

Coordination and Beyond: Social Functions of Groups in Open Content Production

Andrea Forte
Drexel University
aforte@drexel.edu

Haiyi Zhu
Carnegie Mellon University
haiyiz@cs.cmu.edu

Niki Kittur
Carnegie Mellon University
nkittur@cmu.org

Amy Bruckman
Georgia Institute of Technology
asb@cc.gatech.edu

Vanessa Larco
Playdom
vanessa@gatech.edu

Robert E. Kraut
Carnegie Mellon University
robert.kraut@cs.cmu.edu

ABSTRACT

We report on a study of the English edition of Wikipedia in which we used a mixed methods approach to understand how nested organizational structures called WikiProjects support collaboration. We first conducted two rounds of interviews with a total of 20 Wikipedians to understand how WikiProjects function and what it's like to participate in them from the perspective of Wikipedia editors. We then used a quantitative approach to further explore interpretations that arose from the qualitative data. Our analysis of these data together demonstrates how WikiProjects not only help Wikipedians coordinate tasks and produce articles, but also support community members and small groups of editors in important ways such as: providing a place to find collaborators, socialize and network; protecting editors' work; and structuring opportunities to contribute.

Author Keywords

Open Content, Peer Production, Wikipedia, WikiProjects, Groupwork

ACM Classification Keywords

H.5.3 [Information Interfaces]: Group and Organization Interfaces – Collaborative computing, Computer-supported cooperative work, Web-based interaction

General Terms

Human Factors

INTRODUCTION: THE CHALLENGES OF OPEN CONTENT DEVELOPMENT

Organizing hundreds of thousands of volunteers in a single project is hard. Anyone who's collaborated on even a small group project like a conference paper knows that sharing a workbench can quickly become complicated. Eric Raymond's description of "a great babbling bazaar of differing agendas and approaches" that yields a seemingly

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CSCW'12, February 11–15, 2012, Seattle, Washington, USA.
Copyright 2012 ACM 978-1-4503-1086-4/12/02...\$10.00.

miraculously stable community of developers has become the quintessential characterization of open source development [26]. Understanding the bazaar as a complex social space has kept researchers busy for over a decade: how do people organize themselves to accomplish feats like open source and open content? In this paper, we examine open content production in Wikipedia from the perspective that organizing volunteer collaborators involves more than getting people to pull in the same direction, it also involves supporting the well-being of workgroups and individuals.

One set of challenges for self-organizing communities involves meeting production goals by optimizing volunteers' efforts and allocating tasks. Volunteers need to know what needs to be done in order to align their interests and expertise with the needs of the community [5, 11, 15]. This is often referred to as the problem of coordination. Another set of challenges involves making the experience interesting, pleasant and beneficial to contributors to attract and sustain their participation. [18, 23, 25]. This is often addressed by speaking of motivation and incentives. Still another set of challenges involves supporting workgroups who assemble to accomplish a particular set of tasks, for example, through upkeep of morale and trust within a team [29]. Crowston et. al. integrate all of these perspectives in their review of literature on Free/Libre Open Source Software (FLOSS) development. We likewise integrate these perspectives with a focus on Wikipedia. In this paper, we examine the role of nested organizational structures in supporting these varied aspects of community health in Wikipedia by not only serving as a mechanism for orchestrating article production, but also by supporting sub-communities and individual community members.

RELATED WORK: WHAT'S GREAT ABOUT GROUPS?

