A Quantitative Explanation of Governance in an Online Peer-Production Community

Chandan (Dan) Sarkar,[1] Donghee Yvette Wohn,[2] Kurt DeMaagd [4]
Media and Information Studies
Michigan State University
East Lansing, MI 48824
{sarkarch,wohndong, kdemaagd}@msu.edu

Cliff Lampe [3]
School of Information
University of Michigan
105 S. State St.
Ann Arbor, MI 48109
cacl@umich.edu

ABSTRACT
In this paper, we examine how user ratings of content produced for an online community are taken into account by administrators when they decide whether to delete content. Incorporating about 10 years of server data from the online peer-production community Everything2, we looked at how specific features of voting predicted deletion of posts. We found that not all types of voting are the same: negative voting of users was the strongest factor explaining deletion of a Write-up. Receiving a positive vote from a member with higher status decreases the chances of deletion, while receiving a positive vote from a user with neutral status has a very little effect on the deletion of content.

Author Keywords
Online community; moderation; user voting; feedback; online governance; content editors; decision-making.

ACM Classification Keywords
H.5.3. [Information interfaces and presentation]: Group and Organization Interfaces.

INTRODUCTION
Online peer-productions communities are based on user-generated content and emergent social systems that help govern user behavior in those sites [1]. Wikipedia, Flickr, YouTube and other content-based sites have varying rules and norms about evaluating user-generated content, but in general, users can often provide feedback about the value of the content to those who have the ability to delete that content. Deletion privileges are usually reserved for site administrators, with notable exceptions, like Wikipedia.

Some online communities have a role for users who evaluate and delete content. We are interested in whether or not editors' deletion of articles is informed by feedback provided by other users when such information is available, or if they make decisions independently of that feedback. *Research Question: What is the relationship between user feedback on content and editorial deletion decisions?* Decision-making processes are an integral part of online community governance [1,8]. Understanding these processes and how they are associated with governance structure has broader implications for design, infrastructure, and sustainability for online communities.

COMMUNITY GOVERNANCE
Preece [8] explains that the process of community self governance is basically a process of decision making that involves decision makers (administrators or editors) and policies and norms association with decision making. One of the most important decisions that are made in peer-production communities is the decision on whether or not content should be retained or deleted.

In the context of peer production communities such as Wikipedia or Usenet, researchers have studied how content policies are enforced [3,4]. Kollock and Smith [4] described global and local-level guidelines within Usenet to describe the administrative governance. In these communities, there are a variety of governance structures, ranging from tight control of content quality in the case of most open source projects, to the more distributed deletion authority of Wikipedia [3, 10]. Wiki structure communities such as Wikipedia, also function following the *Iron Law of Oligarchy*— in which power structure and decision making rest on a small group of individuals [5]. Although some scholars argue that Wikipedia policy-making is becoming more de-centralized over time still policy-making is being performed by a core small group of content editors [3].

All these studies show that this small group is responsible for decision-making, but we don't know how their decisions are informed—in particular, if they take into consideration the opinions of the users. If indeed editors' decisions to delete content are informed by general users' opinions, then we would expect that positive feedback from other users would affect preservation of the content while negative feedback from other users would predict deletion.

METHOD
Everything2
To understand how feedback from users in the form of ratings informs administrator deletion behaviors, we analyzed data from the server logs of the peer-production

community, Everything2.com. This site has been studied by the authors in other work [7] and two of the authors have been long-term participant-observers of the site.

Everything2 started in 1999 as an open source encyclopedia, but has evolved over time to include more personal journaling and creative writing. Everything2 is a reasonable case to study for the research question due to following reasons: it has a heterogeneous set of users, a large corpus of articles, and the administrators granted us access to server level data. Everything2 has an active user base of about 3,000 contributing members, and receives approximately 590,000 unique visitors and 1.2 million page views per month as of September, 2011.

The primary contribution of authors on Everything2 is a "Write-up", similar to an article on Wikipedia in that it's an original contribution of content. However, each "Write-up" is solely owned by one author, and only administrators may edit or delete it.

