Rules and Rule-Making in the Five Largest Wikipedias

Sohyeon Hwang, Aaron Shaw
Department of Communication Studies, Northwestern University
sohyeonhwang@u.northwestern.edu, aaronshaw@northwestern.edu

Abstract

The governance of many online communities relies on rules created by participants. However, prior work provides limited evidence about how these self-governance efforts compare and relate to one another across communities. Studies tend either to analyze communities as discrete entities or consider communities that coexist within a hierarchically-managed platform. In this paper, we investigate both comparative and relational dimensions of self-governance in similar communities. We use exhaustive trace data from the five largest language editions of Wikipedia over almost 20 years since their founding, and consider both patterns in rule-making and overlaps in rule sets. We find similar rule-making activity across the five communities that replicates and extends prior work on English language Wikipedia alone. However, we also find that these Wikipedias have increasingly unique rule sets, even as editing activity concentrates on rules shared between them. Self-governing communities aligned in key ways may share a common core of rules and rule-making practices as they develop and sustain institutional variations.

Introduction

Online communities in the same social computing system, site, or platform are often bound to a single set of rules, but not always. Communities may be governed by central agents, like a platform owner or a political state; however, many communities also develop and sustain their own extensive rules and self-governance institutions, sometimes in conjunction with other communities operating autonomously within a larger site. How these autonomous rules develop and relate to each other, especially in large and long-sustained communities, is a key aspect of online governance that has received little attention in prior research.

Most prior studies of online community governance focus on a single community or a site as a whole, and overlook patterns of governance across multiple communities. Moreover, the literature provides conflicting intuitions and evidence about whether the self-governing efforts of communities will be similar because they are shaped by shared sociotechnical characteristics (e.g., Shaw and Hill 2014) or different due to the sheer diversity of communities' membership, goals, or organizational culture (e.g., Kiene, Shaw, and Hill 2018). Recent analyses that investigate community self-governance practices across a system in a comparative perspective suggest both elements (of similarity and difference), but do not provide insights into how communities' autonomous rules and rule-related activity develops over time (Müller-Birn, Dobusch, and Herbsleb 2013; Chandrasekharan et al. 2018; Strimling and Frey 2020; Fiesler et al. 2018).

Figure 1: Examples of a Wikipedia rule page for the "Neutral Point of View" rule (NPOV) from the English (left panel) and Japanese (right panel) language editions.

In this paper, we investigate both comparative and relational dimensions of self-governance practices in similar online communities. We focus on the five largest language editions of Wikipedia: a set of related communities pursuing common goals using shared technical infrastructure, but with distinct, autonomous governance arrangements. Using data from the Wikimedia API, our analysis has two parts. First, we compare patterns of rule-related editing in the five language editions, replicating and extending the longitudinal analysis of English Wikipedia from Keegan and Fiesler (2017). Second, we examine overlaps and differences in communities' rule sets over time, leveraging Wikipedia's "interlanguage link" feature that connects pages on equivalent topics across language editions.

We find that the language editions in our study follow similar patterns in rule-related editing over twenty years, having a rapid initial surge in activity followed by stabilization

and a shift towards discussing rules rather than editing them. However, the five Wikipedias also follow distinct trajectories. Further, the language editions' rules overlap less over time, producing increasingly differentiated rule sets even as most activity focuses on widely-shared rules.

This work makes three main contributions: (1) we provide evidence that large, well-established Wikipedia language editions follow similar patterns of rule creation and organizational formalization, replicating and extending prior work that focused predominately on English Wikipedia; (2) we show that communities with the same goals, technical infrastructures, and similar organizational trajectories develop substantial and sustainable institutional variations; and (3) we highlight the importance of broad, widely-shared rules created early on in communities' histories.

Background

Rules and Rule-making in Online Governance

Rules in online communities serve critical social and organizational functions related to governance. In Ostrom (1990)'s analysis, rules are established institutional arrangements that coordinate behavior, get enforced, and have consequences when broken. While governance certainly extends beyond rules, rule-making, or enforcement, rules—especially written or formalized ones—create central, visible places for group members to learn, review, and contest norms.

Rules come about for a variety of reasons and change in a variety of ways. In some cases, rules might be set by founders who have specific community goals (Fiesler and Dym 2020) or emerge out of consensus amongst community members (Kiesler et al. 2012). In other cases, they might arise or change rapidly. For example, Kiene, Monroy-Hernández, and Hill (2016) described how rules on a Reddit community called *NoSleep* changed after the community saw a sudden influx of users, with the rules becoming narrower in scope and more articulated for new users.

In these ways, rules serve as natural histories of group governance (March, Schulz, and Zhou 2000) and shed light on how governance has changed over time. For example, Keegan and Fiesler (2017) examined the evolution of rules on English Wikipedia and found a "policy environment marked by on-going rule-making and deliberation [...]" even after a decade, alongside declining revision activity and uneven attention among rules. One limitation of many studies looking at rules in online settings is that they frequently consider just one community or an entire site. As we elaborate in the following section, looking across and between communities is essential to understanding online governance.

Governance in Multiple Online Communities

Online community governance often happens at several levels via distinct community-level practices as well as the various relationships between them. Prior studies of online community governance have examined a range of sites, platforms, and federated services including Wikipedia, Fandom (formerly Wikia), Reddit, World of Warcraft, Minecraft servers, Facebook Groups, and more. A comprehensive review exceeds the scope of this paper, but several salient patterns stand out. First, prior research looking beyond a single community has tended to pursue either a comparative or a relational approach. In general, comparative research considers patterns *across* communities as comparable, independent units of analysis whereas relational research emphasizes the connections and interactions *between* them. Some key aspects of each motivate our study of rules and rule-making.

Comparative analyses Online communities often share many attributes and they provide an ideal opportunity to understand organization and governance practices through *comparative* studies, which can yield more robust and generalizable results (Hill and Shaw 2019). However, HCI's historical focus on novelty has discouraged replicating findings across empirical settings (Wilson et al. 2011) and a limited body of prior literature has engaged in comparative work.

