

The Evolution and Consequences of Peer Producing Wikipedia’s Rules

Brian C. Keegan & Casey Fiesler

Department of Information Science
University of Colorado, Boulder
Boulder, Colorado, USA 80309

Abstract

Wikipedia’s policies, guidelines, and other rules can be revised edited by anyone at any time, creating a rule environment that has changed substantially over its 15-year history. Wikipedia provides a remarkable benchmark for understanding the potential and pitfalls of self-governance in a knowledge commons and empirically informing theories of networked governance and institutional analysis. Using a corpus of 725,000 revisions made to 2,012 pages about rules and rule discussions since 2001, we explore the dynamics of English Wikipedia’s rule-making and maintenance over time. Our analysis reveals a policy environment marked by on-going rule-making and deliberation across multiple regulatory levels more than a decade after its creation. This dynamism is however balanced by strong biases in the attention and length towards older rules coupled with a diminishing flexibility to change these rules, declining revision activity over time, and a strong shift toward deliberation.

Understanding how large social systems govern themselves is a fundamental question in sociology, political science, and organizational studies. When we speak of “rules”, we refer to explicitly encoded regulations that are distinct from “norms” and other tacit and shared expectations. These rules organize human actions and coordinate our social lives and manifest themselves in classic bureaucratic forms as organization charts, standard operating procedures, contracts, and sanctions. Wikipedia has adopted a model of content production as the “encyclopedia that anyone can edit” and then applied this radically egalitarian model to its own governance: Wikipedia’s policies, guidelines, and other rules can be revised edited by anyone at any time. How can a large-scale social system be effectively governed if its rules are not concrete walls marking clear boundaries but are rather drifting dunes that shift over time?

Policy creation and maintenance with online peer production projects like Wikipedia are essential components of keeping the community operating (Butler, Joyce, and Pike 2008). But much of the governance-related research on Wikipedia has focused on processes like administrative roles, conflict resolution, or project coordination rather than attending to the dynamics of how the rules themselves have changed over time. Understanding the structure and evolu-

tion of policies on Wikipedia has implications for designing governance both for organizations adopting knowledge sharing services as well as for understanding long-term organizational processes within a large-scale socio-technical system. In the face of on-going challenges to recruit new users and sustain existing content (Suh et al. 2009), questions about how Wikipedia governs its content and users by adopting algorithmic tools and slowing policy creation have taken on heightened importance (Halfaker et al. 2013).

This paper revisits earlier work that explored Wikipedia’s policy space (Beschastnikh, Kriplean, and McDonald 2008; Butler, Joyce, and Pike 2008; Kriplean et al. 2007; Black et al. 2011) to examine how Wikipedia rule editing has changed over the past 15 years as well as the consequences of editors’ engagement with rulemaking. These policies are not monolithic, but represent distinct genres for different classes of content or user behavior (Morgan and Zachry 2010). We also contribute to the debate around whether Wikipedia’s flat hierarchies, decentralization, and little managerial control are evidence of “ad hococracy” (Konieczny 2009) or if the centralized contributions and declining mobility of editors into its administrative ranks is evidence of oligarchic organizing (Shaw and Hill 2014).

We analyze how rule-making patterns on English Wikipedia have changed over time, how patterns of user contributions have evolved, and how participation in rule-making on Wikipedia alters editors’ behavior. Across Wikipedia’s complex rule ecosystem, we find evidence for declining revision activity and an increased emphasis on deliberation. Most users’ engagement with rule-making is intense but fleeting, but the rule ecosystem nevertheless benefits from a core of dedicated users who work together across many types and levels of rules. Employing sequence analysis methods (Keegan, Lev, and Arazy 2016), we also identify significant changes in editors’ mobility between different types of articles that enables alignments between rules and articles.

Background

Dynamics of rules

Rules are not static and understanding their dynamics, beyond an administrative focus on implementation or a legal focus on interpretation, requires understanding how they

are selected, adapted, and diffused throughout an organization. Changes in rules and policies are central to both legal and organizational studies. Rules can be seen through a variety of perspectives: (1) conscious, rational efforts to organize efficiently, (2) proliferating organisms that continually create rules to maintain legitimacy, (3) elements used by members to construct, understand, and influence the organization, and (4) accumulated knowledge and experience over time. Rules are conspicuous symbols within organizations, their impersonality substituting for direct managerial supervision, their power focusing organizational attention and discourse, and their legibility transmitting organizational knowledge (March, Schulz, and Zhou 2000).

Rule use and rule making are central for understanding the operation of an organization. Rules may be administrative documents enumerating standards for conduct (“rules-in-form”) or generally known practices and widely-enforced practices (“rules-in-use”). Rule making and rule use unfolds at the *operational* level of day-to-day decisions, whereas *collective choice* level of individuals interacting to make rules at the operational level. Finally, the *constitutional* level defines the rules governing who may make decisions at the collective choice level (Ostrom and Hess 2007).

In online communities, the rules that community members follow often come from multiple sources — such as law, Terms of Service, community guidelines, or social norms — and negotiating multiple sources of rules can be a challenge (Fiesler, Feuston, and Bruckman 2015). In cases where rules conflict, community members may turn to rules that are most internal to that space, such as norms and community-created rules, which have been shown to have more staying power than externally imposed rules (Ostrom and Hess 2007). Rules created by the community also tend to be easier to understand and more closely aligned with existing norms and practices (Pater et al. 2016). Wikipedia’s rules, created through peer production, are one example of rules that are internal to the community.

Rule-making on Wikipedia

Wikipedia employs a number of robust, formal processes and policies to govern itself since its foundation. Wikipedia’s policies operate in distinct, but overlapping, forms of policy networks: editors come together to co-author and maintain policies, discuss proposals about policies on their talk pages, and reference these policies in other deliberative forums. These rules — whether by convergence or design — emulate many of Ostrom’s principles for governing common goods such as localized solutions, participatory rule-making, and technologically-mediated monitoring (Viégas, Wattenberg, and McKeon 2007).

Wikipedia emphasizes consensus-formation, graduated sanctions, and local enforcement of rules, all of which contribute to decentralized governance processes that allow Wikipedia to scale to millions of articles and users in spite of coordination and enforcement costs (Forte, Larco, and Bruckman 2009). These policies are an important mechanism by which a large-scale social system like Wikipedia is able to control user behavior, even as rules evolve and new ones are created (Kriplean et al. 2007). Analysis of changes

to policies emphasized variability and decentralization in the creation, modification, interpretation, and enforcement of these policies, but provided no metrics by which to evaluate the level of decentralization or stabilization in the revision and enforcement of policy.