There are three different types of feedback that users can provide to others' Write-ups through the system. The first is page-views, which can be a measure of how much attention the Write-up is garnering. Page-views are a measure of traffic—the online version of people voting with their feet. Although it is ambiguous as to why people are choosing to view that content.

The second type of feedback is a more direct measure of valence: voting. Registered users who have contributed content can rate Write-ups with an Up-vote (positive feedback) or Down-vote (negative feedback). Users are allocated with a fixed number of votes per day and can only vote once per Write-up. These votes are anonymous.

The third type of feedback is one that can only be provided by registered site members who have a certain amount of experience on the site. Known as a "Cool," this is the same as an up-vote in terms of positive valence, but is more exclusive in that there are fewer members (with high status) who can grant it.

While registered users of the site can provide feedback about content, only a small group of users, known as "content editors", can delete content. Users do not have permission to delete their own content, though editors often do so, on request. Consequently, all deletion decisions are made by these content editors.

By default, the system informs authors when their content has been deleted. The administrators can choose to "cloak" a deletion however so the user is not informed. Content editors on the site can also leave a small message as to why the content was deleted, but they have a large amount of discretion as to whether and when to do so.

Analysis

We collected all 36,925 Write-ups that had been deleted from the site over its ten-year history. Server logs were used to select 37,000 non-deleted Write-ups with similar aggregated length distributions of deleted Write-ups. The non-deleted Write-ups were selected randomly using repeated measures through a computer script until a match in aggregated length is obtained between the samples. We then collected the total number of Up-votes, Down-votes, page views, words (length) for each of the Write-ups.

We wanted to see if the three types of feedback factors outlined above contributed to the likelihood of a Write-up being deleted or not. Our research question was explored using the Hazard model (COX proportional Hazard model with time dependent covariates) [2]. This method is similar to binary logistic regression in that it is used to predict a categorical dependent variable, but takes into account the probability that the case would survive. In other words, for our data, Write-ups that are currently not deleted have the probability of being deleted in the future; the model takes into consideration this probability (future censoring). This analysis is used when the dependent variable is affected by time and future likelihood of change [6] and is fairly robust in dealing with time dependent explanatory variables [9]. We only matched length between deleted and non-deleted Write-ups because Hazard model will account the influence of covariates over time on non-deleted Write-ups.

In our case, the outcome variable (Write-ups deleted/ not deleted) is exponential in distribution and associated with censoring, since non-deleted Write-ups could be deleted in the future. Moreover, our covariates (Up-votes, Down-votes, Cools, Page-views) are likely to be affected over time, as the longer the Write-up sustains in the community, the more Cools and Votes it would likely receive over time. As Write-ups have been on the site for different amounts of time, we used rate instead of total number in terms of measuring our independent variables to reduce the influence of time on covariates (to some extent). For example, instead of total Up-votes, we looked at the rate of Up-votes, as it is likely that Write-ups which are posted on the site for a longer period will have increased levels of voting in general. Rate was determined by dividing the total numbers of instances by how long the Write-up had been on the site (lifespan of Write-up). For deleted Write-ups, lifespans (time to events) were measured by subtracting posting time from deletion time. For non-deleted Write-ups, lifespan (time to events) was measured by subtracting posting time from server log captured time. The predictor variables we used in our model are described in Table 1.

The length of the Write-up was added to account for potential proxies for users' responses to a Write-up. Length is likely a signal of quality to users, even though quality will obviously vary. Finally, we included the ratio of positive/negative votes in our model since it may not be the rate of Up-votes or Down-votes but the proportion that may have affected editors' decisions. The outcome variable is categorical (deleted or not deleted). Non-deleted Write-ups include Write-ups not yet deleted but could be deleted in future outside your study time and Write-ups that will never be deleted.