Comparative analyses of self-governance in communities point to both similarities and differences. Large, long-surviving communities exhibit similarities in governance complexity and rule articulation (Frey and Sumner 2019), parallel trajectories of governance and contributor behavior as communities grow (Müller-Birn, Dobusch, and Herbsleb 2013; TeBlunthuis, Shaw, and Hill 2018; Shaw and Hill 2014), and reflect prevailing patterns in how we think about community governance structure (Schneider 2021). At the same time, self-governance in online spaces can embrace immense opportunities for institutional flexibility and diversity (De Filippi et al. 2020). Despite the shared patterns noted above, rules and norms also differ across communities, even within the same site or platform (Fiesler et al. 2018; Chandrasekharan et al. 2018; Strimling and Frey 2020). Rules may remain context-dependent and tailored to the community at hand even when they overlap, presumably as a product of diverse community goals, content, membership, technical affordances, and histories.

To advance these lines of inquiry, we ask: **RQ1.** *How do patterns of rule-making over time compare across autonomous communities with shared goals and technical infrastructure?* Few prior studies have (to our knowledge) conducted longitudinal comparative analysis of rules and rule-making activity across communities that share both common goals and technical infrastructure as well as the autonomy to pursue divergent organizational practices. Such environments create opportunities for institutional diversity while holding some potential sources of that diversity constant.

Relational analyses of communities Whereas comparative work treats communities as separate entities, a relational approach considers a range of possible interactions and interdependencies between them. Communities, especially those within a single site, do not exist in isolation, but instead within a complex ecosystem of related, overlapping groups. Such communities can compete with and complement each other (TeBlunthuis and Hill 2021; Hwang and Foote 2021), as users move between and engage with them (Tan and Lee 2015).

In terms of governance, communities within a site or platform may be unified by shared constraints and bounds, as the design choices (Malik and Pfeffer 2016) and culture (Massanari 2017) of a site shape activity. Jhaver, Frey, and Zhang

(2021)'s characterization of "multi-level governance structure[s]" in social computing systems emphasizes how community governance is both vertically entangled with "top-level" actors such as platforms and horizontally connected to other communities. In the case of sites or projects with shared goals and technical infrastructure, such as different language editions of Wikipedia, we might expect that communities' rules are even more deeply intertwined.

Conversely, even seemingly similar self-governing communities have diverse rules, norms, and institutional arrangements (e.g., Fiesler et al. 2018; Frey, Krafft, and Keegan 2019; Strimling and Frey 2020). This follows the principles and findings of Ostrom (1990), who argues that rules developed within an autonomous collective have more staying power than those imposed externally. In other words, rules created by communities themselves should last longer, leading to increasing differences among communities.

Together, the prior work considering the relationships between communities' governance practices in a single site emphasizes that they are both intertwined and diverse. More precise characterizations of how communities relate in their self-governance efforts remain warranted, especially in situations where communities share key attributes. To this end, we ask **RQ2.** *How do the sets of rules become more or less similar over time among communities with shared goals and technical infrastructure?*

Governance in Wikipedia(s)

We pursue RQ1 and RQ2 through an analysis of self-governance in a sample of Wikipedia language editions. Wikipedia presents a unique governance environment that deviates from many settings of earlier work in important ways.[1] A massive, global multilingual project, Wikipedia involves millions of contributors across several hundred language editions, which have been substantially oriented in vision and goals by the first language edition, English, which began in 2001. The projects have proceeded in tandem with the Wikimedia Foundation (WMF), a U.S.-based non-profit organization created in 2003.

The governing relationship of the Wikipedia language editions to each other, the WMF, and other entities encompassed by the larger project is amorphous. The WMF provides financial and organizational support to language editions in addition to managing and owning all Wikipedia infrastructure. While WMF staff engage in supra-project level decision-making alongside community leaders, the volunteer editor communities retain mostly autonomous control over their own governance, and individual language editions have many sub-communities. In sum, the entire ecosystem is not reducible to a hierarchical, nested arrangement where the WMF "owns" the language editions, although it does hold resources that allow it to project unique authority.

Several aspects of this governance arrangement inform how we pursue the research questions above. In the comparative component of our inquiry (RQ1), we focus on the rules and related activity within language editions. On a day-to-day level, the projects operate as autonomous governance environments and so we consider them as such here. Studies looking at multiple language editions emphasize the differences across them as they have distinct sets of users and content (Hara, Shachaf, and Hew 2010; Hecht and Gergle 2010; Sen et al. 2015; Hale 2014; Kim et al. 2016), suggesting different organizational practices and needs. At the same time, given that they are connected by their shared mission and various sociotechnical mechanisms, the governance of each community is intertwined. We pursue this relational dimension of our study (RQ2), through an analysis of overlaps in language editions' rule sets.

Methods

To examine patterns of rules and rule-making in communities sharing goals and technical infrastructure, we look at a sample of the five largest language editions of Wikipedia at the time of data collection. We define language edition size in terms of the number of "active" editors, defined as editors who made at least 1 edit in the last 30 days: English, French, German, Japanese, and Spanish[2]. Because language editions operate using shared technical infrastructure and toward the same goals, they are ideal cases to study how self-governance varies and relates across communities. As active and established communities, these language editions also have many people likely to engage with rules as well as long histories over which to observe rule-making behaviors.

To compare patterns of rule-making across communities (RQ1), we examined the revision histories of rule pages. In each wiki, rules can be written and revised by any contributor, and every revision is logged, providing a comprehensive record of rule-making activity. Rule pages are managed differently per language edition, but often appear within the "project" namespace, a distinct sub-domain in a Wikipedia dedicated to project governance and administration.[3]

To study the evolution of rules in relation to each other (RQ2), we leverage "*interlanguage links*" (ILLs).[4] ILLs are intended to connect conceptually equivalent pages across Wikipedia language editions, although exceptions exist for concepts that may collapse or fork across languages. Language editions primarily manage ILLs through a central knowledge graph, Wikidata. In Wikidata, individual entities are assigned unique object identifiers (called "Q identifiers" or "QIDs"). Pages on different language editions that refer to the same QID are (automatically) connected by an ILL.

ILLs can exist for rule pages as well as content article pages and a given ILL need not be reciprocal between language editions. The variable extent to which language editions are conceptually aligned in their articles formed the basis of several important prior studies (Hecht and Gergle 2010) and the design of the Omnipedia system (Bao et al.

[1] See Jemielniak (2014) for a full exposition of Wikipedia's organizational history and form.