Nested organizational structures are an important feature of large self-organizing communities [24]. Examples of nested structures include local chapters of large international organizations like the Red Cross or Wikimedia Foundation. Ostrom explains that such structures arise in self-organized resource regimes to

facilitate decentralized decision making; when organizations get big, nested structures allow people whose work is directly affected by decisions to continue playing a role in making them. Forte et al. identified WikiProjects as an example of Ostrom’s “nested organizational structures” in Wikipedia [8]. WikiProjects are groups of editors who manage the production of Wikipedia articles related to specific topics and themes. Once a writing project becomes as large as Wikipedia, it becomes difficult to create policies and maintain social norms—particularly editorial conventions—that apply equally well to all the diverse areas of the site [8]. Setting editorial guidelines for writing about pharmacology topics requires different considerations than, for example, writing about marine life, Catholicism, sexuality, or Star Wars. Decentralization allows those with relevant expertise to make decisions about how to proceed with local affairs. We ask:

Q. What are the mechanisms by which WikiProjects and their internal structures facilitate specialized work in Wikipedia?

Our analysis was influenced by Joseph McGrath’s typology of group functions. This typology underscored the importance not only of production activities that target the tasks that need to be done, but also of activities that maintain the health of the group itself and support its individual members [20]. McGrath’s typology juxtaposes three group functions with four modes of operation (Table 1). Jonathan Grudin has observed that the typology is useful for understanding computer-mediated group work because it can help researchers and designers not overlook important group functions: “A rational approach to supporting project activity is to ask ‘what are these people trying to do and how can we help them?’ Unfortunately, this leads to an exclusive focus on the Execution mode of the Production function: Performance” [9]. By focusing attention narrowly, designers and researchers run the risk of misapprehending failures and successes.

	Production	Group Well-Being	Member Support
Inception	Production demand and opportunity	Interaction demand and opportunity	Inclusion demand and opportunity
Problem-Solving	Technical problem-solving	Role network definition	Position and status attainments
Conflict Resolution	Policy resolution	Power and payoff distribution	Contribution and payoff distribution
Execution	Performance	Interaction	Participation

Table 1: McGrath’s typology of group modes and functions (McGrath 1991)

Of course, production is an important aspect of group work and a number of studies have examined the ways

that Wikipedians coordinate production tasks. For example, Kittur et. al. suggested that better task routing tools could help volunteers more efficiently direct their efforts [11]. Cosley et. al. developed an intelligence task routing tool for Wikipedia called SuggestBot [5]. Krieger et. al. followed up on this work by interviewing Wikipedians to understand how individuals in different roles manage tasks and, based on their findings, prototyped a tool called WikiTasks, to help facilitate better task coordination [15]. From this perspective, coordination is overhead that can be reduced to economize volunteer time. This is a rational response to Viegas et. al’s findings in 2007 that an increasing proportion of editing activity was taking place on areas of the site dedicated to coordination activities, not on article composition itself [31].

Kriplean and Beschastnikh et. al. took a different approach to understanding coordination; they analyzed discussions to understand how policy is used to support coordination activities in Wikipedia [1, 16]. From their perspective, the work of coordination is critical to the process not only of producing content and arriving at consensus, but also provides opportunities for volunteers to interact, define roles and use power relationships, which are important features of group health. Kriplean et al. also examined mechanisms by which Wikipedia editors identify valued contributions through awarding one another barnstars [17]. Although it does not directly serve the goal of writing encyclopedia articles, awarding barnstars serves the function of member support. These perspectives are aligned with Grudin’s observation that “non-production functions are not directly tied to a group’s task, but in the long run they contribute to accomplishing it” [9].

In this study, our findings complement and build on the existing literature on Wikipedia and peer production by positioning WikiProjects as nested organizational structures engaged in all three of McGrath’s functions, including not only production activities, but also maintaining the well-being of the group itself and supporting its members. McGrath characterizes groups as “complex, intact social systems that engage in multiple, interdependent functions, on multiple, concurrent projects, while partially nested within, and loosely coupled to, surrounding systems” [20]. This corresponds with our characterization of WikiProjects as nested organizational structures that exist to address specific needs within the broad goals of the English-language edition of Wikipedia.