Variables	Explanation
Rate of Up-votes (*M*=.79, *S.D.*=.14)	No. of Up-votes per Write-up /Write-up lifespan
Rate of Down-votes (*M*=.56, *S.D.*=.06)	No. of Down-votes per Write-up/ Write-up lifespan
Rate of Cools (*M*=.14,*S.D.*=.73)	No. of Cools per Write-up / Write-up lifespan
Rate of Page-views (*M*= .05, *S.D.*=.74)	No. of Page views per Write-up/ Write-up lifespan
Votes Ratio (*M*=.007,*S.D.*=.14)	No. of Up-votes /No. of Down-votes
Write-up Length (*M*=42.5, *S.D.*=129.20)	Length of Write-up in words

Table 1: Predictor Variables

RESULTS

Table 2 shows the results of the Hazard Rate Analysis. The EXP(B) column represents the Hazard, or odds ratio. Due to the large sample size, statistical significance should be examined with the effect size as indicated by the odds ratio.

The interpretation of this model is similar to Logistic Regression. In this case, the hazard's ratio for the rate of Page-views is 1.00, which suggests that the variable in the model has no effect on the deletion of Write-ups over time. On the other hand, the Hazard ratio for rate of Down-votes is 1.126, implying that a single unit increase in Down-vote (Rate of Down-vote) will increase the probability of a Write-up being deleted over time by 12.6% {(1.126-1)*100%} at the $p<.001$ significance level). Clearly, a strong correlation is present between Down-votes and deletion of Write-ups. The Hazard ratio for Cools is .301, suggests that for every single unit increase in Cool (Rate of Cools) on a Write-up, the probability of deleting a Write-up will *decrease* by 232%. This suggests that the positive evaluations made by experienced high status users are being taken into more account by content editors when deciding whether or not to delete a Write-up.

The Hazard ratio for Up-votes is 1.005. This suggests that for a single unit increase in Up-vote (Rate of Up-votes) the probability of deleting a Write-up will marginally increase by 0.5%. The result is counter intuitive, but the effect size was very small, so this should be taken into consideration when interpreting the data. One possible explanation could be that content editors did not pay attention to ordinary members' voting behavior for making their decision. The Hazard ratio of Write-up Length is .996, suggests that for every single word increase in length the probability of deleting a Write-up will *decrease* by 0.4%.The chi-square difference between null model and full model is significant at $p<.001$, suggesting that the effects size of the covariates explaining the deletion of the Write-ups event is significant. The coefficients of Rate of Up-votes, Rate of Down-votes and Rate of Cools are more than twice than the corresponding standard errors, confirms the time dependencies of the covariates.

	β	Wald	SE	EXP(B)
Rate of Page-views	.000	.012	.911	1.000
Rate of Up-votes	.005	47.287	.001	1.005*
Rate of Down-votes	.119	66.097	.015	1.126*
Rate of Cools	-1.200	9652.976	.012	0.301*
Votes Ratio	-.760	1.235	.684	0.467
Write-up Length	-.004	897.595	.000	0.996*

**p< .001(Chi Square difference between null and full model is significant at p<.001*, one-degree of freedom(df=1) chi-square distribution of covariates explains time proportionality of the hazard ratio)*

Table 2: Hazard Model Predicting Deletion of Write-up

DISCUSSION

In this paper we show that deletion decisions by editors within a peer production community are being informed by users' action. While there is some relationship between user feedback and editorial deletions, it's clear that editors' decisions are only slightly guided by that feedback. While some editors do use feedback as a guide what to delete, they are not obligated to base their judgments on votes. Positive votes had a statistically significant but practically meaningless effect on deletion. The ratio of Up-votes to Down-votes did not end up being statistically significant, but we still suspect that the relationship between up-votes and down-votes is telling in this process. Page-views had no effect on editorial deletion decisions. The length of the Write-up was statistically related to the survival of a Write-up, but the low coefficients suggest they are not key factors in editorial decisions.

The strongest predictor that a Write-up would not be deleted was the presence of "Cools". Cools are given to Write-ups by experienced members of the site, suggesting that editors placed more importance on the opinions of these experienced, or high-status members. It could be that informing editorial decisions would be helped by applying stronger signals of the rater—such as expertise or credibility—than simple ratings. The study shows that quality of Write-ups has very slight positive influence on editors not deleting a Write-up. In general, while features of the content and the feedback the content has received from other users appears to be informing editorial deletion decisions, those variables are only explaining part of that decision-making process. Content editors are still clearly making autonomous decisions on Everything2 about what content to keep. This may be positive, in that these members are experienced users who may have a better global sense of how to govern content on the site than an aggregate of votes on content could provide. On the other hand, it may create disgruntled users who think their opinions are not reflected in editorial decisions.