[2] As of April 2022, these language editions remain the largest in terms of active editor count.

[3] For more, see the English Wikipedia project namespace page explaining the concept and usage of namespaces within Wikipedia: https://en.wikipedia.org/wiki/Wikipedia:Namespace.

[4] See the English Wikipedia help page on ILLs for a conceptual overview: https://en.wikipedia.org/wiki/Help:Interlanguage_links.

		EN	FR	DE	JA	ES	Total
Rule pages	*N*	313	127	122	113	105	780
	Edits	266,048	31,437	40,591	16,679	27,609	382,364
	Contributors	46,767	4,879	7,952	3,441	5,165	67,020
Rule talk pages	*N*	309	120	115	112	90	746
	Edits	1,091,530	37,098	138,088	37,562	10,237	1,314,515
	Contributors	45,517	2,664	8,275	3,168	2,726	61,888
Total	*Contributors*	80,478	6,424	13,601	5,523	7,130	111,497

Table 1: Rule and Rule talk pages count (N), edits, and contributors per language edition (columns, in order: English, French, German, Japanese, Spanish) for the rule pages. Note that Contributors are unique within page-type and so the final row also provides total unique contributors for both page types.

2012). To our knowledge, no prior work has analyzed this aspect of Wikipedia in order to investigate the relationships between community governance practices.

Data

Our data comes from the revision histories of all rule pages and corresponding talk pages for the five sampled language editions from their start date through December 2020 (nearly twenty years for each language edition). Rule pages are managed differently across language editions, which often define specific types of rules in incommensurate ways. As a result, constructing a sample of rule pages across language editions required multiple steps: (1) define general criteria by which to determine whether pages were rules or not; (2) identify candidate pages per language edition likely to contain at least all pages that would meet the criteria; and (3) apply the criteria to candidate pages to determine the sample to be included in analysis. This provides a general method of extracting comparable samples of rule sets across language editions.

Our criteria to determine whether a page was a rule or not were based on key dimensions of rules noted in our Background: we considered whether the page articulated widely-accepted normative content about "how things should be done" in a language edition that would, when invoked, likely be considered a salient ("official") expectation applicable to all contributions and could, if violated, have consequences (Ostrom 1990). This excludes pages that only contain links to other rules, technical details about how things work in a project, or historical information about previous routines. It also excludes proposals for new rules, reflections on the rationale behind rule(s), debates over the legitimacy of rule(s), or other guides that might influence behavior but lack the normative weight and enforceable quality of rules.

We identified four sources for candidate pages likely to contain at least all rule pages in each language edition:

1. Relevant "list of…" pages in the project namespace. For example, there is a "list of policies" page in English Wikipedia.[5] This includes pages such as WP:Richtlinien in German and WP:Recommandation in French.
2. The policy/guideline (or analogous terms) "category" pages. We look with a max depth of 1 sub-category as preliminary exploratory analysis indicated that pages meeting our criteria are generally listed within this range and further sub-categories extend beyond rule pages quickly.
3. Rule (Policy, Guideline, Richtlinien, Recommandation) navigation boxes on rule pages, which contain links to more rules (E.g., see the blue boxes in the bottom-right of both pages in 1).
4. Each wiki's page of shortcuts.[6] Shortcuts invoke or link to rules (and other oft-referenced pages) elsewhere on a wiki (e.g. "Neutral point of view" invoked with the shortcut "WP:NPOV" on English Wikipedia).

Using these sources, we identified 801 candidate pages as of February 2020 that might be rules. We then applied the definition described above to each page to filter out non-rule pages. For languages unfamiliar to members of the research team, we used Google Translate to aid in this process. This (applied May 2021) yielded a sample of 780 rule pages across the five language editions. For the entire sample, we retrieved full revision histories for each rule page and corresponding talk page using the WikiMedia API. We also retrieved all interlanguage links, described below, for these pages.[7] Summary information about the sample, broken down by language edition, appears in Table 1.

Analysis

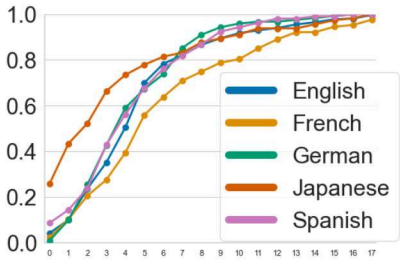
Our analysis consists of two parts: we first examine rule-making patterns using the revision histories of rule pages and their talk pages; we then combine this data with ILL data to evaluate rule overlaps over time. For both, we primarily examine counts, proportions, and descriptive statistics of various attributes of rule-related editing activity.

For all analyses, we compare communities in relative time (i.e. in terms of age rather than calendar time): data for each language edition starts at year 0, defined as the point at which rule-related activity in that Wikipedia began. For

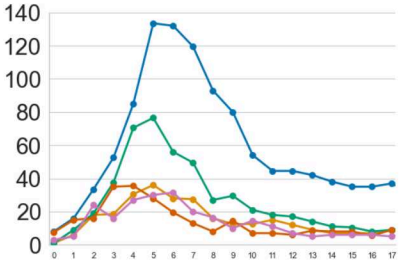
[5] https://en.wikipedia.org/wiki/Wikipedia:List_of_policies

[6] See, e.g., https://en.wikipedia.org/wiki/Wikipedia:Shortcut

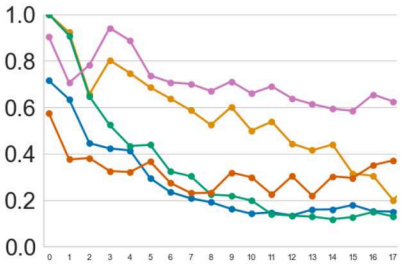
[7] Our API scripts are based on those used in Keegan and Fiesler (2017). All data and code for our work is available as part of the replication materials for the paper to be published at https://doi.org/10.7910/DVN/WSMV2H.



(a) Cumulative distribution function of proportion of rules created



(b) Median number of rule-related revisions



(c) % revisions on rule vs talk

Figure 2: Trends in rule-related revisions (revisions on rule pages or their corresponding talk pages) by years since the first rule-related revision. 0 on the y-axis thus means the first year per language edition of rule-related revision activity.

graphs plotting trends over time, we stop at year 17 (the 18th year), the age of the youngest language editions in our sample at the time of data collection. We align the data in this way for two reasons: (1) our comparative inquiry focuses on patterns in the life-cycles of communities; and (2) we consider relationships between communities through the analysis of ILLs. The Discussion section considers the implications and limitations of this choice.