Wikipedia has several levels of policies that vary in their adherence, enforcement, and flexibility.

Policies. A set of rules with very high levels of adherence and enforcement governing topics around article content, user conduct, and procedural processes like article deletion and administrative enforcement.¹ These are “rules-in-form” at the collective choice level with constitutional-level policies around administrative processes.

Guidelines. A set of “best practices” with moderate levels of adherence and enforcement that further expand on policies about behavior, content, deletion, editing, article naming, topic notability, and stylistic standards.² These are “rules-in-use” typically operating at a combination of the collective choice and operational levels.

Essays. A set of approximately 1,400 opinions from editors that typically lack broader consensus or formal approval and may be intended for specific cases.³ These are “rules-in-use” typically operating at the operational level.

Failed proposals. A set of 474 proposed policies or guidelines that did not obtain sufficient consensus within the community to warrant elevation.⁴ These may reflect a set of “rules-in-use” like essays for which there may be consensus in the margins, but they also mark instances of rule-making that the community explicitly rejected.

These rules are also employed by different kinds of actors in different situations, varying from consensus-formation between editors in an article talk page all the way up to formal arbitration processes. Crucially, one of the “pillars” is the “Ignore all Rules” that enables sanctioned rule breaking when a general rule might be inappropriate for a specific case (Joyce, Pike, and Butler 2013). The Wikipedia community members enforce rules through other socio-technical capacities such as changing the affordances of the underlying MediaWiki software (Müller-Birn, Dobusch, and Herbsleb 2013), automating editing actions taken by both people and software agents (Geiger and Ribes 2010), and implementing technical standards (Niederer and Dijck 2010).

Questions about the oligarchical tendencies in peer produced governance have provided mixed results. Earlier findings suggest that the mutability of policies have limited oligarchical power on the English Wikipedia (Konieczny 2009), matched by a broad decline in administrator influence over article content (Kittur et al. 2007a). In contrast, comparative results across several hundred wikis concludes there is a general tendency for early contributors to monopolize positions of formal authority (Shaw and Hill 2014). There has likewise been a general shift away from direct work on article content towards indirect work on discussion and policy pages (Kittur et al. 2007b). Users’ literacy with

¹Wikipedia:List of policies

²Wikipedia:List of guidelines

³Wikipedia>About essay searching

⁴Category:Wikipedia failed proposals

Wikipedia policies influences their ability to seek out and understand these administrative processes, the policies that are persuasive in these contexts, and the forums to engage in deliberation likewise allow them to prevail over less literate users (Ford and Geiger 2012).

Despite the extensive amount of discussion around many policies, rule-making on Wikipedia has not accumulated a definitive case law that is used as precedent for decision making. Instead, Wikipedia refines policies while still maintaining a level of ambiguity that allows users to claim and re-interpret them strategically (Matei and Dobrescu 2010). Wikipedia policies can be edited directly by most editors and these policy pages also have “talk pages” where users deliberate over changes. Much as judges cite both legal codes and prior cases as precedents for their decisions, editors’ discussions cite Wikipedia policies during discussions (Beschastnikh, Kriplean, and McDonald 2008).

RQ1: How have rule-making patterns on Wikipedia changed over time?

Temporal mis-aggregation

Prior work has examined what behavioral features predict users’ promotion to administrative roles (Leskovec, Huttenlocher, and Kleinberg 2010). While these “admins” are intended to serve more janitorial than managerial roles (Burke and Kraut 2008), they nevertheless have substantial discretion and “hard” power to protect pages from revision or to block users from editing. Other work has explored the role of shared leadership to help coordinate projects (Zhu, Kraut, and Kittur 2012). Less explored in these analyses is whether those editors engaged in the administrative work of discussing and revising policies are distinct from other social roles on Wikipedia, the membership in this group over time, and whether these editors have any “soft” power over others (Aaltonen and Lanzara 2011). Discussions around the enactment of policies tend to have low levels of meaningful deliberation but the structure of these conversations nevertheless differentiates meaningful roles within the collaboration (Black et al. 2011).

RQ2: How have patterns of user contributions to Wikipedia rules changed over time?

Behavioral consequences of rule-making

Although Wikipedia is often popularly framed as lacking any coherent oversight, Wikipedia’s complex and dynamic environment depends substantially on formal organizational structures. But rule-making and governance are fundamentally exercises in power, re-defining the boundaries of conduct and codifying norms that privilege some approaches over others. The work of creating, supporting, or contesting rules on Wikipedia have substantive costs by distracting talented volunteers from the core mission of writing encyclopedia articles, creating more coordination overhead, as well as introducing potentially adverse incentives for elites to

consolidate their power (Shaw and Hill 2014). Wikipedia’s model of self-governance emulates many strategies employed by successful offline communities (Benkler 2006; Viégas, Wattenberg, and McKeon 2007), but these symmetries are not always flattering. In the face of crises like slowing growth of new editors and content generation (Suh et al. 2009), bureaucratic participation may provide adverse incentives for editors to lobby for self-serving rules rather than to collaboratively write an encyclopedia.

RQ3: How does participation in Wikipedia rule-making alter editors’ behavior?

Research setting

Wikipedia is a peer-produced online encyclopedia that has become one of the largest reference works and most-trafficked websites in the world. In the absence of any hierarchy for recruiting contributors, assigning tasks, or evaluating submitted content, Wikipedia relies on a set of policies and guidelines to “describe best practices, clarify principles, [and] resolve conflicts.”⁵ Like Wikipedia’s encyclopedia articles, pages describing Wikipedia policies can be created, edited, and revised by registered users. As such, Wikipedia policies are not fixed but are subject to constant revision in response to new cases and precedents. Policies play an important role in adjudicating disputes on Wikipedia by appealing to general principles during conflicts and controversies (Beschastnikh, Kriplean, and McDonald 2008; Butler, Joyce, and Pike 2008; Kriplean et al. 2007).

