WikiProjects: Collaborative Product Development Teams

PINAR OZTURK AND JEFFREY V. NICKERSON, Stevens Institute of Technology

1. INTRODUCTION

During the last decade, peer production has spawned a number of different products: encyclopedias such as Wikipedia, maps of the world like OpenStreetMap and operating systems such as Linux. Such peer production has changed the ways we exchange ideas, organize communities and create knowledge [Faraj et al. 2011].

The set of self-organized teams called WikiProjects is an example of a peer production environment embedded in the broader ecosystem of Wikipedia. Each WikiProject consists of participants with a variety of motivations, expertise and levels of interest performing tasks designed around shared goals and interests. Hence, they are ideal structures for analyzing the factors that affect peer production because they function as a space for regulating and coordinating decentralized work at the level of the team. In this paper, we examine the interactions and activities of WikiProject participants and how those interactions affect the development process of wiki articles.

2. BACKGROUND

Even though online collaborative product development environments seem to differ from their traditional offline counterparts, they share many parallel attributes. Similar to traditional product development processes [Ramesh and Tiwana, 1999], peer production involves individuals with diverse backgrounds and expertise working together on developing products. And despite their seemingly static nature, these digital peer production environments are also in constant flux due to membership instability and fluctuating participation, resulting in ever-changing team boundaries. The products and routines (wiki articles and policies) developed in peer production environments are realized through collective action of the participants embodying their collective knowledge. On the other hand, unlike traditional product development processes, in a peer production process there is no well-defined organizational structure. While the tasks required for product development are predefined in traditional environments, in peer production the development process is done through evolution and self-organization. Participants choose the tasks they like to work on and the decisions are made in a decentralized manner. Traditional product development processes are optimized by determining the best sequence of activities that promote team effectiveness. Peer production processes are optimized developing the conditions that foster product evolution.

Many previous studies have examined peer production environments, especially Wikipedia. A significant amount of earlier research has focused on editors' motivations to contribute [Forte and Bruckman 2008; Wasko and Faraj 2005; Lampe et al. 2010]. Other research has focused on the internal organization of Wikipedia [Kittur et al. 2007; Butler et al. 2008] as well as the effects of collaboration and conflict on quality [Viégas et al. 2004; Halfaker et al. 2011]. Recent research has emphasized the importance of understanding how the collaboration and coproduction proceeds over time on different peer production platforms [Faraj et al. 2011; Kane et al. 2014]. In this vein, this study focuses on understanding how the interactions and activities of WikiProject participants affect the development process over time.

On WikiProjects, the interactions and activities related to the development process take place at several logical levels. At the root level of the tree are the articles themselves, narratives edited by the efforts of many. At a higher level, editors discuss sets of pages aligned around particular topics on project pages. The project page is a statement of intention for the group, and often includes a membership list, as well as project guidelines and a list of pages it considers of interest to the group. The project talk page is where members document, justify, or discuss changes that they intend to make to a page or set of pages in the group. Here our focus is on the project talk pages and their

relation to actions taken on the article pages. Do the plans and intentions expressed in the pages result in action? If so, how quickly? What factors affect the cycle-time from planning to execution of planned edits? These questions are important because they get at the heart of the mechanisms at work in online mass collaborative product developments.

In order to calculate and analyze the cycle-time, we gathered information on a topic-focused project from Wikipedia as of January 2015. All information from the project talk pages was searched for article URLs, and all the edit information from those identified URLs made by any of the talk page participants were extracted, including their time stamps. By building an event history model, we were able to discover the cycle-time, the difference in time between project talk page article mention and the article edit: the interval between initial planning and task completion.

3. EVENT HISTORY MODEL

An event history model examines how long it takes until the event (action) of interest occurs at a particular time interval (event time). The event of our model is the editing of the wiki articles. The event time is defined as the time lag between a planning discussion on the Project talk page and the edit of the article by one of the discussants. The covariates used in the model are self-assignment, the importance of the article as shown on the project page, the experience of the individual editors and the editors' project membership.

Self-assignment is defined as the editing of an article by the same editor who mentioned it on the talk page. Editors' specific mentions of article links might be an indicator of their interest in specific topics and thus may affect their later article-editing activity. That is, editors mentioning articles may be more likely to go and edit the article themselves, rather than leave such tasks for other editors.

Most WikiProjects rate the importance of the articles within their scope on a 4-point scale ranging from low to high. Since articles that are rated as higher priority might receive faster attention, this variable allows us to investigate whether an aspect of structured project organization, the determination of project priorities, affects article-editing activities.

There are two other potential factors related to the editors. One factor relates to experience. Anyone can edit on Wikipedia, but some editors have much longer track records than others. This experience might influence their editing behavior: for example, more experienced editors might operate at a different pace than less experienced editors.

Not only experience but also project membership might affect editing behavior. Even though declaring oneself a project member is not required in order to contribute to the articles within the projects' scope or participate in project discussions, explicit membership can be an important factor in the development of group dynamics. Group members may help to establish group norms and common repertoires, creating a sense of group identity while working together on joint tasks towards common goals. Hence, adding explicit project membership to our model allows us to investigate whether members' identification with the project creates common goals related to project discussions and influences article-editing activities. Membership in other WikiProjects might also affect the editing behavior of the editors since a Wikipedia article might be covered under multiple projects and hence might be of interest to many development groups.