METHODS AND PARTICIPANTS

Our work is based on the premise that social systems cannot be understood in isolation from the interpretations of the individuals who create those systems. We chose to begin with an interview-based approach to obtain a rich

understanding of how members experienced WikiProject membership. We then followed up with a quantitative examination of editing activity on the site to understand how interviewees’ perceptions might be made manifest in electronic trace data. This mixed methods approach reflects a phenomenological view of sociology, which suggests that social systems are reflexively constructed based on people’s understanding of them [28, 32].

In our first round of interviews, we sought to understand governance in Wikipedia as a site-wide phenomenon. As such, we interviewed eleven individuals who had been involved in the site for many years, in many different capacities (See Tables 2 and 3 for roles and experiences of interviewees). Interviewees were recruited via public postings on mailing lists, by identifying some of the earliest posters on policy pages, and via snowball recruitment, in which participants recommend other informants. These interviews provided a rich description of organizational structures that had evolved in the community over time, such as the arbitration committee and WikiProjects. The theme of decentralization emerged in all of our interviewees’ experiences and it became clear after several interviews that WikiProjects had become important organizational features of the site.

In our second round of interviews, we focused mainly on one WikiProject in order to understand them more deeply as local organizational structures. How do they work? What role do they play in supporting the goals of the community? How do they affect the experiences of members? We chose to focus on the WikiProject Military History because it is a particularly successful, long-standing project, has an active and prolific membership, and has a mature set of editorial guidelines. To recruit a strategic sample of interviewees, we posted messages on the WikiProject discussion page, identified the earliest contributors in the project, and again used snowball recruitment. Our nine follow-up interviews were with Wikipedia editors who were active in Military History, including founding members of the project, people who hold administrative positions in the project, and individuals who edit articles as part of the project without playing explicit leadership roles. In all, we interviewed 20 individuals, fifteen of whom were active in one WikiProject or another. All of these interviews were conducted on the phone or in person and recorded, with the exception of one interview, which was conducted via an exchange of emails.

Normally, research participants’ identities are held in confidence and interview data are anonymized to protect interviewees’ privacy. Questions about anonymity arise when participants have made significant contributions to a public project and desire credit [3]; moreover, public figures like Jimmy Wales are impossible to anonymize. With consent of participants and permission from our

Institutional Review Board (IRB), we have identified Wales by name and use “light disguise” (Ibid.) for other Wikipedians.

Participant	Months Active	Approximate Edit Count
1	64	47,500
2	28	18,500
3	67	20,500
4	33	10,000
5	57	7,200
6	58	54,200
7	76	74,500
8	73	31,500
9	33	88,000
10	64	18,500
Jimmy Wales	88	3,800
12	54	14,000
13	68	1,500
14	22	300
15	3	few
16	11	unknown
17	26	12,000
18	8	unknown
19	53	5,000
20	36	57,500

Table 2: Participant editing activity in Wikipedia at time of interviews

Once the full set of interviews had been transcribed, they were iteratively examined to identify emergent themes using a grounded approach [30]. Initial affinity clusters of data were discussed by two researchers and examined by a third to assess the credibility of the interpretation. Successive iterations were completed by one researcher to compare and refine central concepts. McGrath’s work provided a conceptual vocabulary or *sensitizing concepts* for analyses of qualitative data, but was not used as a theoretical proposition against which to “test” our data. Blumer describes the role of sensitizing concepts in qualitative work in contrast to definitive concepts: “definitive concepts provide prescriptions of what to see, sensitizing concepts merely suggest directions along which to look” [2]p. 7]. McGrath’s theory sensitized us to ways that participation in Wikiprojects may address many different needs for individuals, groups and for the Wikipedia project as a whole.

The qualitative analysis was used to inform a series of targeted quantitative analyses that will be presented alongside qualitatively derived narratives throughout the next sections. The quantitative analyses are based on archival data from 379 projects from 2001 to 2008.