Given this complexity, we limit our analysis of Wikipedia policy ecosystem in several ways. First, we focus on the processes surrounding the making of Wikipedia’s “rules-in-form” rather than examining Wikipedia “rules-in-use” that are pervasive and on-going (Beschastnikh, Kriplean, and McDonald 2008; Kriplean et al. 2007). Written rules leave clearer traces of changes but changes to written rules are likewise subject to different approval, interpretation, and sanctioning processes than unwritten norms (March, Schulz, and Zhou 2000). Far from being stable, restricted, or subject to formal approvals, these rules-in-form are actively contested and revised over time, open to all registered editors to edit, and are subject to consensus-based acceptance. The need to document changing precedent across a distributed collaboration, centrality of these documents in discussions, and on-going need to recruit and socialize newcomers into best practices suggests that these rules-in-form should quickly adapt to reflect rules-in-use.

Second, we examine only the policies, guidelines, essays, and failed policies rather than the broader ecology of algorithms, bots, and templates that also serve as governance mechanisms on the platform (Geiger and Ribes 2010; Müller-Birn, Dobusch, and Herbsleb 2013). These formalized rules are nevertheless central signifiers of the communities’ values and the binding precedent for enforcing compliance with behavioral and content rules across the entire community (Forte, Larco, and Bruckman 2009). Because

⁵Wikipedia:Policies and guidelines

Wikipedia keeps a detailed archive of the who, what, and when of changes made to these policy pages, the evolution of rules can illuminate how governance practices evolve (Butler, Joyce, and Pike 2008).

Finally, we emphasize the changes in these rules over time as an essential dimension for understanding the governance processes on Wikipedia. While many policies were established early in Wikipedia's history, they have nevertheless continued to develop and evolve in response to new practices such as writing about contentious editing (Joyce, Butler, and Pike 2011) or current events (Keegan 2013). There is a "life cycle" to policies in which entire new policies can be proposed for discussion, existing policies are revised in light of new precedents, and policies can also be demoted in the face of changing standards.

Data

The English Wikipedia's policies, guidelines, essays, and failed proposals are identified by membership in their respective categories. Rule pages' membership in these categories is strongly enforced by the use of "templates" that explicitly mark the pages as the specific rule type for readers of the rule. The categories for policies and guidelines contain sub-categories going down multiple levels, but the list of articles used were only extracted to a maximum depth of 1 sub-category down. This produced a set of 62 policy pages (including the pillars), 175 guideline pages, 1,476 essay pages, and 311 failed rule pages. This corpus of 2,012 unique pages includes small overlaps between several rule types, which we ignore in subsequent analyses.

Like many platforms supporting online knowledge collaborations, the MediaWiki software running Wikipedia includes a complete revision history of every change to a page. These revisions document the changes in content as well as meta-data such as the contributor and timestamp. We used a custom Python script employing the python-wikitoools library to retrieve the complete revision histories for rule pages from the English Wikipedia's MediaWiki API. Note that the data returned by Wikipedia's API may omit revisions deleted by administrators that are blatant copyright violations or contain offensive/disruptive material, but these omissions should be rarer on rule pages than on typical Wikipedia articles. This generated a corpus of 265,248 revisions made to all 2,012 rule pages. We repeated the same steps on the respective talk discussion pages for each of the rule pages, which generated a corpus of 460,124 revisions.

Results

RQ1: Changes in rule-level behavior

Research Question 1 asked, "How have rule-making patterns on Wikipedia changed over time?" We can compare the aggregate patterns of policy-making across the rest of Wikipedia's rule ecosystem by looking at changes in the number of revisions, size, and discussion activity. In the subsections below we find evidence of declining revision activity, sustained growth in the amount of content on rules, and a shift toward deliberation over legislation.

Declining revision activity Figure 1 plots the average annual revisions to each of the four classes of rules, stratified by the year the rule was first created. Taking the average annual revision activity to policies on the left as an example, we find evidence of declining rule-editing activity.

First, editing activity peaked between 2006 and 2007 and fell significantly for all subsequent years. The "golden era" for rule-making occurred between 2004 and 2007, which coincides exactly with the era when Wikipedia was expanding most rapidly. Second, older policies (between 2001 and 2004) consistently have the highest levels of editing activity, even after new rules are introduced. Third, new policies across all years attract minimal revision activity initially. However, there is a significant split between pre-2007 and post-2007 rule-making with the latter never "taking off" with an increasing number of revisions.

Similar patterns of peaked activity, early rule dominance, and slow starts play out for guidelines and essays over the same time frames as well, although peak average activity on policies around 2007 was about 50% higher than peak average activity for guidelines and three times higher than peak average activity on essays. All these metrics show general declines in activity for rule-making in more recent years, emulating more general trends in Wikipedia engagement (Suh et al. 2009; Halfaker et al. 2013)

Growing policy size Not all revisions result in similar amounts of content. Figure 2 plots the average annual size of rule articles (measured in bytes of content) to each of the four classes of rules, stratified again by the year the rule was created. Similar trends play out across policies, guidelines, and essays as before. The oldest rules tend to be much longer than more recently-created rules, although all rules tend to start off around the same size. Rules grow at constant rates with no diminishing returns on rule size, but earlier rules grow in size more quickly than more recent rules.

Because older rules are much longer than more recent rules, they exhibit more complexity and potentially greater "quality" than other rules pages. On one hand, these findings are perhaps unsurprising given that older rules have had more time to accumulate changes. But newer rules may also reflect more pressing concerns and provide outlets for editors to exert influence without precedent or coordination costs while the meaning of older rules stabilized. The disproportionate attention to the oldest rules implies they remain active sites of deliberation about contemporary issues — rather than being settled issues — extends prior findings about their influence (Kriplean et al. 2007; Heaberlin and DeDeo 2016).

Sustained deliberative intensity Wikipedia's rule-making process emulates (and perhaps exaggerates) the consensus-formation processes employed on its articles. While these deliberations might play out in a minor way within the revision comments editors leave for each other, each rule's talk page is intended to be the primary site for discussion about potential changes to rules. Comparing the revision activity of policies' talk pages to the policies' pages reveals substantial differences in behavior.