RQ1: Comparison across language edition rules. We replicate part of the analysis conducted by Keegan and Fiesler (2017) and apply it to our sample of rule pages. That paper investigated patterns in rule-related editing activity in English Wikipedia alone. We adapt the analyses to accommodate some of the variations between the English language edition and the others in our study. In particular, we adapt the investigations into patterns in rule-page revisions (the number of, the lag between, and the size of revisions over time) to examine rule-related editing activity. We also evaluate the findings in Keegan and Fiesler (2017) that revisions go increasingly to talk pages and older rules. One difference is that we consider rules as a general group per language edition instead of distinguishing between classes of rules (e.g., policies vs. guidelines). We also do not conduct user-level analyses, focusing instead on rule pages nested within language editions as the units of analysis.

RQ2: Relationships between language edition rules. We examine ILLs to understand how rules in all five language editions relate to each other. For each rule page in each year, we count the number of ILLs it has to one of the other four language editions; therefore a rule in any language and year will have between zero and four ILLs. We first compare the distributions of rules and revision activity across number of ILLs over language editions over time. Our measure of links per year is not dependent on when a user added an ILL but rather on when two pages with current ILLs coexisted. This helps us avoid misattributing changes in ILL status as changes in overlaps when they are actually due to changes in procedures surrounding ILLs. We then group the rule pages by number of ILLs and consider the most and least connected rules (those with four and zero ILLs respectively) in terms of revision activity, talk activity, and several other attributes. This comparison reveals how rule-related activity patterns vary between the widely-shared rules and community-specific ones.

Results

Comparing Rule-related Activity

RQ1. How do patterns of rule-making over time compare across autonomous communities with shared goals and technical infrastructure?

Rule creation follows similar, but distinct patterns across the wikis in our sample. Figure 2a shows the cumulative proportion of rules created in each year over the lifetime of each wiki. In general, most rules are created early. For all language editions, years 2–6 see the most rule page creation. Japanese Wikipedia is an exception, with 25% of its rules created in its first year (year 0), which makes that first year the year with most rules created for the wiki.

To investigate why Japanese Wikipedia diverged in this regard, we examined the 29 rules created in year 0 of Japanese Wikipedia. Most (25 of the 29) had at least one ILL page in this first year. The earliest ten revisions of each of the 29 rule pages showed that at least 16 of them started as translations of its corresponding contemporary on English Wikipedia (which began rule-making two years before Japanese). Also, many of these early rules were "core" rules such as *Neutral Point of View* and *Be Bold* as well as basic guidelines to essential aspects of the encyclopedia such as *Manual of Style* pages and rules about deletion and site administrators. This suggests that Japanese Wikipedia bootstrapped its core rules from the older English Wikipedia, rather than pursuing a divergent rule-making approach.

Editing activity on rule pages also suggests broad similarities across language editions, shown in Figures 2b and 2c. The median number of rule-related revisions each year per rule, shown in Figure 2b, rises initially, peaks around year 5, and then declines in all language editions. In contrast, the proportion of rule-related revisions to rule versus corresponding talk pages, shown in Figure 2c, starts declining over the first decade of most projects and then flattens out. For both of these metrics, we note substantial variations between the projects, despite similar overall patterns.

We next apply a coarse periodization scheme from Halfaker et al. (2013) to define *early*, *growth*, and *decline* eras

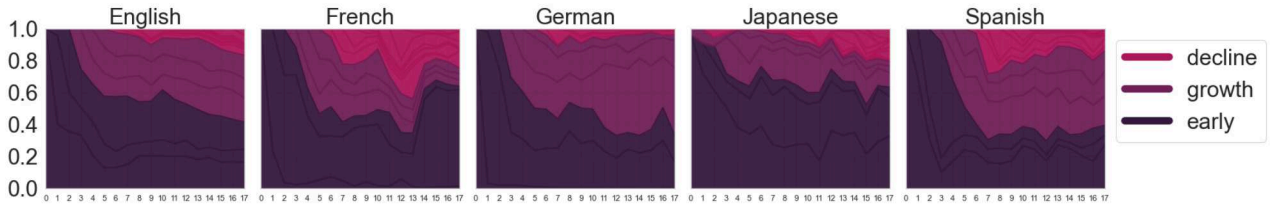


Figure 3: The proportion of rule-related revisions being made to rules that were created in "eras" of the language edition over the first 18 years of rule-making on each Wikipedia.

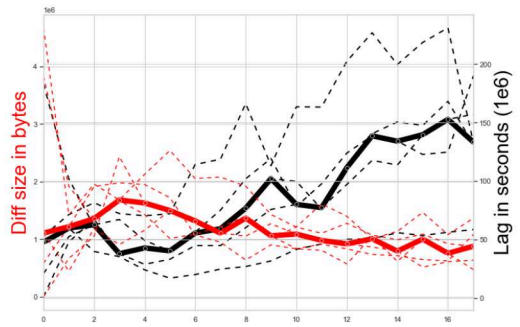


Figure 4: Annual medians of diff sizes/lags pooled from monthly median of average rule page diff sizes (red, in bytes) and lags (black, in seconds) between revisions per language edition. Each language edition is represented by a dashed line; the bold line is the median trend across language editions.