4. RESULTS

To compare coefficients among variables, we standardized independent variables by transforming the mean to 0 and the standard deviation to 1. Therefore, the coefficients in the table should be interpreted as the effect, where the percentage increase or decrease in editing rate is associated with one standard deviation increase in the independent variables.

Table 1. Event History model *** $p < 0.001$

Dependent variable	Article editing rate		
N	25174		
Number of events	9391		
	exp(coef)	**se(coef)**	**p**
Covariates			
Explicit Project Membership	0.808	0.0339	***
Other Project Membership	0.765	0.0553	***
Editor Experience level	1.002	0.0146	
Self Assignment	2.446	0.0275	***
WikiProject Importance	1.154	0.0069	***

The estimated hazard ratio in the model shows that self-assignment increases the article-editing rate of the editors by a factor of 2.4. In other words, an editor who mentions an article on the talk page is 2.4 times more likely to go and edit the mentioned article compared to other editors participating the discussions. In addition, the WikiProject-determined article importance also increases the article-editing rate by 15%, which suggests that collaborative efforts drive editors to the higher importance articles first.

The model also shows that explicit project membership has an opposite effect on the article-editing rate, where those who identified themselves as project members were 20% less likely to edit a project than non-members. A possible explanation for this finding is that project members spend more time on creating and organizing the organizational routines and procedures on project pages, rather than joining the collaborative efforts in the project talk pages. An analysis of the editing activities on the project pages show that 10 project members produced 208 edits on the project page while 70 other project members produced 156 edits and 201 nonmembers produced 385 edits (Pearson's Chi-Square test, χ -squared=72.69, df=2, p<0.0001).

While the editor experience level did not have any significant effect on the article-editing rate, the experience levels of the members and non-members showed an interesting distribution (Min:2, Q1: 1370, Median 27460, Q3: 30470 and Max: 390200). The members of the project who have edited the article pages all had high experience levels between Q3 and max, but nonmembers had either low experience levels (between Q1 and median) or high experience levels (between Q3 and max). Such a distribution indicates that there are indeed some very experienced non-member editors that edit articles and it could be that they no longer join projects, because they want to focus their attention on other tasks. This would be consistent with research that posits that editors come and go based on their preferred response pattern [Kane et al. 2014]. Thus, editors may have heterogeneous preferences with respect to their response patterns, and the mixture of these preferences yields the distributions seen here.

5. CONCLUSION

Our analysis shows that the assigned importance of the wiki article as well as the editors' self-assignment increases their article editing rate, possibly leading to a faster development process. While explicit membership declarations of the editors have a negative effect on their article editing rate, their activities on project pages show that they focus more on creating and organizing the routines and procedures on project pages. This distinction might be representative of different roles in traditional product development teams. While team members may be more interested in performing their assigned tasks, managers may be more interested in creating a productive environment for product development through setting up routines and schedules.

Acknowledgements
This material is based upon work supported by the National Science Foundation under grants IIS-0968561, IIS-1211084, IIS-1422066, and CCF-1442840.

REFERENCES

Butler, B., Joyce, E., & Pike, J. (2008). Don't Look Now, but We'Ve Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* CHI '08 1101–1110.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge Collaboration in Online Communities. *Organization Science* 22(5) 1224–1239.

Forte, A., & Bruckman, A. (2008). Why do people write for Wikipedia? Incentives to contribute to open–content publishing. *Proceedings of 41st Annual Hawaii International Conference on System Sciences (HICSS)* 1–11.

Halfaker, A., Kittur, A., & Riedl, J. (2011). Don't Bite the Newbies: How Reverts Affect the Quantity and Quality of Wikipedia Work. *Proceedings of the 7th International Symposium on Wikis and Open Collaboration* WikiSym '11 163–172.

Kane, G. C., Johnson, J., & Majchrzak, A. (2014). Emergent life cycle: The tension between knowledge change and knowledge retention in open online coproduction communities. *Management Science*, *60*(12), 3026-3048.

Kittur, A., Chi, E., Pendleton, B. A., Suh, B., & Mytkowicz, T. (2007). Power of the few vs. wisdom of the crowd: Wikipedia and the rise of the bourgeoisie. *World wide web* 1(2) 19.

Lampe, C., Wash, R., Velasquez, A., & Ozkaya, E. (2010). Motivations to Participate in Online Communities. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* CHI '10 1927–1936.

Ramesh, B., & Tiwana, A. (1999). Supporting collaborative process knowledge management in new product development teams. *Decision support systems*,*27*(1), 213-235.

Viégas, F. B., Wattenberg, M., & Dave, K. (2004). Studying cooperation and conflict between authors with history flow visualizations. *Proceedings of the SIGCHI conference on Human factors in computing systems* 575–582.

Wasko, M. M., & Faraj, S. (2005). Why should I share? Examining social capital and knowledge contribution in electronic networks of practice. *MIS quarterly* 35–57.