

Collaborative Sensemaking during Admin Permission Granting in Wikipedia

Katie Derthick¹, Patrick Tsao², Travis Kriplean², Alan Borning²,
Mark Zachry¹, and David W. McDonald³

¹Human Centered Design & Engineering, University of Washington

²Department of Computer Science & Engineering, University of Washington

³The Information School, University of Washington
Seattle, Washington, USA

{derthick, pattsao, kriplean, borning, zachry, dwmc}@uw.edu

Abstract. A self-governed, open contributor system such as Wikipedia depends upon those who are invested in the system to participate as administrators. Processes for selecting which system contributors will be allowed to assume administrative roles in such communities have developed in the last few years as these systems mature. However, little is yet known about such processes, which are becoming increasingly important for the health and maintenance of contributor systems that are becoming increasingly important in the knowledge economy. This paper reports the results of an exploratory study of how members of the Wikipedia community engage in collaborative sensemaking when deciding which members to advance to admin status.

Keywords: contributor systems, system administration, collaboration tools, sensemaking

1 Introduction

Decision making in an online community is often a difficult process. In large online communities like Wikipedia, administrators are key to keeping the community functioning, and those administrators come from the ranks of regular users. The decision to grant administrative permissions has important consequences for the entire community. In the process of deciding which candidates should be granted administrative permissions, reviewers must consider a user's behavior relative to a set of criteria and come to some reasonably shared understanding of the merits of a new potential administrator. This study examines the collaborative activity of an online community deciding who is given administrative privileges.

Understanding the processes of admin promotion in Wikipedia by its own members is challenging. Prior work has considered the edit histories of contributors to identify the characteristics of strong admin candidates (Burke & Kraut, 2008). A complementary study (Forte et al., 2009) provides an insider's perspective on the process, using interviews with administrators and editors to reveal that the deliberation process has become increasingly difficult and unappealing to candidates. These and other studies (e.g., Leskovec et al., 2010) illustrate that the decision

making is complex, needing to account for diverse data and subject to intense scrutiny. However, researchers have yet to consider how the processes and tools available in deliberations about candidates for administrator status are employed. Our work is intended to help close this gap by considering the role of shared analytical tools in collective sensemaking (Heer & Agrawala, 2008).

2 Method

We conducted an exploratory study to understand the processes, tools, and information that Wikipedians use to decide the outcome of Requests for Adminship (RfA) cases. In our study, we collected and analyzed three sources of data: (1) semi-structured interviews with active Wikipedia editors who participate in the RfA decision process, (2) the archived record of successful and unsuccessful RfAs as maintained by Wikipedia, and (3) a review of the tools that Wikipedians use for viewing and distilling the contributions and activities of an RfA candidate. We designed this multi-perspective view of the RfA process to gain a rich understanding of the interplay among people, processes, and tools in the work of collaborative sensemaking about the promotion of select community members to admin status.

2.1 Interviews

We conducted a series of semi-structured interviews with 10 experienced editors to discover how they think about their online interactions with other Wikipedians, and particularly how they develop their understandings of specific individuals. A primary focus area in these interviews was participation in the RfA process, including typical ways of participating in RfA cases and the use of different information tools (e.g., edit history counters). The interviewees worked from locations across the U.S., so all were interviewed remotely, using either telephone, instant message, or a combination of these methods. Interviews were transcribed and each transcript was coded by multiple team members using an open theme coding process.

2.2 Content Analysis

We reviewed the online discussion transcripts of 6 RfA cases, examining instances in which candidates were approved, in which candidates were denied administrator status, and in which the candidate withdrew. We discussed all the transcripts individually and then identified prominent themes in the deliberative exchanges.

2.3 Tool Review

When an RfA case is presented to the community for consideration, a set of tools are provided on the nomination page, allowing participants to explore different user characteristics and counts of system activities for the admin candidate. See Figure 1

for an example of the basic count- and history-based tools that point to information about all candidates. This group of tools offers a set of links to information on regular and special pages in Wikipedia. The one exception in this group is “count [quick],” which resides on Wikimedia Toolserver.

[\(talk · contribs · deleted · count \[quick\] · logs · block log · lu · rfar · rfc · rfcu · ssp · spi\)](#)

Fig. 1. A set of link-based tools to regular and special pages in Wikipedia that provide counts in different categories related to a given admin candidate’s actions and related status indicators in the system.