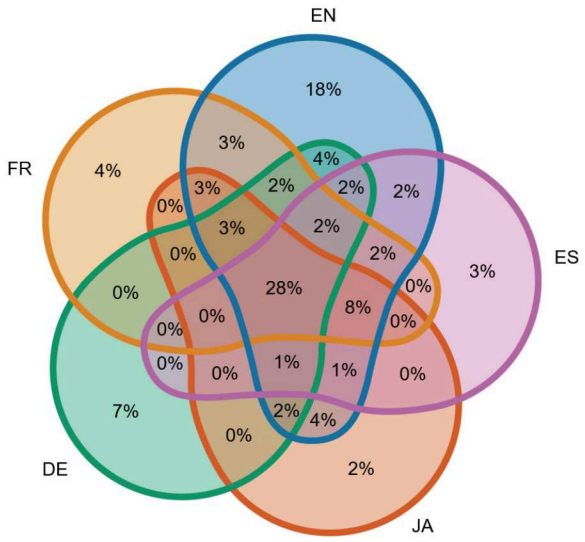


Figure 5: Percentages of language edition rule set overlap. Note that the diagram is not to-scale.

for each project based on the number of active editors.[8] Figure 3 shows the proportion of rule-related revisions in a year belonging to each of these three eras faceted by language edition. Rules created in the *early* and *growth* eras dominate activity across all five language editions. However, the degree to which *early* versus *growth* era rules were edited varies by language edition. For example, the Japanese language edition has the highest proportion of *early* era rule revisions, while the Spanish language edition has a larger proportion of *growth* era rule revisions. In none of the five Wikipedias do rules created in the *decline* era attract a large proportion of annual rule revisions.

Across the lifetime of all five language editions, rule-related revisions become less frequent, get smaller, and tend to shift from rule pages themselves to talk pages. We visualize the first two patterns as a single time series plot of the "grand medians" (over all language editions) of revision latency (time between revisions) and revision size (byte length) in Figure 4. As the projects age, the latency increases and the byte size rises at first and then declines. As with other measures, the degree of increase and decrease in these trends varies by language edition (the dashed lines in Figure 4). Figure 2c shows that rule-related revisions are increasingly on talk pages. Together, these shifts suggest that rule-related work involves increasing discussion rather than changing rule text itself. However, we also note that Spanish Wikipedia currently has a much larger percentage of its revisions on rule (rather than talk) pages, about 60% to the other language editions' 10-40%.

In sum, the five largest Wikipedia language editions have followed similar but distinct patterns in rule-making over about twenty years. Although communities differ on every measure, they have striking commonalities. Most rule pages are created and most actively edited early on as the projects grow. In the last year, language editions appear to have relatively stable sets of rules. While rules remain, in the words of Keegan and Fiesler (2017), "dynamic" as the projects age, editing activity on them stabilizes and shifts to rules' talk pages. Throughout, rule-related editing has remained concentrated among rules created earlier on.

Similarity and Overlap in Rules Sets

RQ2. How do the sets of rules become more or less similar over time among communities with shared goals and technical infrastructure?

Although the language editions show similar trends in patterns of rule-related activity, the interlanguage links (ILLs)

[8] Halfaker et al. (2013) defines the eras in the following way: *early* begins at the community's creation; *growth* begins in the year the number of active editors starts to increase exponentially; and *decline* begins in the year this growth peaks or tapers off.

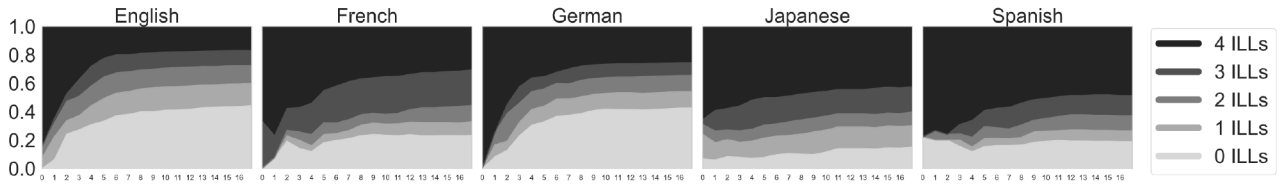


Figure 6: The proportion of rules in a focal language edition in each year (X-axis) with the number of ILLs to the other language editions in the set (area shading).

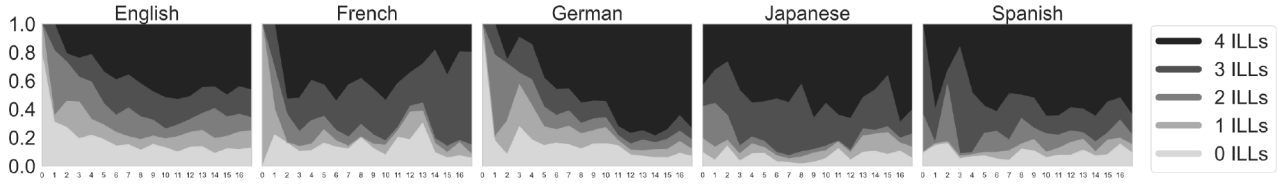


Figure 7: The proportion of rule-related revisions in a focal language edition in each year (X-axis) made to rules with the number of ILLs to other language editions in the set (area shading).

between their rule sets reveal substantial differences.

The number of rules shared between the five largest Wikipedia language editions fluctuates over the different possible combinations of the language editions. Figure 5 shows a 5-way venn diagram that visualizes the percentages of rules shared across language editions based on ILLs at the time of data collection.[9] About 28% of all rule pages are shared across all five language editions. The number of rule pages on only one language edition (259 rules with 0 ILLs) rivals and then surpasses the number of those shared by all five language editions (221 rules with 4 ILLs).

This distribution of connectedness by ILLs has also shifted over time, with the language editions' rule sets becoming slightly more differentiated from each other. Figure 6 shows the proportion of rules having zero to four ILLs each year, faceted by language edition. We note that overall proportions are skewed by the number of unique rules in the English and German language editions, indicated by the lightest shaded area in those two panels. However, as of the time of data collection, all of the language editions contain a substantial proportion of rules that have either no ILLs or link to only one other language edition in the set (approximately 20% or more). Also, for all five language editions, the proportion of these 0-1 ILL rules has grown over time. In other words, the majority of rules in all five wikis are shared and connected by ILLs, but the rule sets increasingly contain rules that are unique or shared with just one other language edition. Moreover, the proportion of rules connected by ILLs in each wiki has stabilized (although this may be a function of the stabilization of the number of rules as well).

Rule-related revision predominately and increasingly goes towards rules with more ILLs. Figure 7 shows the proportion of rule-related revisions (including talk page revisions) going to rules with between zero and four ILLs. What distinguishes rules shared across all language editions from those that exist only on one language edition in our sample? Table 2 reports several summary statistics. Comparing the medians of the two groups, rules with 4 ILLs (221 rule pages) typically see 3.6 times more revisions and 3.9 times more unique contributors than those with zero ILLs (259 rule pages). Moreover, 4-ILL rules are created slightly earlier, with a median creation year of 2004, while 0-ILL rules have a median creation year of 2007. The 4-ILL rules also have 70.8 times more unique pages linking to them (wikilinks) than the 0-ILL rules.