Figure 3 plots the median number of revisions made to

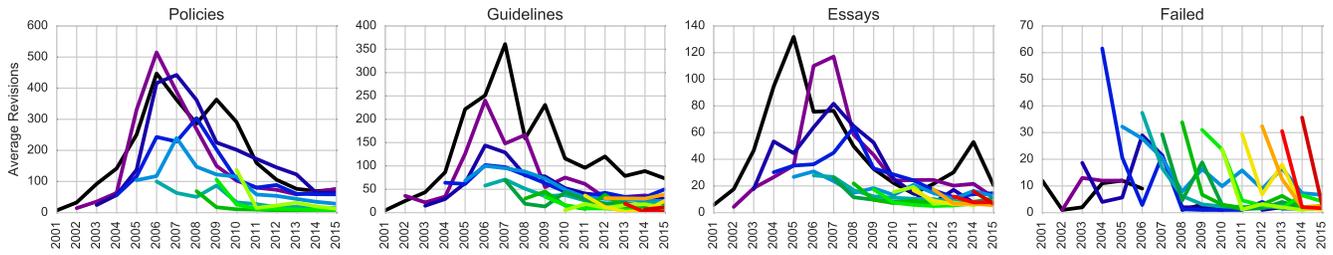


Figure 1: The average revisions to rule pages stratified by year the rule was created for each of the four classes of rules.

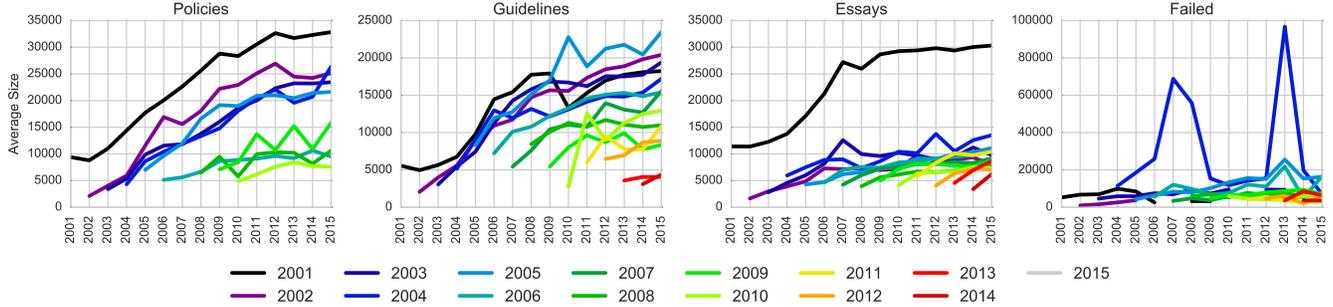


Figure 2: The average annual size for rule pages, segmented by year the rule was created.

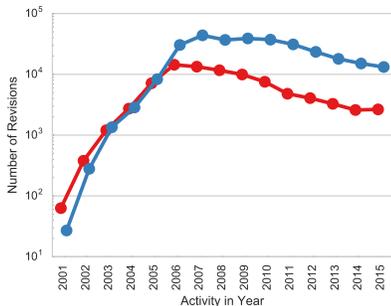


Figure 3: The total number of revisions to the policy (red) and policy talk (blue) pages by year.

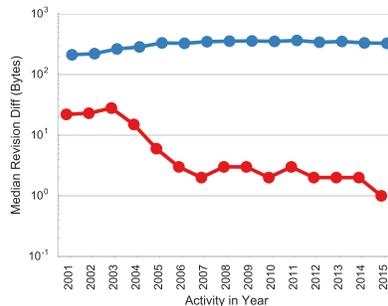


Figure 4: The median content difference for revisions to the policy (red) and policy talk (blue) pages by year.

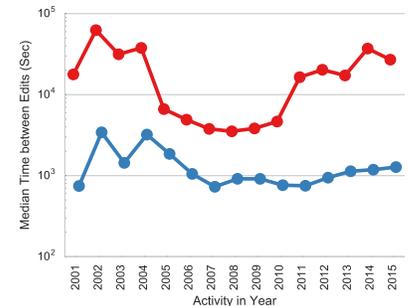


Figure 5: The median latency between revisions to the policy (red) and policy talk (blue) pages by year.

policy pages (red) and policy talk pages (blue) between 2001 and 2015. Two interesting regimes emerge; a tightly coupled growth stage from 2001 through 2006 where the number of changes to policies and their talk pages tracked each other closely. Wikipedia editors follow a “Bold-Revert-Discuss” (BRD) cycle⁶ of proposing changes, having other editors revert them back to a consensus version, and then starting a discussion. These cycles appear to be relatively tight and scaled with each other through 2006. However, from 2007 onward, there is a marked divergence between revisions to rules themselves and discussions about them. Revision activity falls monotonically for both types of pages and talk page-related editing drops off much more slowly than revisions to the rules themselves. This decoupling of page and discussion editing activity suggests a second regime corre-

⁶Wikipedia:BOLD, revert, discuss cycle

sponding to an intensification of policy discussions unconnected to attempts to revise the policies themselves.

The differences between policy-making and policy-taking become more pronounced when we examine other features of the revision logs. Figure 4 plots the median number of changes in the length of policy pages (red) and policy talk pages (blue). The size of revisions to the policy talk page are always larger than the size of revisions to the policy page. This is unsurprising because deliberations invite editors to contribute several sentences at a time while changes to the content of a policy page might be on the order of a few words or phrases. There is a substantial divergence in the behavior of these revision differences over time as policy deliberations involve longer contributions to policy talk pages over time but smaller changes to the policy pages themselves.

Finally, the time elapsed (latency) between successive re-

visions provides a metric to measure the intensity of activity: smaller latencies suggest more intense coordination and editing. Figure 5 plots the median latency for policy pages (red) and policy talk pages (blue). The time between revisions on the policy pages is always longer than on the policy talk pages, reflecting the higher pace of changes as a part of deliberations. The policy page latency demonstrates two interesting regimes of latencies between changes: between 2005 and 2010, the time between policy changes was substantially faster than in the “scaling” era beforehand or the more recent “institutionalized” era. The shift in policy page editing tempo may reflect normative changes to make smaller, faster changes rather than slower, larger changes, but reverted back to small and slow changes by 2011.

RQ2: Changes in editor-level behavior

Research Question 2 asked, “How have patterns of user contributions to Wikipedia rules changed over time?” Using the same corpus of data about revisions made to rules and rule talk pages, we shift the unit of analysis from articles to users. In particular, we perform “time-aware analyses” (Barbosa et al. 2016) of user cohorts and differentiate users by the year they began contributing to rules in the sample, not their first edit to Wikipedia. Across cohorts, we find evidence of inconsistent editor engagement over time, complex overlaps in contributions across the rule ecosystem, and substantial variation in editors’ first rule contributions.