This mixed-methods approach has several benefits. Our qualitative data provides rich and grounded evidence of the importance of Wikipedians' activities in WikiProjects, but has limitations of generality and representativeness. Meanwhile, our quantitative data has the advantage of summarizing behavior on a large scale, but does not provide data on whether the activities so described are truly meaningful influences on members' behavior. By combining these approaches we provide converging evidence that has the advantage of both rich, meaningful interpretations as well as large-scale tests of representativeness and generality. Although the mixed-methods approach does not preclude alternative hypotheses, the convergence of approaches can help us be more confident that the phenomena we are studying are both real and widespread.

Wikimedia Foundation founder, Jimmy Wales
Arbitration Committee member/former member (6 participants)
Regular user / No special technical or social designations (5 participants)
Involved in Wikimedia Foundation (3 participants)
Developer of MediaWiki software (1 participant)
Involved in WikiProjects (15 participants)
Users with access privileges including: administrator, checkuser, oversight, bureaucrat, steward, developer (10 participants)

Table 3: Wikipedian roles represented among participants at time of interview

THE STORY OF WIKIPROJECT MILITARY HISTORY: A COMMUNITY WITHIN A COMMUNITY

In interviews, we found that scaling the community from fewer than 10 members in January 2001 to hundreds of thousands of editors in 2011 [34] involved the development of decentralized governance structures called WikiProjects:

As [Wikipedia] gets larger it's kind of hard to know everyone who's there, so... people try and find new ways of dealing with issues when they don't know everything that's going on... WikiProjects are formed to kind of focus on articles in a particular area. And they can develop policies that just relate to their area. So it enables people to still work together within a small community even though Wikipedia itself is a really huge community. (I1)

The fact that it's gotten so big, it's hard—people can't keep up to date with everything that happens anymore. So in a sense I think the growth of WikiProjects has been partly a reaction to that. In that, you know, you no longer feel part of a community of thousands and thousands of people. So people set up smaller communities. (I10)

WikiProjects as an idea appeared early in the history of Wikipedia but have become increasingly important features of the organization. Kittur et. al. showed that by late 2007, over 1% of edits on Wikipedia were made to WikiProject pages – note that these are not edits to articles that are within the scope of projects, but edits to the pages on which WikiProjects themselves are managed [13]. WikiProject Military History was one of the earliest projects. It was created in October of 2002 with the title of WikiProject Battles with the goal of standardizing all the battle articles on Wikipedia. “Specifically,” remembers one of the earliest members, “we wanted to add an information box to those articles.” (I13)

WikiProject Battles was a small, fairly quiet project for a few years. Then in mid-2005, one user adopted the practice of placing banners advertising the WikiProject on the talk pages of battle-related articles. This drew in more participants and the project quickly grew. In October of 2005, WikiProject Battles merged with WikiProject Wars to form WikiProject Military History. Starting in 2006, WikiProject Military History began electing members into the role of Coordinators:

I think originally we started with a total of 3 [coordinators]... I think on the second election, there were 5, on the third election there were 7, on the fourth election there were 9. (I20)

Just as Wikipedia spawned nested structures to manage local interests as it scaled, Military History has created internal structures also. Task Forces are still smaller, less formal groups within the WikiProject that focus on a specific topic or effort—today, a goal such as the one WikiProject Battles started out with would be taken on by a Task Force. Task Forces benefit from the resources of their parent project and are still able to function as an independent group.

They have task forces by eras—divided up into time eras, WW1 and WW2 and so on and so forth and then they also have special task forces that deal with weapons and weaponry and fortifications and those types of things. (I18)

The Task Forces are pretty much autonomous, but they use the infrastructure that the central project has set up like review processes and templates and announcements. (I20)

WikiProject Military History has 1170 active members, 673 members who are no longer active (have not edited in three months), and 750 Featured Articles (these statistics are manually maintained at [33]). In his 2007 *Guide to Military History on the Internet*, Simon Fowler rated Wikipedia as the best general source for military research. Of the military pages, he said: "The results are largely accurate and generally free of bias." Wikipedians and knowledgeable outsiders perceive WikiProject Military History as a success.