		0 ILLs	4 ILLs
Median	Rule creation year	2007	2004
	Rule page length (char bytes)	8,100	14,251
	# Wikilinks	449	31,806
	# Revisions	195	371
	# Unique contributors	53	234
Total	% revisions rule-to-talk	29.3%	23.3%

Table 2: Attributes of rule pages with 0 ILLs and those with 4 ILLs, at the end of 2020. The number of Wikilinks is from the Link Count API provided by Wikipedia. Note that the number of revisions and unique contributors are for rule and rule-talk pages.

Overall, our analyses examining ILLs revealed that language editions' rule sets became increasingly differentiated from one another, though this trend slowed by year 5 in each wiki. While the number of 0-ILL rules have grown, the proportion of these community-specific rules per language edition varies. For example, the presence of community-specific rules is more obvious in the two larger language editions, English and German. Notably, we find that community-specific rules are not the focus of rule-related revision activity in any community, which instead goes towards rules shared across all five language editions. A comparison of widely-shared rules versus community-specific ones indicates that the former are created earlier, have longer pages, see more revisions and contributors, and

[9]The areas plotted in the diagram are not to scale.

are more widely referred to on other pages.

Discussion

Our results contribute several novel insights into how rules and rule-making have supported self-governance in the five largest language editions of Wikipedia over nearly twenty years. Rule-related editing activity peaked in early years, and has since stabilized, slowed, and shifted to discussion. The communities have created and elaborated a common core of rules through similar rule-related editing behaviors. They have also combined these shared rules and practices with some localized ones: the rule sets of the language editions have become increasingly diverse over time, as indicated by the growth of community-specific rules, even as editing activity has remained focused on widely-shared rules, which tend to be older, longer, and more heavily linked. We elaborate on these themes and their implications below in light of the prior work that motivated this study.

Stabilization in self-governance Our analyses extend and reproduce several earlier findings using almost 100 years of data aggregated across five communities. We find a rise-and-decline pattern in rule-related editing in all five language editions, mirroring patterns of more general activity and rule-related editing on English Wikipedia (Halfaker et al. 2013; Keegan and Fiesler 2017) as well as activity in other large wiki communities (TeBlunthuis, Shaw, and Hill 2018). Also consistent with prior work on English Wikipedia (Keegan and Fiesler 2017), all language editions have seen fewer and smaller rule-related revisions over time, with a growing proportion of them going to talk pages.

Why do we see these similarities in patterns of rule-related activity across these communities? We find that the five largest Wikipedias have all experienced rule-related stabilization alongside formalization of self-governance. Prior work on English Wikipedia and other communities has characterized similar patterns to those we observe as indicative of "diminishing flexibility to change [rules]" (Keegan and Fiesler 2017), "policy calcification" (Halfaker et al. 2013), and "entrenchment" (Halfaker et al. 2013; TeBlunthuis, Shaw, and Hill 2018). In this vein, our findings resonate with existing claims that the formalization of institutions in self-governing online communities accompanies the consolidation of organizational authority (Shaw and Hill 2014; TeBlunthuis, Shaw, and Hill 2018). Future work might consider how editor experience relates to patterns of rule-related activity, including the enforcement of rules, in order to elaborate and test this affinity between governance formalization and authority consolidation more fully.

Generalizability of findings Patterns shared across the five language editions suggest commonalities that may generalize even more widely to other peer production systems and online communities engaged in self-governance. Earlier studies of English Wikipedia or any single community alone could not exclude the possibility that observed patterns derived from linguistic, organizational, or other project-specific factors (such as scale, age, or specific events) unique to that community (Hill and Shaw 2019). By comparing five Wikipedias we eliminate some of these potential alternative explanations. However, it is possible the size or longevity of the language editions in this sample or other commonalities between the communities (e.g., they are all Wikipedias) limit the generalizability of the findings. Future work considering the experiences of younger, smaller Wikipedias and more diverse types of online communities will be necessary to address this concern empirically.

Developing diverse institutions. Although the language editions are aligned in key aspects—they share goals and technical infrastructure—the similar trajectories noted in the preceding section do not point to identical institutional arrangements. We found that the number of community-specific 0-ILL rules grew to surpass the number of widely-shared 4-ILL rules. Our results extend prior findings highlighting the diversity of institutional rules across communities within a single system (Chandrasekharan et al. 2018; Fiesler et al. 2018; Strimling and Frey 2020).

Zero-ILL rules exemplify one way that the communities sustain differences that merit further explanation. We note that several other differences relate to project age and size. In particular, the sheer scale of the German and English language editions renders some points of comparison difficult between these two projects and others in our sample. Some differences, such as the bulk importation of rules at the beginning of the Japanese language Wikipedia, may reflect project sequencing, learning, and influence across projects. Other variations, such as Spanish Wikipedia's propensity to sustain a uniquely high proportion of rule-related activity on rule pages (rather than talk pages) require further investigation. Such variations may reflect differences in, for example, communities' organizational culture (Kiene, Shaw, and Hill 2018), the actions of influential leaders, or other factors.

Our observational analyses cannot provide conclusive explanation of the causes or effects of these differences. However, given that our sample holds the goals and technical infrastructure of these Wikipedias near-constant, differences between them are not likely explained by these factors. Our evidence suggests instead that institutions of online community self-governance have substantial and sustainable variations (Frey, Krafft, and Keegan 2019; Ostrom 1990). Moreover, it suggests that these institutional variations may scaffold the relationships between communities, which may range from competitive to complementary in nature TeBlunthuis and Hill (2021).

The role of older, shared rules. Although overlaps in rule sets decreased, rule-related revisions increasingly concentrated on older rules shared across all five language editions. These shared 4-ILL rules, particularly those created early in projects' histories, include those core to the Wikipedia mission such as rules connected to the "Five pillars" of Wikipedia. Compared against 0-ILL rules, they to be older as well as more actively edited by a larger group of contributors, and are far more widely referenced. A preliminary, post-hoc exploration following this analysis suggests that many of the 4-ILL rules are essentially obligatory because they reflect the founding principles shared by all Wikipedias. As language editions maintain the same goal of peer produc-

ing a reliable online encyclopedia under a certain ethos, the shared rules likely help coordinate this shared mission. We suggest that the centrality and breadth of early, shared rules would also explain why revision activity skews to them. Given that many of the original core rules began with English Wikipedia, this may also suggest a path dependency resulting in similar institutional arrangements.

