusing on rule text, we turn to rule practice via policy invocation. Specifically, we compare the topical contexts in which these core, shared rules are invoked across language editions. Because older, core rules receive more attention and use by far in prior work looking at English Wikipedia [Beschastnikh et al. 2008; Keegan and Fiesler 2017], we expect these rules to be most invoked across language editions but also be the most variable in contexts of their use. However, in looking at the contexts of policy invocation, we expect distinct cultural or geo-spatial concerns to emerge, as language editions have demonstrated knowledge diversity that is culturally-biased in content production [Hecht and Gergle 2010].

5. DISCUSSION

Our preliminary findings show significant variance in the rule pages and rule organization across Wikipedia language editions. As most work on governance in Wikipedia focuses on the English language edition, the apparent heterogeneity of rule text and organization across language editions in Wikipedia highlights the need for an investigation of self-governance moving beyond English-speaking contexts. More broadly, this project demonstrates how Wikipedia provides a unique opportunity to conduct multi-community analyses of collective self-governance. Understanding how communities operating under the same technical system towards the same goals self-govern in (dis)similar ways can provide more nuanced understandings of effective governance, such as the impact of conflicting localized rules at a collective level or how one community can learn from another's practices.

In ongoing work, we continue this project by investigating the topical contexts in which core rules are invoked across language editions. We also plan to expand analyses that evaluate the significance of unique and/or conflicting rule pages. Further work should examine in greater detail how shared rules across language editions converge or diverge (both in text and practice) over time to shed light on the complex dynamics of heterogeneous governance practices in diverse, large-scale collaborative systems.

6. ACKNOWLEDGEMENTS

Thanks to the other members of the Community Data Science Collective for their feedback on an earlier version of this work. Support for this work came from the National Science Foundation (grants IIS-1617129, IIS-1617468).