Inconsistent editor engagement We know from Figure 1 that editing activity on rule articles peaked in 2006–2007. But was this activity driven by long-time rule editors responding to the influx of new editors or does it represent new rule editors reframing Wikipedia’s governance processes? Figure 6 plots the average number of revisions made by users to rule pages in the corpus, stratified by the year they began editing. There is strong evidence that the latter is the case: new, not pre-existing, rule editors were responsible for an extremely high level of activity during these crucial years.

But just as quickly as they burst onto the regulatory scene, this activity diminished rapidly over subsequent years. Next year’s cohort of users picked up much of this slack and also made dozens of contributions across rules pages in their first year. However, there’s also a troubling trend toward more recent cohorts making fewer and fewer contributions in their first year of rule editing compared to previous cohorts. We explore the mechanisms behind this dynamic in more detail in the results for RQ3. However, it is not the case that the first rule editor cohorts were responsible for the bulk of Wikipedia policy-making. This “churn” through users across cohorts suggests that initially enthusiastic contributors either become burned out, are driven away, or secure the policy changes they desired and return to editing other pages. However, average rule editing across cohorts does not return to 0 but sustains at a few edits per year.

The average number of revisions by users per cohort does not capture the extent to which editors in these cohorts made contributions across all the rules available for them to edit in a particular year. Figure 7 plots the fraction of rule pages edited by user cohort by year. Corroborating

our findings above, there is a substantial decay in editor engagement year-over-year. Almost every rule page is edited at some point in a cohort’s first year, but cohorts’ contributions across other articles approaches 0.

More troublingly, there is a decreasing tendency for post-2007 editors to engage with the entire policy environment available to them in their first year. While there are fewer rules in earlier years, which makes it easier for all rules to be revised by earlier cohorts, later cohorts also have substantially more users. This finding highlights the risk of new rule editors failing to be fully socialized into the broader rule ecosystem as more rule pages are increasingly overlooked in the crucial first year of rule editing. These patterns of inconsistent editor engagement within cohorts over time and declining first-year policy engagement raises important questions about the legitimacy of rules that new users either are unaware of or are irrelevant to their editing practices.

RQ3: Policy-induced behavioral change

Research Question 3 asked, “How does participation in Wikipedia rule-making alter editors’ behavior?” Are rules a pit stop in an editor’s contribution trajectory to support an editing agenda? A lateral move in an attempt to develop legitimacy and demonstrate competence for administrative promotions? An instance of specializing in policy administration to the exclusion of other editing opportunities? We explore these questions by examining the consequences of editors’ engagement with rule-making.

For each editor contributing to any rule page or rule talk page in our sample, we identified the timestamp of their first revision to articles our sample. We then extracted all the contributions they made across the English Wikipedia for the four weeks preceding and following this first rule edit. For the analyses below, we segmented this corpus of first-rule-edit-centered user revisions into the pre- and post-groups. We observed significant differences in editors’ level and location of activity as well as significant shifts in their sequential transitions between article namespaces.

Changes in behavior and location How did users’ aggregate behavior differ in the four weeks before compared to the four weeks after their first rule edit? Figure 8 shows these pre/post changes for four variables. First, there was a small but significant increase in the average number of page namespaces (described in more detail in the next paragraph) editors participated in after their first revision. Second, there is a moderate and significant decrease in the median size (in bytes) of editors’ average revisions. Third, the average number of unique pages editors contributed to increased significantly in the post-period. Finally, the average number of revisions editors made also increased significantly in the four weeks following their first rule contribution. These results suggest that editors’ participation in rule-making disrupts prevailing editing practices, but our subsequent analyses paint a more nuanced picture.

We expand the revision and namespace results in Figure 9. A namespace is a very high-level way of categorizing different classes of Wikipedia pages. Contributions to articles go in namespace “0”, article discussion pages go to names-

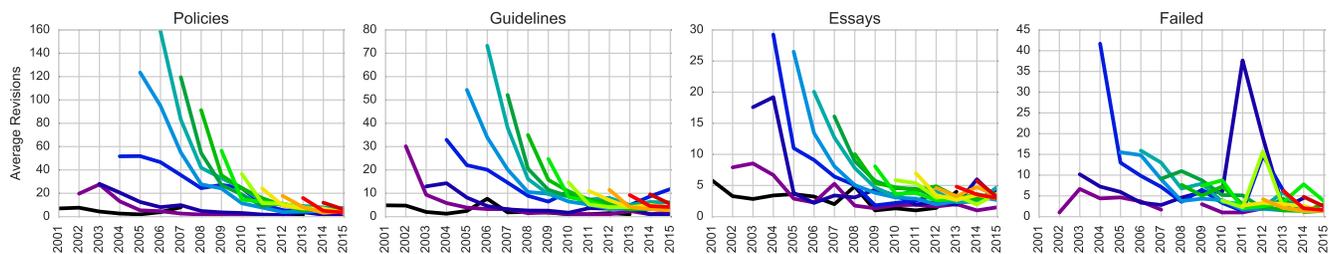


Figure 6: The average annual editor revisions to rule pages, stratified by year the editor began editing rules (colored line).

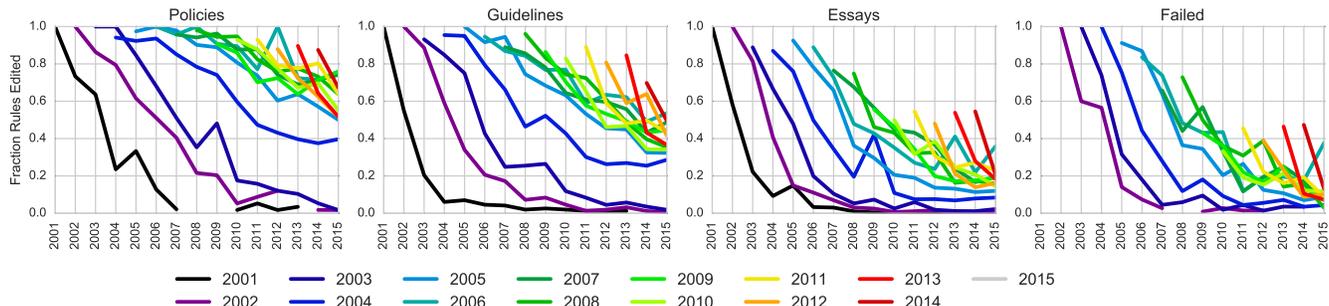


Figure 7: The total fraction of rule pages edited by year, stratified by year the editor began editing rules (colored line).