This predefined set of tools appears as a block of links on the nomination/deliberation page for each admin candidate. Figure 2 below shows an additional “toolbox” of links available at the end of each RfA nomination. These links all point to tools residing outside Wikipedia itself on the Wikimedia Toolserver. Like the information tools in Figure 1, these tools allow reviewers to consider the candidates’ actions based on count-based criteria, some of which extend beyond Wikipedia itself. We examined this tool set, observing the affordances of each available tool for supporting an editor attempting to arrive at an evidence-based view of a candidate administrator.

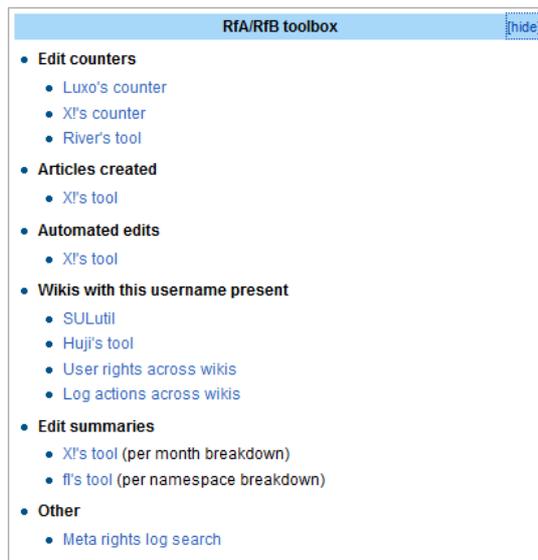


Fig. 2. A toolbox provided with each RfA nomination, allowing reviewers to explore different dimensions of the candidate’s actions and presence in Wikipedia and other Mediawiki projects.

3 Results

To understand how Wikipedia participants decide who should be promoted to admin status, we considered the formal process from the perspective of the participants, from evidence of their deliberative work, and from the system tools used to support the process. Our interviews and reviews of processes and tools yielded detailed insight into the RfA process. These results are presented below with selected quotes to illustrate each of three themes that emerged from our analysis: styles of interaction, social networks, and counts of online activity versus the contexts in which those activities occur. These results are then explored in the discussion section.

3.1 Styles of Acting and Interacting

A primary means by which an online community develops an understanding of its members is by interpreting the actions and interactions of individual participants. A sense of who a person is develops as that individual's behavior is assessed—whether that behavior is related to individual acts performed in the shared space of the system, or whether it is his or her behavior when engaged in dialogic exchanges with others. Within Wikipedia, an editor's actions and interactions are considered important sources of evidence when the community deliberates about the potential promotion of a candidate editor to admin status. Actions and interactions on user talk pages, on article talk pages, in edit summaries, and in private backchannels such as email or IRC chats, become evidence about which the community can deliberate.

When deliberating about RfA candidates, editors seek evidence that an editor has been civil when engaging with others during their work. As one interview explained:

I tend to value civil interactions. . . . Wikipedia has a few people who are good content contributors, but are also possessed of tempers and tend to blow up at people. In [such] case[s], it's weighing value: are their contributions worth the people who might leave because they belittle or harass them, or the environment it creates for others? – Dylan

Editors who engage in controversial issues with a level-headed approach are also viewed favorably, particularly since working productively in a heated exchange is a necessary part of doing administrative work. As one editor explained,

Participation in highly controversial articles will gain people ideological enemies, but at the same time present a very stringent test of the person's ability to deal with complex and emotive issues, interact[ing] with disruptive editors. [These] qualities are often in demand of administrators. – Taylor

Yet another kind of interaction style that is considered when admin candidates are being reviewed is how they have engaged “newbies,” whose initial contributions are sometimes more damaging than valuable. A productive, encouraging interaction style when dealing with new editors is important because the community as whole depends on the infusion of new contributors to conduct the overall work of the system. Admin

candidates, thus, are expected to exhibit an interaction style that encourages new contributors to work productively within community norms. One interviewee used a personal example to illustrate how this interaction style may be exhibited when interacting with a new contributor who appears to be engaged in disruptive editing:

Well, they get the standard warning template(s) that all vandals get, but I'll also add a personal note underneath the template. Nothing too cuddly or kind -- they're disrupting the encyclopedia, after all -- but "I'll grant you, that was hilarious...but please stop before I have to report you to the admins" is the type of language I'd use. What I don't do is advise them on contributing in a constructive way -- I figure that even if they have the means they pretty much need to decide that on their own, as I did. – Bill

As this quote suggests, the ways that an admin candidate has interacted with newbies is scrutinized to ensure that the candidate is oriented more toward enculturating newcomers than in executing the rules of Wikipedia. Interpreting such past interactions, however, is not easy. A strong candidate for adminship must know the system rules and norms, but also find productive ways of interacting with people:

If they've nominated a new editor's bad page for deletion, [that action] is often the subject of scrutiny at RfA. – Dylan

While much of the sensemaking that reviewers engage in while considering admin candidates focuses on arriving at a collective assessment of how productive their interaction styles have been, there is a related consideration about the network of editors within which the candidate is interacting.

3.2 Social Networks in Sensemaking

When developing shared understandings in their deliberations about admin candidates, Wikipedians consider social networks. Participants described how they conceptualized the nature of social bonds, how those bonds are built, and how those social bonds influence the way they come to understand what others are saying.

What Constitutes the Social Network?

The creation of a social network in a system like Wikipedia is about establishing social bonds. While some participants talked about meeting "random" people who just happened to be interested in similar topics, others didn't always see it that way. Some form of homophily emerges when people show up at a topic and begin to express similar ideas or when people come to a Wikipedia policy page and begin to discuss a policy decision. Of the many interactions participants might have, some begin the formation of social bonds, and some people become "friends." When asked about how he defined friendship on Wikipedia, one interviewee explained:

I have real-world friends I interact with on Wikipedia, but by the standard of interactions on-wiki, I wouldn't call them "wiki-friends". "Friendship" online is

held to a lower standard than the real world, so I suppose you could classify a wiki-friend as someone you regularly associate with in a common topic or area of interest, and agree with, or have a rapport with. – Dylan

In Dylan's view, friendships in Wikipedia are based in shared interests, having rapport and seeing things generally the same way. Another participant elaborated on this point to include collaboration on projects and shared work as characteristic of building a social bond:

Every editor has a group of friends they talk with, collaborate with, create projects with, and so forth. My group consists of editors with whom I enjoy speaking, people I agree with on policies, people whose talk pages I watch, people I trust to take the right course of action on matters dealing with policy. – Manny

The comment by Manny reinforces many of the common conceptions of how the social network is built—the idea that repeated interactions and the enjoyment of those interactions are important. Additionally, as his quote indicates, bonds form around having similar ideas about Wikipedia policies. Manny mentions that agreeing on policy results in trust with policy decisions. Manny also mentions using a watchlist to watch for changes in other users talk pages. Watching is an important part of building a stronger social bond. Dylan explained the watchlist concept at length:

I should point out that a lot of Wikipedia relationships are defined by who watchlists what pages. That is, Person A has Person B's user talk page on his or her watchlist, and so when Person C comments on Person B's page, Person A might chime in ... Without watchlists, keeping up to date on communication or changes is reduced to manual refreshing, which inhibits further communication or interaction. So wiki-friends or collaborators often have similar watchlist circles, reinforcing the same types of connections. – Dylan

Many of our participants had the everyday conception of the social network as simply individuals who interact frequently with each other. That knowing of one another through interaction can result in two types of judgments. Given that Wikipedia is big, it is unlikely that an observer can determine the complete social network, which may not be well connected.

Influence of Social Networks

The potentially disconnected nature of a social network has implications for sensemaking in the RfA process. In that process, first the candidate for administrative privileges is nominated. Other editors can ask questions and comment on the candidate. The social network of the individuals who nominate, support, or oppose a candidate is one aspect of how others involved in the deliberation come to understand what a candidate has done and how to interpret his or her responses to questions.

In one interview, Taylor talks about knowing who is nominating and supporting a particular candidate. In the quote Taylor is trying to explain one way to understand a person's reputation. Taylor says:

“Reputation” is hard to qualify or explain, but you start to recognize editors and see how they act and what they say. If you see a candidate that has been nominated by several well-known editors who you have always found to make useful and sensible comments, this will suggest that people whose judgment you trust have a positive opinion about the candidate. This isn’t mentioned in the official criteria, but I’m sure it plays a large part in shaping people’s first impressions. – Taylor

Editors find that a candidate’s circle of friends is useful for understanding RfA candidates: In a much more succinct way this same comment was made by Marshall who says

“To some degree I look at how other editors whose opinion I usually agree with have voted.”