REFERENCES

- Yochai Benkler, Aaron Shaw, and Benjamin Mako Hill. 2015. Peer production: a form of collective intelligence. In *Handbook of Collective Intelligence*, Thomas W. Malone and Michael S. Bernstein (Eds.). MIT Press, Cambridge, MA, 175–204.
- Ivan Beschastnikh, Travis Kriplean, and David W. McDonald. 2008. Wikipedian self-governance in action: motivating the policy lens. In *Proceedings of the AAAI International Conference on Web and Social Media*. AAAI. <http://www.aaai.org/Papers/ICWSM/2008/ICWSM08-011.pdf>
- Laura W. Black, Howard T. Welsler, Dan Cosley, and Jocelyn M. DeGroot. 2011. Self-Governance Through Group Discussion in Wikipedia Measuring Deliberation in Online Groups. *Small Group Research* 42, 5 (Oct. 2011), 595–634. DOI: <http://dx.doi.org/10.1177/1046496411406137>
- Eshwar Chandrasekharan, Mattia Samory, Shagun Jhaver, Hunter Charvat, Amy Bruckman, Cliff Lampe, Jacob Eisenstein, and Eric Gilbert. 2018. The internet’s hidden rules: An empirical study of reddit norm violations at micro, meso, and macro scales. *Proc. ACM Hum.-Comput. Interact.* 2, CSCW (Nov. 2018), 32:1–32:25. DOI: <http://dx.doi.org/10.1145/3274301>
- Casey Fiesler, Jialun “Aaron” Jiang, Joshua McCann, Kyle Frye, and Jed R. Brubaker. 2018. Reddit rules! Characterizing an ecosystem of governance.. In *Proceedings of the AAAI International Conference on Web and Social Media*. AAAI, Stanford, CA, 72–81. <https://aaai.org/ocs/index.php/ICWSM/ICWSM18/paper/view/17898>
- Andrea Forte, Vanesa Larco, and Amy Bruckman. 2009. Decentralization in Wikipedia governance. *Journal of Management Information Systems* 26, 1 (July 2009), 49–72. DOI: <http://dx.doi.org/10.2753/MIS0742-1222260103>
- Aaron Halfaker, R. Stuart Geiger, Jonathan T. Morgan, and John Riedl. 2013. The rise and decline of an open collaboration system: how Wikipedia’s reaction to popularity is causing its decline. *American Behavioral Scientist* 57, 5 (May 2013), 664–688. DOI: <http://dx.doi.org/10.1177/0002764212469365>
- Noriko Hara, Pnina Shachaf, and Khe Foon Hew. 2010. Cross-cultural analysis of the Wikipedia community. *Journal of the American Society for Information Science and Technology* 61, 10 (2010), 2097–2108. DOI: <http://dx.doi.org/10.1002/asi.21373>
- Brent Hecht and Darren Gergle. 2010. The Tower of Babel Meets Web 2.0: User-generated Content and Its Applications in a Multilingual Context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’10)*. ACM, New York, NY, USA, 291–300. DOI: <http://dx.doi.org/10.1145/1753326.1753370> event-place: Atlanta, Georgia, USA.
- Pike Jacqueline C. Joyce, Elisabeth and Brian S. Butler. 2012. Rules and Roles vs. Consensus: Self-Governed Deliberative Mass Collaboration Bureaucracies. *American Behavioral Scientist* 57, 5 (2012), 576–94. DOI: <http://dx.doi.org/10.1177/0002764212469366>
- Brian Keegan and Casey Fiesler. 2017. The Evolution and Consequences of Peer Producing Wikipedia’s Rules. In *Eleventh International AAAI Conference on Web and Social Media*. AAAI, Montreal, Canada, 112–121. <https://aaai.org/ocs/index.php/ICWSM/ICWSM17/paper/view/15698>
- Sara E. Kiesler, Robert E. Kraut, Paul Resnick, and Aniket Kittur. 2012. Regulating Behavior in Online Communities. In *Building Successful Online Communities: Evidence-Based Social Design*, Robert E. Kraut and Paul Resnick (Eds.). The MIT Press, Cambridge, MA, 0. DOI: <http://dx.doi.org/10.7551/mitpress/8472.003.0005>
- Aniket Kittur, Bongwon Suh, Bryan A. Pendleton, and Ed H. Chi. 2007. He Says, She Says: Conflict and Coordination in Wikipedia. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ’07)*. ACM, New York, NY, USA, 453–462. DOI: <http://dx.doi.org/10.1145/1240624.1240698>
- Travis Kriplean, Ivan Beschastnikh, David W. McDonald, and Scott A. Golder. 2007. Community, consensus, coercion, control: cs*w or how policy mediates mass participation. In *Proceedings of the 2007 International ACM Conference on Supporting Group Work*. ACM, New York, NY, USA, 167–176. DOI: <http://dx.doi.org/10.1145/1316624.1316648>
- Cliff Lampe and Paul Resnick. 2004. Slash(dot) and Burn: Distributed Moderation in a Large Online Conversation Space. In *Conference on Human Factors in Computing Systems (CHI)*. ACM, Vienna, Austria, 543–550.
- Sorin Adam Matei and Caius Dobrescu. 2010. Wikipedia’s “neutral point of view”: settling conflict through ambiguity. *The Information Society* 27, 1 (Dec. 2010), 40–51. DOI: <http://dx.doi.org/10.1080/01972243.2011.534368>
- Kim Osman. 2013. The Role of Conflict in Determining Consensus on Quality in Wikipedia Articles. In *Proceedings of the 9th International Symposium on Open Collaboration (WikiSym ’13)*. ACM, New York, NY, USA, 12:1–12:6. DOI: <http://dx.doi.org/10.1145/2491055.2491067>
- Elinor Ostrom. 1990. *Governing the commons: The evolution of institutions for collective action*. Cambridge University Press, New York, NY.
- Wikipedia. 2020a. Help:Interlanguage links. (2020). https://en.wikipedia.org/wiki/Help:Interlanguage_links
- Wikipedia. 2020b. Wikipedia:Policies and guidelines. (2020). https://en.wikipedia.org/wiki/Wikipedia:Policies_and_guidelines
- Anita Williams Woolley, Ishani Aggarwal, and Thomas W. Malone. 2015. Collective Intelligence in Teams and Organizations. In *Handbook of Collective Intelligence*, Thomas W. Malone and Michael S. Bernstein (Eds.). MIT Press, Cambridge, MA, 143–168.