FINDINGS: SUPPORT MECHANISMS THAT GO BEYOND PRODUCTION

Iterative coding of interview transcripts resulted in a set of categories that primarily describe dimensions of support that WikiProjects provide their membership as well as counterexamples of failure to provide support. Interviewees' descriptions of their WikiProjects experiences fell into four primary categories of *support mechanisms*: A) helping editors find expert collaborators, B) providing networking and socializing opportunities, C) structuring participation, and D) protecting editors' work. In the following sections, we will discuss how these perceived support mechanisms contribute to different modes of both production and non-production functions of groups. Support mechanisms do not map perfectly onto production or non-production functions, rather, they may support multiple modes and multiple functions. This discussion will be informed both by interviewees' experiences and by data collected in complementary, targeted quantitative analyses.

Getting Help, Finding Collaborators

One important support mechanism associated with WikiProject Military History is that it allows members to find help and expert collaborators when they need it. Instead of foundering alone with an article, writers of military history articles who need help know that there are interested experts to whom they can turn.

The thing about Military History is that you could go there if you had a question and say "Does anyone know anything about-" and more often than not, you could find someone that knew about it. (I16)

People can kind of specialize—in War Film, for example, there is a task force called War Film so people who have an expertise in this particular area can work on articles over there, of their expertise, while still being able to interact with people with other expertises [sic]. (I18)

That's the thing about Wikipedia, is that the more involved you get in it, the more you realize the sort of back alleys, how it works... a more sophisticated user can direct you to resources like WikiProjects and Task Forces that can help you find other users to collaborate with. (I19)

In terms of the member support function of groups, finding skilled and interested parties with whom to collaborate appears to be an important part of sustaining individuals' interest in participating in WikiProjects.

An important characteristic of Military History that distinguishes it from other projects is its number of active editors. A critical mass of editors may be required for a project to "ignite" and become an engine for sustaining the participation of its membership:

[WikiProject] Archeology is not organized and History is a bit organized, but it doesn't really work. And yeah therefore, I mainly work in Military History.

Interviewer: When you said it doesn't really work, what does that mean?

It's not a place you can go and get help if you have a problem. You want to improve an article, you want somebody to review an article and give you hints on what could be done better and for these issues you need at least a minimum number of active members. (I17)

While the above quote reflects the experience of a single interviewee, it suggests that without a certain number of people involved, a project may lose the critical mass it needs to maintain itself. Many projects have become inactive over the years as founders fail to engage new recruits or lose interest themselves. Here we apply a mixed-methods approach building on qualitative interview data to identify hypotheses that we can quantitatively measure as well.

In many online communities activity follows a power law function, in which a small number of communities are highly active but the vast majority have little to no activity. For example, out of the thousands of projects created in SourceForge, only 10% have more than three members [27]. Power law distributions have been shown to be a robust way to describe phenomena ranging from the network of Internet links to the growth of animal populations [7, 22]. In particular, truncated power laws (in which two different power laws are joined by a breakpoint) have been used to identify sub-processes that may affect how communities evolve [21, 35]. We hypothesize that if indeed there is a "critical mass" needed for a community to thrive, we should see it as a deviation from the standard power law distribution. In other words, we might expect a standard power law distribution for projects which have achieved a critical mass, but fewer projects than we would expect below that critical mass, as such projects would not survive for long (e.g., if newcomers decide to join other projects instead).