The previous two comments show that both the social network around the candidate *and* the social network around the reviewer matter. The intersection of these social networks is hopefully the people who are most qualified to judge the many aspects of the candidates’ attributes.

While the social network around the candidate and the reviewer are important, Taylor notes that the network is not adequate as the sole basis for a judgment:

This is probably a particularly useful shortcut since Wiki[pedia] is such a large place and the number of people you interact with directly is a tiny fraction of the whole. However, like I say this is a shortcut that may well mislead, so I wouldn’t make a decision based entirely on who supports or opposes, it matters very much what they say and what the candidate says! And of source on what the candidate has done. – Taylor

In addition to social network information, activity counts and the contexts for interpreting them are important in the process of deliberating about RfA candidates.

3.3 Considering Counts and Contexts

Wikipedia editors often strive to live up to a set of shared values, such as Assume Good Faith, as they sense of the behaviors of other editors, particularly when they are making some decision that has consequences for that editor (e.g., when he or she is an admin candidate).

Thinking broadly about an editor, however, can be a difficult information-seeking endeavor considering the vast edit histories that many people have. To address this challenge, participants rely on tools to gain insight into an editor’s history, such as edit counts. Such counts can then be considered in defined distributions across time, articles, or namespaces. Counts thus sometimes perform a thresholding function:

I check their edits. I see what they’ve edited, where they’ve edited. Like if you’ve only got a few hundred edits, you’re likely not completely familiar with

Wikipedia enough to become admin. Similarly, if you've got thousands of edits, but only to a small section, such as only articles about one TV show, or one sports team, or something, you've not had enough interaction throughout the entire site to be familiar with it all – Marshall

In considerations of RfA candidates, raw edit counts are sometimes applied to specific kinds of work, such as consideration of Articles for Deletion (AfDs):

I'll oppose if they don't have a track record in AfDs, for example, because I think it's important for admins to know how to "properly" judge and close AfDs.
– Dylan

Raw edit counts are used to gauge distributions of work across the system:

In general, I like to see a good "mix" in contributions. I generally vote against candidates whose major contributions are to Wikipedia space, instead of articles.
– Dylan

However, while counts can give a baseline sense of the attention and work of a candidate in the system, our interviewees also noted that attention should be paid to *qualities* of actions rather than just their counts. This inclination for contextual interpretation of actions is illustrated in Bill's reflection on his sensemaking habits:

An important thing to consider when interacting with any editor -- particularly one who *appears* to be acting disruptively, like a vandal -- is to look at their contributions. Not all editors do this...frankly, I suspect most don't. But I always do. Sometimes someone who appears to be a vandal, or a provocateur, or just someone who doesn't have good intentions, is actually acting in perfectly good faith. You can infer this if you are about to warn someone for vandalism but look at their user contributions beforehand and note that, say, up until a day ago this person was doing great work... So perhaps what looks like vandalism, or harsh words, or whatever, to *me*... perhaps that's just someone coming across the wrong way. – Bill

Unfortunately, existing tools that Wikipedians use do not assist in the iterative exploration between summations and contextualized action. Even when data is provided that goes beyond numbers, it can be difficult for editors to evaluate whether the data is part of a pattern that accurately characterizes the editor, or an isolated event. This evaluation is complicated by the fact that other editors are often involved in the process of evaluating the editor in question, providing selected diffs as evidence in the discussion forum where the candidate is being considered. A "diff" is a page in Wikipedia that shows the differences between edited versions of a page. The diffs an editor chooses to provide may not be motivated by a principle as respected as Assume Good Faith, where an editor tries to create a false impression of a pattern in another editor. Taylor illustrates this point with a hypothetical case where one editor attempts to persuade others that a candidate exhibits bad behavior:

Imagine somebody says “X is incredibly rude; look at this comment here (link).” You click on the link and it shows somebody saying something that might be interpreted as rude, but you lack all context. It is like seeing everything people say in sound bites rather than conversation. – Taylor

Consequently, he elaborates, the work of principled sensemaking about the candidate requires additional work:

You need to do a lot of detective work. For instance, you can see when the comment was made and search back to find the conversation in a talkpage’s archives. You can then read the context and perhaps find out that they were being sarcastic, or making a joke to a friend. – Taylor

Actions often need to be considered in their context to be appropriately evaluated in a fashion that aligns with the evaluator’s values. But contextualizing actions is difficult in hindsight, particularly when the evaluator was not present when the action was taken. Of course, this qualitative/quantitative tension in large data analysis is a familiar tension for academic researchers.