IMPLICATIONS FOR PRACTICE

Sustained management of peer-production communities is dependent on positive relationships between administrators and users who provide feedback on the quality of content. If the user-base sees these editors as ignoring their feedback, then they may become frustrated, members may feel that their voices are not being acknowledged. This may lead to a decrease in participation and eventually members may leave the community. Understanding how site administrators incorporate user feedback into their content deletion decisions can inform the development of better tools to communicate that information. For example, “Cools” were strong signals of quality, which may have helped administrators make decisions. Such signals could be added as design elements to be used as feedback by trusted members. Tools that aggregate these types of feedback cues could reduce the efforts of peer-production administrators in incorporating them into their practices.

LIMITATIONS

Although we found some correlations between administrators’ decisions on deletions of user-generated content and user ratings, additional methods of collecting data might like interviews or content analysis would add more depth to these findings. A limitation of our study is that we were unable to look at how other channels of communication such as messaging or chatting may have affected content editors’ decision to delete content. Votes are anonymous. It is not possible for editors to distinguish who made the Vote before deletion. On the other hand Cools are not anonymous. Editors can distinguish whether Cools applied to a Write-up are from editors or higher status members. In our sample, 1116 distinct users submitted Cools. We did not account separately for Cools submitted by editors and non-editors separately as editors are also users and contributors in the community. Never the less this may added slight bias in our results for Cools, though only l.76% of the Cools were from content editors in our sample. In addition, our data are derived from a snapshot of user behavior, thus we cannot determine causality, nor account for users’ and administrators' beliefs or opinions regarding their participation.

CONCLUSION

In this paper we have looked at how voting behaviors as feedback in an online community over a period of 10 years supports the editorial decision making process. This study supports the notion that positive feedback from higher status members positively affects administrators’ decision making about not deleting a Write-up while receiving a negative feedback from an ordinary user significant increases likelihood of deletion. These findings suggest that more nuanced feedback cues and the providers of the feedback are important for online community leaders who need to plan ways to encourage members to participate, and to retain them in their ongoing social structures.

ACKNOWLEDGMENTS

The authors wish to thank the administrators and users of Everything2.com. This work was supported by NSF grant IIS-HCC-0812429.

REFERENCES

1. Ackerman, M.S., Swenson, A., Cotterill, S., and DeMaagd, K. I-DIAG: from community discussion to knowledge distillation. *Communities and Technologies*, (2003), 307–325..
2. Cox, D. R. and Oakes, D. *Analysis of Survival Data*. Chapman and Hall, London, England, 1984.
3. Forte, A., Larco, V., and Bruckman, A. Decentralization in Wikipedia Governance. *Journal of Management Information Systems 26*, 1 (2009), 49-72.
4. Kollock, P. and Smith, M. Managing the virtual commons□: Cooperation and conflict in computer communities. In S.C. Herring (ed.) *Computer-Mediated Communication: Linguistic, Social, and Cross-Cultural Perspectives*, (1996), 109-128..
5. Konieczny, P. Governance, organization, and democracy on the Internet: The iron law and the evolution of Wikipedia. *Sociological Forum 24*, 1 (2009), 162-192.
6. Kraut, R., Wang, X., Butler, B., Joyce, E., and Burke, M. Beyond information: Developing the relationship between the individual and the group in online communities. *Information Systems Research*, In Press (2010).
7. Lampe, C., Wash, R., Velasquez, A. and Ozkaya, E. Motivations to participate in online communities. *Proceedings of the SIGCHI conference on Human factors in computing systems*,(2010),1927-1936, ACM, Atlanta, Georgia, USA.
8. Preece, J. Sociability and usability in online communities: Determining and measuring success. *Information Technology 20*, 5 (2001), 347-356.
9. Singer, J. D. and Willett, J.B. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press, 2003.
10. Viégas, F.B., Wattenberg, M., and Mckeon, M.M. The hidden order of Wikipedia. International conference on Online communities and social computing, (2007), 445–454.