To answer this question, we examined the activity of each project as measured by the average number of edits to project pages per month (i.e., the pages where most of the coordination and discussion about the project itself happened, rather than the articles themselves). Figure 1 shows the resulting distribution. As is evident from the figure, this is not a standard power law distribution. There is no "long tail"; instead, there are many projects that have a moderate amount of activity, but very few projects with very low activity. This suggests that there may be forces at play that reduce the likelihood of low-activity projects, for example, if editors preferentially join higher-activity projects because of expectancy-value judgments

as suggested by interviewees (e.g., expecting that the project will not provide them with significant benefits) or cognitive issues (e.g., being overloaded with choices such that the likelihood of seeing low-activity projects is low [7]). The “elbow” of the distribution appears to be around 5-10 edits/month, suggesting that even relatively low but consistent activity may be enough to keep a project alive.

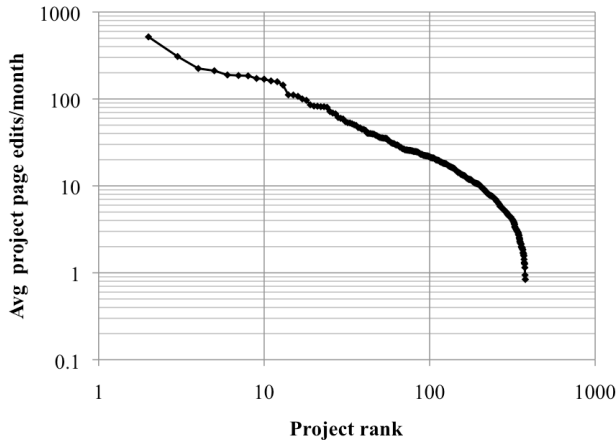


Figure 1: Distribution of project activity (avg edits per month) across a sample of 379 projects. MilHist is the leftmost (most active) project. (WikiProject Spam has been removed from this graph as an outlier, (~5000 edits/month) as it is used as a place to report spam.)

Networking and Social Opportunities

In addition to providing a place to find expert collaborators, WikiProject Military History creates a local work environment where editors get to know one another. Kollock observes that, in online communities, identity persistence and the ability to identify (and assess) past actions are critical factors in motivating collaboration and help members feel that their contributions will someday be reciprocated [14]. Working with people whom one knows, trusts, and whose strengths and weaknesses are known is a different experience than working with strangers who share a common goal. This sense of contribution and reciprocation is an aspect of member support.

If somebody prominent in the Civil War area contacted me and said “Hey, I like your article, but here are some comments on improving it,” I would whole heartedly be excited to get that sort of feedback, but when some anonymous person on the internet comes up with his advice to me on why he thinks it should be improved, I don’t consider their advice to be any more valuable than my own. (I12)

It also helps resolve the group-support problem of role-network definition [20], for example by allowing the group to identify individuals who are able to fill important organizational roles.

WikiProjects, these conglomeration of people, it is a point for networking, meeting a person in networks—you get to know an editor, for example by reviewing these articles and next time you have an article, you ask him for example to help on an article you want to improve—these projects do enhance the ability to form such networks. (I17)

The cooperation with other coordinators and with other editors went very well and I quite liked it. (I6)

In WikiProject Military History, there are also explicit opportunities for members to have fun with the editing process. Organized events such as contests and awards have been touted as community-building activities [10] and help create a sense of group identity and promote group health.

We have monthly newsletters that go out with things like project news and announcements, we have a number of award programs that we maintain to recognize editors that have done various useful things or have written good articles and so forth, we have announcement templates of a variety-some for individual task forces others for the project as a whole that editors can put on their user pages and keep track of everything that is going on and we have periodic special events or drives where there’s some large task that we need to get accomplished so we actually go out and invite editors personally to come and participate. (I20)

Setting up these kinds of activities is precisely the kind of organizational “overhead” that happens on coordination pages of the site. Yet, we see that it is these activities that help create a sense of belonging and support the group and its members. One WikiProject Military History member noted that it was the level of coordination activities that kept him interested in participating in the project:

I just kind of fell into the Military History group because they seemed to me to be the most organized group out there.

Interviewer: And what made you feel that they were the most organized? What is it about them?