4 Discussion

Through our study we identify themes that are critical to decision making in the Wikipedia Request for Adminship process. The interview data yielded insight into the different understandings of what community participants should be sensitive to when considering the promotion of someone from their ranks into an administrative role.

What participants in the RfA process consider important in their decision varies notably from contributor to contributor. Although within Wikipedia contributors can review community-produced lists and commentary about what makes a good administrator, there is no requirement that contributors consult such materials when participating. As one participant pointed out, even if such lists or commentaries are consulted, participants in the RfA decision weigh all considerations differently. Further, what a given participant values most in RfA considerations shifts over time. The editors who actively participate in RfA cases (the “regulars”) shifts over time, changing the collective sensitivity to different potential considerations.

Through an analysis of RfA cases and the tools that are used in the process, we identified techniques that reviewers use to focus the attention of others involved in the deliberation on specific evidence to support their interpretations, and how that evidence can be a source of shared understanding and differing opinions.

The deliberation about whether to promote a candidate to adminship is facilitated within the system itself via an asynchronous wiki-based discussion. The deliberation is public and open to participation from the entire community. The candidate is also expected to participate. The candidate answers a set of questions, some of which are standard, some of which are focused on key events in the community and perspectives that the candidate may have about the practices of the community. Following the questions, discussion ensues where participants use evidence to argue for “support” or

“oppose” positions. Sensemaking thus occurs as a collaborative effort of the community as they collectively consider the merits of the candidate.

Sensemaking is currently an artifact-based process in which selected pieces of evidence available in the system are introduced and interpreted by the deliberators. These interpretations are subject to debate as participants decide which evidence is relevant to their decision and what the evidence means. The deliberative forum is seeded with a set of tools for retrieving the most commonly considered evidence in deliberations. These tools include such things as a counter of past contributions in different segments of the system and a history of the candidate’s status since joining the system. As the central means of generating shared artifacts for consideration, these tools play a significant role in how sensemaking occurs in online systems.

One key part of this sensemaking is the shared interpretation of evidence. The technical tools are very effective at identifying “diffs” contributed by the candidate. Isolated contributions rarely tell the entire story. During deliberation, participants use the tools to explain and provide additional details that contextualize interpretations of the candidate’s actions. Some of the more hotly contested deliberations surround discontinuous behavior by the candidate, such as when a candidate has an early history of being a poor member yet changes to be a strong and valuable member.

5 Conclusion and Future Work

Our results suggest a number of important requirements for the design of tools to support collaborative decision-making in large-scale online communities. In particular, systems like Wikipedia suffer from their wealth of activity and behavior data. While each action is recorded by the system, the relationships among interrelated actions are difficult to uncover and interpret. Thus, new tools are needed that can help large-scale communities understand and interpret mass interaction data.

Acknowledgments. This research was supported by NSF award IIS-0811210.

References

1. Burke, M., and Kraut, R. (2010). Taking up the mop: Identifying future Wikipedia administrators. In M. Burnett, M. Costabile, T. Catarci, B. Ruyter, D. Tan, M. Czerwinski, and A. Lund (eds.), *Proceedings of the Twenty-Sixth SIGCHI Conference on Human Factors in Computing Systems*. New York: ACM Press, pp. 3441–3446.
2. Forte, A., Larco, V., & Bruckman, A. (2009). Decentralization in Wikipedia Governance. *Journal of Management Information Systems*, 26, 1, 49-72.
3. Heer, J., & Agrawala, M. (2008). Design considerations for collaborative visual analytics. *Information Visualization* Basingstoke-, 7, 1, 49-62.
4. Leskovec, J., Huttenlocher, D., & Kleinberg, J. M. (2010). Governance in Social Media: A case study of the Wikipedia promotion process. *Artificial Intelligence*, 98-105. AAAI. Retrieved from <http://arxiv.org/abs/1004.3547>