The rules shared across all five language editions also elicit ongoing deliberation. Prior work on English Wikipedia has underscored the importance of deliberation in communities' understandings of rules, particularly for core, ambiguous rules such as "Ignore all rules" established early in Wikipedia's history (Matei and Dobrescu 2010; Black et al. 2011). This deliberation also includes critical discussion of rules on other pages such as *essays* (Morgan and Zachry 2010). Building on this set of prior work, the active talk pages we observe suggest that each language edition continuously engages in a process of discussing, adapting, and internalizing widely shared rules. In addition to longitudinal comparisons of rule content across language editions, further work examining how communities interpret and invoke shared rules can clarify if attention to shared rules is a sign of intertwined governance or divergence.

Both the growth of community-specific rules and active discussion of shared rules in each language edition suggest that communities foster their own distinct institutional understandings, echoing the importance of internally understood rules in communities (Ostrom 1990). Furthermore, the sustained importance of older, shared rules highlights an relationship between communities that differs from competitive ones described in much work of the work that focuses on overlaps in ecologies of communities (Wang, Butler, and Ren 2012). Instead, the overlaps of the shared rules underscore how communities co-exist in complementary ways (TeBlunthuis and Hill 2021; **?**).

Limitations

Our study contains several other limitations that point to directions for future work. For example, as noted above, our sample of Wikipedias contains the five *largest* Wikipedia communities. This includes a wide size range in terms of active editors, but is neither exhaustive nor representative as we selected communities in order to compare a set similar in activity and age. For example, a small community created recently may heavily copy rules from larger and older communities, pointing to a different explanation. Future work should examine a larger and more diverse sample along dimensions of community size and age. Moreover, the research setting of Wikipedia limits the generalizability of our findings: unlike many sites where each community may have different topics, Wikipedia language editions are unified by a shared mission. Research should consider cross-community rule development on other sites.

We are also limited in describing relationships between communities. While ILLs are centrally managed via Wikidata, they are imperfectly maintained and cannot encompass the range of possible relationships between pages and concepts. For example, some ILLs may simply have yet to be created, thus contributing to the growing number of rules with fewer ILLs. Since we look at current ILLs to consider overlaps retrospectively, our data also does not reflect changes in ILL maintenance. Other ILLs may exist despite substantial variations in the content of the pages in question. One-to-one mappings of concepts across language editions via ILLs often break, with one page on one language edition becoming several on another Hecht and Gergle (2010). Within the scope of this project, these issues underscore the larger pattern of community-specific rule development. Future work might address these constraints and examine other dimensions of intercommunity interactions, such as rule page connections via cross-lingual editors or the diffusion of specific rules, which may point to clearer mechanisms of rule overlap and change. In particular, as noted above, a close examination of the content of rules that are widely-shared versus those that are community-specific can provide a clearer image of *why* communities diverge in rule development.

Finally, our decision to align language editions temporally for the purposes of comparison limits certain possibilities of comparison within our data. Temporal alignment puts communities on parallel footing and obscures sequential relationships in the data. Alignment is appropriate for our study because we focused on general life-cycles of communities' governance and explore relationships between communities through ILLs. Future work can consider event sequences directly.

Conclusion

This study examined rule-related editing across the five largest Wikipedia language editions to investigate how self-governing communities that share goals and technical infrastructure compare and relate. We found that language editions experience similar trajectories in rule-making and each continued to pay attention to broad, widely-shared rules created early on. However, the language editions also developed distinct institutions through community-specific rules and active discussions about core rules. The tendency towards rule stabilization in all language editions highlights how key early years are to communities' institutional formation. The variations sustained across communities underscore how shared goals, technologies, and even organizational development could still result in diverse sets of rules. Finally, the continued attention to shared rules and the growth of community-specific rules also indicate potential mechanisms by which Wikipedia language editions coordinate around common principles and internalize these principles in distinct ways.

Acknowledgements

The authors would like to thank Brian Keegan and Casey Fiesler who generously shared their code for analysing rule page revision histories on English Wikipedia; Agnes Horvat, Darren Gergle, members of the Community Data Science Collective, and the reviewers, all of whom provided thoughtful feedback on this project; and undergraduate research assistant Serene Ong, who assisted with data collection and cleaning. Support for this work came from the National Sci-

ence Foundation (IIS-1910202, DGE-1842165) and Northwestern University.

References

Bao, P.; Hecht, B.; Carton, S.; Quaderi, M.; Horn, M.; and Gergle, D. 2012. Omnipedia: Bridging the Wikipedia Language Gap. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '12, 1075–1084. New York, NY, USA: ACM.

Black, L. W.; Welser, H. T.; Cosley, D.; and DeGroot, J. M. 2011. Self-Governance through Group Discussion in Wikipedia Measuring Deliberation in Online Groups. *Small Group Research* 42(5): 595–634.

Chandrasekharan, E.; Samory, M.; Jhaver, S.; Charvat, H.; Bruckman, A.; Lampe, C.; Eisenstein, J.; and Gilbert, E. 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): 32:1–32:25.

De Filippi, P.; Frey, S.; Schneider, N.; Tan, J. Z.; and Zhang, A. X. 2020. Modular Politics: Toward a Governance Layer for Online Communities. *arXiv:2005.13701 [cs]* .

Fiesler, C.; and Dym, B. 2020. Moving Across Lands: Online Platform Migration in Fandom Communities. *Proc. ACM Hum.-Comput. Interact.* 4(CSCW1): 042:1–042:25.

Fiesler, C.; Jiang, J. A.; McCann, J.; Frye, K.; and Brubaker, J. R. 2018. Reddit Rules! Characterizing an Ecosystem of Governance. In *Proceedings of the International AAAI Conference on Web and Social Media*, 72–81. Stanford, CA: AAAI.