Well they have that they are very active-if you go to their Wikiproject page they’re (laughter)-they’re very organized-you know with the style guides and writing articles-they have contests with the bronze star award and those kinds of things. (I18)

Moreover, working in smaller social groups affords opportunities for learning about how Wikipedia works from supportive peers. Learning is a social process that requires relationships with others, access to help, and models of good work [19]. In earlier work, we examined the processes by which newcomers to the community learned to be productive Wikipedians [4]. We found that participation in a WikiProject can provide important

opportunities for new editors to learn how to contribute to Wikipedia successfully.

I am just using these other articles to kind of get used to how to write articles and style guides and that kind of thing and I plan to, over the next few months, go over to WikiProject Systems, kind of full time... I used [Military History], having a learning place for me to learn what good article writing is all about. (I18)

[In other WikiProjects] I can find certain people that are very passionate about some things, but there wasn't the same level of caring that there was [in Military History]... for me to go to computer networking and contribute, I wasn't going to be learning anything from that. (I16)

These qualitative findings suggest that active projects afford their members greater opportunities for meaningful interaction with other editors. To provide further evidence for this claim, we also examine members' interaction history upon joining a project. Figure 2 shows 30,189 editors' interaction activity before and after joining a project. Interaction activity is defined as total edits made to other editors' user talk pages or by other editors to their own user talk page; these data represent 1,386,100 interactions in total.

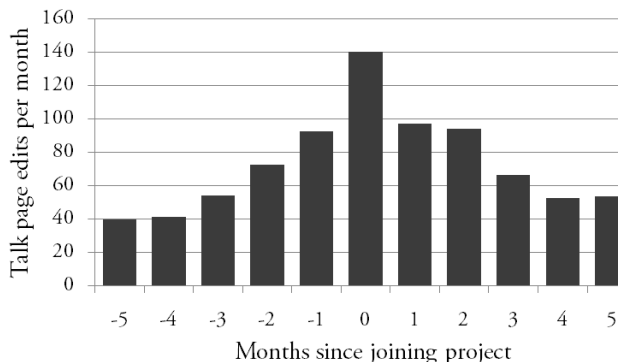


Figure 2: Average member interaction activity per month before and after joining a project

As can be seen from the figure, editors' interaction activity increases upon joining a project. After an initial peak of interaction activity when editors join a project, social activity ebbs off again; however, five months after joining, average interaction activity remains higher than before joining a project. The gradual rise before the peak and apparent drop of interaction activity following the peak can be attributed in part to the fact that in any given time period, some users join and others drop out of Wikipedia altogether. So although all editors are active at the time when they join a project (at the peak), some may have begun editing immediately before joining and others may have dropped out soon afterwards. Since interaction activity is still elevated 5 months after joining, it appears

that joining a WikiProject has an effect on interaction activity, and could indicate an impact on overall editor retention. However, an alternative interpretation is that heightened commitment and interaction with other project members partially cause editors to join a project.

5.3 Structuring Participation

Contributing to a massive enterprise like Wikipedia can be a daunting experience. Where can one best contribute? What needs to be done? In an analysis of edit histories, Kittur and Kraut found that, upon joining a WikiProject, members focused their editing activity on pages managed by that project [13]. In interviews, WikiProjects members described *how* projects influence their editing activities. In some cases, WikiProjects help structure members' participation explicitly by organizing to-do lists and task forces:

Usually what [task forces] ask is for some people to commit to doing something, they get a goal, some task to do. (I19)

WikiProjects also structure participation in implicit ways, as a byproduct of content standardization. When there are models and standards to which editors aspire, bringing an article up to par involves addressing concrete deficiencies, not just "improving it" in an abstract sense. Systematically structuring content allows potential collaborators on an article to easily see what's missing and how they might contribute.