Subject	Talk	Name	Description
0	1	Main	Articles, lists, & redirects
2	3	User	User pages & sandboxes
4	5	Wikipedia	Policy, essay, & processes
6	7	File	Media file descriptions
8	9	MediaWiki	Auto-generated pages
10	11	Template	Infoboxes, nav boxes
12	13	Help	Software help
14	15	Category	Categorized pages
100	101	Portal	Topics & WikiProjects

Table 1: Namespace descriptions.

pace “1”, user talk pages are in namespace “3”, rules are in namespace “4” and rule discussions in namespace “5”. There are 35 namespaces in Wikipedia, but many of these are infrequently used so we focus on 18 of the most active namespaces. These namespaces are described in Table 1.

Figure 9 visualizes changes in editors’ average revision activity across these 18 active namespaces before and after their first rule edit. There are significant differences in engagement across most of these namespaces. Most namespaces have significant increases in activity after editors’ first rule edit, suggesting that rules editing does not end up cannibalizing editors contributions elsewhere on Wikipedia. Many of these differences are relatively small, but the namespaces with the largest changes over the baseline include user talk pages (“3”), project (“4”) and project talk pages. The latter two are the namespaces to which the rules and rule discussions belong. The large post-hoc increase in activity in the user talk namespace (“3”) suggests rule editing drives them towards increased communication with

other users. Whether this conflict is related to conflicts over their edits or attempts to coordinate activity would require quantitative or qualitative content analysis methods.

It is also important to note that editor contributions activity in the “4” and “5” namespaces were non-zero for the pre-period in this user contribution sample. While all rules and rule discussions are classified in these respective namespaces, other administrative Wikipedia activity like WikiProjects, process requests, maintenance tasks, and directories are also in this space. Thus the non-zero activity before editors’ first rule edits suggests that rule editors are already familiar with Wikipedia administrative processes.

Rule ecosystem transitions Editors working across different parts of the rule ecosystem are potentially engaged in some combination of experimentation and lobbying. Following one mono-causal arrow, editors who attempt to make changes to policy pages and are rebuffed by the community remain motivated and attempt to create new policy that ultimately fails. Following the mono-causal arrow in the other direction, these editors might have proposed a policy and had it fail during deliberations. Undeterred, these users may move their “lobbying” efforts to a policy pages instead.

We look at editors’ revision sequences within the rules corpus to analyze if editors are engaging in arbitrage by either editing policies in reaction to rejected proposals or creating proposals after having policy contributions rejected. Sorting by time and grouping by editors, we can count how often editors’ contributions in the corpus involve transitions between policy, guideline, essay, or failed proposal pages. Figure 10 plots these rule ecosystem transition probabilities for all the revisions made in a year. Self-transitions make up the bulk of the occurrences, so we exclude those to instead

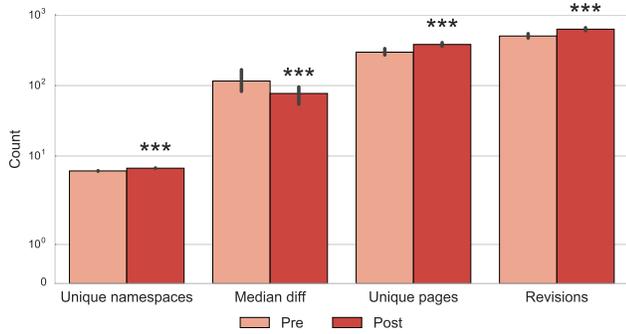


Figure 8: Four measures of user behavior and their counts in the four weeks before (light red) and after (dark red) rule editors' first rule contribution. P-values from t -tests for two related samples are plotted above: *** $p < 0.001$.

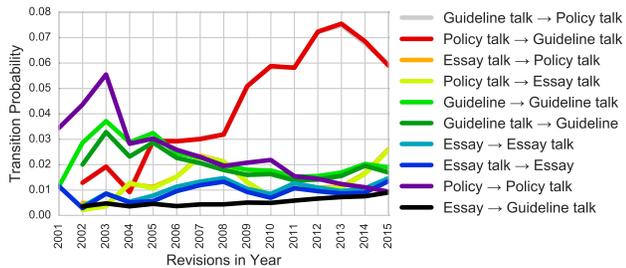


Figure 10: The average probability (y -axis) of editors' sequential contributions transitioning between rule types (colored bars) for revisions made in a given year (x -axis).

focus on transitions between different types of rules.

There is an interesting change in transition probabilities that occurs right around the crucial years of 2006–2007 as Wikipedia growth and rule editing activity peaks (Halfaker et al. 2013). Policy to policy talk and guideline to guideline talk were the highest probability transitions in the early years, potentially reflecting a deliberative mode of making bold changes and discussing them. This dynamic decays over time and post-2007, a new order dominated by editors moving sequentially between guideline talk and policy talk emerges instead. Discussion prompts more discussion as norms shift toward proposing and developing consensus around rules, rather than discussing and then making changes to the rules. By 2015, the four highest probability transitions involve editors moving from one class of rule talk to another class of rule talk. This corroborates other work identifying decreasing direct work on articles and increas-

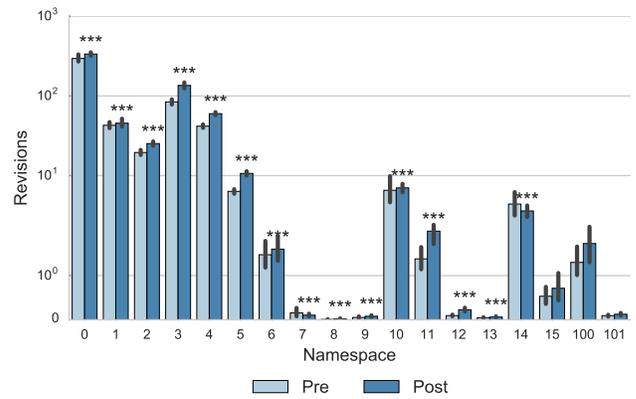


Figure 9: The average number of users' revisions by article namespace in the four weeks before (light blue) and after (dark blue) their first rule contribution. P-values from t -tests for two related samples are plotted above: *** $p < 0.001$.