Frey, S.; Krafft, P. M.; and Keegan, B. C. 2019. "This Place Does What It Was Built for": Designing Digital Institutions for Participatory Change. *Proc. ACM Hum.-Comput. Interact.* 3(CSCW): 32:1–32:31.

Frey, S.; and Sumner, R. W. 2019. Emergence of Integrated Institutions in a Large Population of Self-Governing Communities. *PLOS ONE* 14(7): e0216335.

Hale, S. A. 2014. Multilinguals and Wikipedia Editing. In *Proceedings of the 2014 ACM Conference on Web Science*, 99–108. New York, NY, USA: ACM.

Halfaker, A.; Geiger, R. S.; Morgan, J. T.; and Riedl, J. 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): 664–688.

Hara, N.; Shachaf, P.; and Hew, K. F. 2010. Cross-Cultural Analysis of the Wikipedia Community. *Journal of the American Society for Information Science and Technology* 61(10): 2097–2108.

Hecht, B.; and Gergle, D. 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, 291–300. Atlanta, Georgia, USA: ACM.

Hill, B. M.; and Shaw, A. 2019. Studying Populations of Online Communities. In Foucault Welles, B.; and González-Bailón, S., eds., *The Oxford Handbook of Networked Communication*, 173–193. Oxford, UK: Oxford University Press.

Hwang, S.; and Foote, J. D. 2021. Why Do People Participate in Small Online Communities? *Proc. ACM Hum.-Comput. Interact.* 5(CSCW2). doi:10.1145/3479606. URL https://doi.org/10.1145/3479606.

Jemielniak, D. 2014. *Common Knowledge?: An Ethnography of Wikipedia*. Stanford, CA: Stanford University.

Jhaver, S.; Frey, S.; and Zhang, A. 2021. Designing for Multiple Centers of Power: A Taxonomy of Multi-Level Governance in Online Social Platforms. *arXiv:2108.12529[cs]* .

Keegan, B.; and Fiesler, C. 2017. The Evolution and Consequences of Peer Producing Wikipedia's Rules. In *Proceeings of the Eleventh International AAAI Conference on Web and Social Media*, 112–121. AAAI.

Kiene, C.; Monroy-Hernández, A.; and Hill, B. M. 2016. Surviving an "Eternal September": How an Online Community Managed a Surge of Newcomers. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 1152–1156. New York, NY: ACM.

Kiene, C.; Shaw, A.; and Hill, B. M. 2018. Managing Organizational Culture in Online Group Mergers. *Proc. ACM Hum.-Comput. Interact.* 2(CSCW): 89:1–89–21.

Kiesler, S. E.; Kraut, R. E.; Resnick, P.; and Kittur, A. 2012. Regulating Behavior in Online Communities. In Kraut, R. E.; and Resnick, P., eds., *Building Successful Online Communities: Evidence-Based Social Design*. MIT Press.

Kim, S.; Park, S.; Hale, S. A.; Kim, S.; Byun, J.; and Oh, A. H. 2016. Understanding Editing Behaviors in Multilingual Wikipedia. *PloS one* 11(5): e0155305.

Malik, M.; and Pfeffer, J. 2016. Identifying Platform Effects in Social Media Data. *Proc. Tenth International AAAI Conference on Web and Social Media* 10(1).

March, J.; Schulz, M.; and Zhou, X. 2000. *The Dynamics of Rules: Change in Written Organizational Codes*. Stanford, CA: Stanford University Press.

Massanari, A. 2017. #Gamergate and The Fappening: How Reddit's Algorithm, Governance, and Culture Support Toxic Technocultures. *New Media & Society* 19(3): 329–346.

Matei, S. A.; and Dobrescu, C. 2010. Wikipedia's "Neutral Point of View": Settling Conflict through Ambiguity. *The Information Society* 27(1): 40–51.

Morgan, J. T.; and Zachry, M. 2010. Negotiating with Angry Mastodons: The Wikipedia Policy Environment As Genre Ecology. In *Proceedings of the 16th ACM International Conference on Supporting Group Work*, GROUP '10, 165–168. New York, NY, USA: ACM.

Müller-Birn, C.; Dobusch, L.; and Herbsleb, J. D. 2013. Work-to-Rule: The Emergence of Algorithmic Governance in Wikipedia. In *Proceedings of the 6th International Conference on Communities and Technologies*, 80–89. New York, NY, USA: ACM.

Ostrom, E. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. New York, NY: Cambridge University Press.

Schneider, N. 2021. Admins, Mods, and Benevolent Dictators for Life: The Implicit Feudalism of Online Communities. *New Media & Society* .

Sen, S. W.; Ford, H.; Musicant, D. R.; Graham, M.; Keyes, O.; and Hecht, B. 2015. Barriers to the Localness of Volunteered Geographic Information. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 197–206. Seoul Republic of Korea: ACM.

Shaw, A.; and Hill, B. M. 2014. Laboratories of Oligarchy? How the Iron Law Extends to Peer Production. *J Commun* 64(2): 215–238.

Strimling, P.; and Frey, S. 2020. Emergent Cultural Differences in Online Communities' Norms of Fairness. *Games and Culture* 15(4): 394–410.

Tan, C.; and Lee, L. 2015. All Who Wander: On the Prevalence and Characteristics of Multi-Community Engagement. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, 1056–1066. Republic and Canton of Geneva, Switzerland: International World Wide Web Conferences Steering Committee.

TeBlunthuis, N.; and Hill, B. M. 2021. Identifying Competition and Mutualism Between Online Groups. *arXiv:2107.06970 [cs]* .

TeBlunthuis, N.; Shaw, A.; and Hill, B. M. 2018. Revisiting "The Rise and Decline" in a Population of Peer Production Projects. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 355:1–355:7. New York, NY: ACM.

Wang, X.; Butler, B. S.; and Ren, Y. 2012. The Impact of Membership Overlap on Growth: An Ecological Competition View of Online Groups. *Organization Science* 24(2): 414–431.

Wilson, M. L.; Mackay, W.; Chi, E.; Bernstein, M.; Russell, D.; and Thimbleby, H. 2011. RepliCHI - CHI Should Be Replicating and Validating Results More: Discuss. In *CHI'11 Extended Abstracts on Human Factors in Computing Systems*, CHI EA '11, 463–466. New York, NY: ACM.