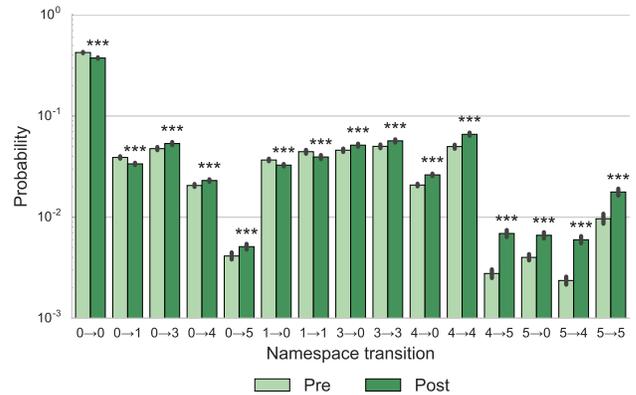


Figure 11: Changes in all rule editors' namespace transition probabilities in the month before and after their first edit to a rule page. P-values from t -tests for two related samples are plotted above: *** $p < 0.001$.

ing indirect work on talk pages (Kittur et al. 2007b).

Namespace transitions We extend the revision sequence analysis approach from Figure 10 to the pre/post user contribution corpus to test how users' tendencies to transition between different namespaces across Wikipedia following their first rule edit. Figure 11 measures the average transition probabilities between namespaces for 15 transitions relevant to our analysis. The differences between the means observed in these transitions are all statically significant at the $p < 0.001$ level using a two-sided t -test for related samples. While we observed that the number of policy namespace revisions increased significantly in Figure 9, did these editors continue to make revisions to articles or is rule-editing correlated with them shifting their contributions to other areas?

Decreases in transition probabilities are observed for many of the article-related namespaces: editing article namespaces sequentially ($0 \rightarrow 0$), editing article talk pages

after articles ($0 \rightarrow 1$), editing articles after article talk pages ($1 \rightarrow 0$), and editing article talk pages sequentially ($1 \rightarrow 1$) all decrease significantly after first-rule editing. If the article namespace is being edited more (Figure 9) but editors are consecutively contributing to these namespaces, where did the activity go?

Namespaces encompassing rules (“4”) and their discussion pages (“5”) are obvious candidates. We find large and significant increases in the transition probabilities for article to rule ($0 \rightarrow 4$), article to rule talk ($0 \rightarrow 5$), as well as their respective reciprocal transition modes ($4 \rightarrow 0$, $5 \rightarrow 0$). The largest increases in transition probabilities are observed transitions between rule namespaces and rule talk namespaces ($4 \rightarrow 5$, $5 \rightarrow 4$) that reflect editors moving between editing and discussing rules. The significant increases in consecutive editing of project ($4 \rightarrow 4$) and the project talk pages ($5 \rightarrow 5$) likewise suggest that contributions to rule pages are not one-off but potentially made cumulatively or in reaction to other editors’ contributions to rules at the same time.

Discussion

Wikipedia’s policy environment has changed substantially over its 15-year history and provides a remarkable benchmark for understanding the potential and pitfalls of self-governance in a knowledge commons (Ostrom and Hess 2007). Our analysis of Wikipedia’s rules found a policy environment marked by on-going rule-making and deliberation across multiple regulatory levels more than a decade after its creation. This dynamism is however balanced by strong biases in the attention and length towards older rules coupled with a diminishing flexibility to change these rules, declining revision activity over time, and a strong shift toward deliberation and discussion.

Rule-making on Wikipedia involves a large numbers of editors collaborating across different rules. These users’ rule-making activity has changed considerably over the history of Wikipedia and is marked by transient editor engagement but also a high degree of overlap in editing membership across the rule ecosystem. The shift towards greater deliberation also manifests in changes in the patterns of Wikipedians’ first edits becoming more focused on talk pages rather than the policies themselves.

There are significant behavioral consequences to rule-making on Wikipedia as well. On one hand, rule-making can be occasions for editors to assemble around their shared values or to disseminate best practices. Reassuringly, rule-making does not lure editors away from productive work like writing encyclopedia articles and towards litigating peripheral issues. We do find evidence of significant behavior changes in overall activity levels as well as distributions of work throughout the encyclopedia once editors begin revising rule-related topics. Examining the distributions of editors’ transitions between editing pages of different types also corroborates earlier findings about a qualitative shift in Wikipedia’s organizational culture during the most acute phases of its growth during 2006–2007.

Implications for theory and practice

These rules face many of the same coordination challenges as its articles in managing distributed co-authorship over time. However, these rules by definition occupy a central and highly symbolic position within the organization because they define its identity, scope its content, and codify acceptable behavior. Prior work has persuasively articulated the importance of these rule ecologies for governing a large-scale social system and knowledge commons. However, the *enactment* and *maintenance* of these structures and how they change over time has received less empirical attention. Motivated by theories of institutional analysis, our research examined large-scale event log data to examine how peer-produced rules emerge and stabilize.

This research has implications for theorizing about the dynamics of rules as organizations increasingly come to rely upon online knowledge collaboration platforms like wikis and repositories to accomplish their missions. Our results suggest that while “rules-in-use” may become stable enough to enable mass collaborations, the “rules-in-form” can remain a site of active revision and deliberation years after their articulation. However, the growth of a community relying on these rules may require a shift in coordination towards a maintenance modality emphasizing more deliberative and iterative changes to rules. Managing these coordination tensions are central to institutional theories but the affordances of online knowledge collaborations like Wikipedia complicate many institutional assumptions about resources, processes, and power (Faraj, Jarvenpaa, and Majchrzak 2011).

Limitations and future work

The structure and dynamics of peer-produced policy offer a compelling empirical setting to understand a variety of social and organizational processes. This paper emphasized the formation and to some extent the governance of policy networks at a macro-scopic level, but qualitative and case study based methods would reveal more specific mechanisms and practices that produce the structural patterns observed. Several policy pages are protected from editing given their legal importance and the effects of rule page protections on editors’ ability to contribute was not considered in this work (Hill and Shaw 2015). Wikipedia’s emerging body of rules and precedent also invites comparative analysis with data from legal systems to illuminate other pressures and trajectories governing peer production systems may face in the long run (Li et al. 2015). Finally, future research might also employ stronger causal inference methods to move beyond descriptive analyses of log data and begin to understand the consequences of different kinds of organizational design levers for this unique mode of peer production.

Acknowledgments

We thank our anonymous reviewers for their detailed feedback and Deborah Keegan for her copy-editing.

References

Aaltonen, A., and Lanzara, G. F. 2011. Governing Complex Social Production in the Internet: The Emergence of a

- Collective Capability in Wikipedia. *Social Science Research Network* (ID 1926138).
- Barbosa, S.; Cosley, D.; Sharma, A.; and Cesar, Jr., R. M. 2016. Averaging Gone Wrong: Using Time-Aware Analyses to Better Understand Behavior. In *Proc. WWW 2016*, 829–841.
- Benkler, Y. 2006. *The Wealth of Networks*.
- Beschastnikh, I.; Kriplean, T.; and McDonald, D. W. 2008. Wikipedian Self-Governance in Action: Motivating the Policy Lens. In *Proc. ICWSM 2008*.
- Black, L. W.; Welsler, H. T.; Cosley, D.; and DeGroot, J. M. 2011. Self-Governance Through Group Discussion in Wikipedia. *Small Group Research* 42(5):595–634.
- Burke, M., and Kraut, R. 2008. Mopping Up: Modeling Wikipedia Promotion Decisions. In *Proc. CSCW 2008*, 27–36.
- Butler, B.; Joyce, E.; and Pike, J. 2008. Don't Look Now, but We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia. In *Proc. CHI 2008*, 1101–1110.
- Faraj, S.; Jarvenpaa, S. L.; and Majchrzak, A. 2011. Knowledge collaboration in online communities. *Organization Science* 22(5):1224–1239.
- Fiesler, C.; Feuston, J. L.; and Bruckman, A. S. 2015. Understanding Copyright Law in Online Creative Communities. In *Proc. CSCW 2015*, 116–129.
- Ford, H., and Geiger, R. S. 2012. "Writing Up Rather Than Writing Down": Becoming Wikipedia Literate. In *Proc. WikiSym 2012*, 16:1–16:4.
- Forte, A.; Larco, V.; and Bruckman, A. 2009. Decentralization in Wikipedia Governance. *Journal of Management Information Systems* 26(1):49–72.
- Geiger, R. S., and Ribes, D. 2010. The Work of Sustaining Order in Wikipedia: The Banning of a Vandal. In *Proc. CSCW 2010*, 117–126.
- Halfaker, A.; Geiger, R. S.; Morgan, J. T.; and Riedl, J. 2013. The Rise and Decline of an Open Collaboration System. *American Behavioral Scientist* 57(5):664–688.
- Heaberlin, B., and DeDeo, S. 2016. The Evolution of Wikipedia's Norm Network. *Future Internet* 8(2):14.
- Hill, B. M., and Shaw, A. 2015. Page Protection: Another Missing Dimension of Wikipedia Research. In *Proc. OpenSym 2015*, 1–4.
- Joyce, E.; Butler, B.; and Pike, J. 2011. Handling Flammable Materials: Wikipedia Biographies of Living Persons As Contentious Objects. In *Proc. iConference 2011*, 25–32.
- Joyce, E. W.; Pike, J. C.; and Butler, B. 2013. Keeping Eyes on the Prize: Officially Sanctioned Rule Breaking in Mass Collaboration Systems. In *Proc. CSCW 2013*, 1081–1092.
- Keegan, B. C.; Lev, S.; and Arazy, O. 2016. Analyzing Organizational Routines in Online Knowledge Collaborations: A Case for Sequence Analysis in CSCW. In *Proc. CSCW 2016*, 1065–1079.
- Keegan, B. C. 2013. A History of Newswork on Wikipedia. In *Proc. WikiSym 2013*, 7:1–7:10.
- Kittur, A.; Chi, E.; Pendleton, B. A.; Suh, B.; and Mytkowicz, T. 2007a. Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. In *Proc. WWW 2007*.
- Kittur, A.; Suh, B.; Pendleton, B. A.; and Chi, E. H. 2007b. He Says, She Says: Conflict and Coordination in Wikipedia. In *Proc. CHI 2007*, 453–462.
- Konieczny, P. 2009. Governance, Organization, and Democracy on the Internet: The Iron Law and the Evolution of Wikipedia. *Sociological Forum* 24(1):162–192.
- Kriplean, T.; Beschastnikh, I.; McDonald, D. W.; and Golder, S. A. 2007. Community, Consensus, Coercion, Control: CS*W or How Policy Mediates Mass Participation. In *Proc. GROUP 2007*, 167–176.
- Leskovec, J.; Huttenlocher, D. P.; and Kleinberg, J. M. 2010. Governance in social media: A case study of the Wikipedia promotion process. In *Proc. ICWSM 2010*.
- Li, W.; Azar, P.; Larochelle, D.; Hill, P.; and Lo, A. W. 2015. Law Is Code: A Software Engineering Approach to Analyzing the United States Code. *Journal of Business & Technology Law* 10:297–374.
- March, J. G.; Schulz, M.; and Zhou, X. 2000. *The Dynamics of Rules: Change in Written Organizational Codes*.
- 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 *Proc. GROUP 2010*, 165–168.
- Müller-Birn, C.; Dobusch, L.; and Herbsleb, J. D. 2013. Work-to-rule: The Emergence of Algorithmic Governance in Wikipedia. In *Proc. C&T 2013*, 80–89.
- Niederer, S., and Dijck, J. v. 2010. Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system. *New Media & Society* 12(8):1368–1387.
- Ostrom, E., and Hess, C. 2007. A framework for analyzing the knowledge commons. In *Understanding Knowledge as a Commons: From Theory to Practice*. 41–82.
- Pater, J. A.; Kim, M. K.; Mynatt, E. D.; and Fiesler, C. 2016. Characterizations of Online Harassment: Comparing Policies Across Social Media Platforms. In *Proc. GROUP 2016*, 369–374.
- Shaw, A., and Hill, B. M. 2014. Laboratories of Oligarchy? How the Iron Law Extends to Peer Production. *Journal of Communication* 64(2):215–238.
- Suh, B.; Convertino, G.; Chi, E. H.; and Pirolli, P. 2009. The Singularity is Not Near: Slowing Growth of Wikipedia. In *Proc. WikiSym 2009*, 8:1–8:10.
- Viégas, F. B.; Wattenberg, M.; and McKeon, M. M. 2007. The Hidden Order of Wikipedia. In *Online Communities and Social Computing*, number 4564. 445–454.
- Zhu, H.; Kraut, R.; and Kittur, A. 2012. Effectiveness of Shared Leadership in Online Communities. In *Proc. CSCW 2012*, 407–416.