If you go to just about any nation and look up something on Wikipedia, you will find a template that helps you to navigate between the different articles that are related. You know, if you go to the US Army page there will be this thing that has links to artillery, a timeline of US Military History. Well, they had that for the United States, the United Kingdom, Russia, but for smaller countries, that didn't exist, so I went in and did one for Argentina. (I19)

I remember getting to the Music project's site. They have the templates and what they all mean. For instance, the little color of the info box signifies if it's like a band or a solo artist or a vocalist and then I was reading what each thing meant and I actually liked that part and I thought, you know maybe I should sign up for it. (I14)

Structuring the work so that many people can effectively contribute is critical to harnessing the power of the crowd for complex and interdependent tasks such as writing articles. For example, Kittur & Kraut [12] demonstrated how a small core of article leaders could implicitly structure an article in such a way that many peripheral contributors could make smaller but still useful contributions while avoiding some of the overhead of explicit coordination and communication. WikiProjects

provide many of the same benefits of implicit coordination, but across a collection of articles rather than a single article. Interviewees’ descriptions of these coordinating functions of WikiProjects suggest that structuring activities can have a motivating effect, whether through implicit mechanisms like content standardization or explicit calls to action. By structuring activities, WikiProjects engage both in group well-being and member support functions. McGrath notes that, in order to maintain their health as a unit, groups need to define networks and roles to establish who will do what, when, and with whom. In WikiProjects, volunteers self-select into these roles; supporting this process is an important aspect of maintaining group well-being. Furthermore, it provides member support for individuals who identify these opportunities for self-selected roles. Member support “includes both the self-selection side and the group assignment side of the practices and policies by which individuals attain positions or roles in the group” [20].

A further implication of WikiProjects structuring of participation is that, for a given activity, a small core of project members may be able to coordinate the activity of many others. To test this, we examined the benefits of increasing numbers of project members on predicting either coordination activity within the project or production activity on articles relevant to the project. We define coordination activity as edits to the project pages themselves, while production activity is defined as edits to articles the project has adopted (by adding a project template to the article). We examine activity from 379 projects over the course of 10,220 project-months, totaling 11 million project-relevant edits. For each type of activity (coordination/production) we run a fractional polynomial regression predicting activity based on the number of members in the project. Since the value of the number of members in a project is likely to be curvilinear, we also include a squared number of members term.

The results are shown in Figures 3 and 4 for coordination and production activity, respectively. These figures show predicted activity based on the number of members in the project. Interestingly, we see very different patterns for coordination and production activity. The convex function in Figure 3 shows that the first few editors drive most of the coordination activity, consistent with the hypothesis that a small core of leaders drives the structuring of project work. In contrast, Figure 4 shows a concave function for production work, with more members leading to even greater gains in article editing. Together, these results suggest that a small group of project members can provide much of the value in coordination activity, structuring the work so that very many editors can be involved in useful.

By structuring activities, coordination efforts make it clear how individuals can start contributing. Making it clear how to become part of the group is critical for a volunteer-based organization. It is important to note that coordination efforts do not necessarily amount to task assignment. In the FLOSS community, it has been found that self-assignment is the most common task assignment mechanism [6]. Making the choice to participate means that volunteers exchange their time and effort for the rewards of affiliation. In the next section, we examine one of the perceived payoffs of participating in a WikiProject.

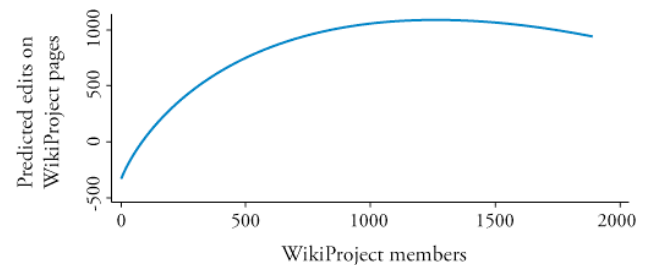


Figure 3: Predicted coordination-related activity (edits to project pages) versus number of project members.

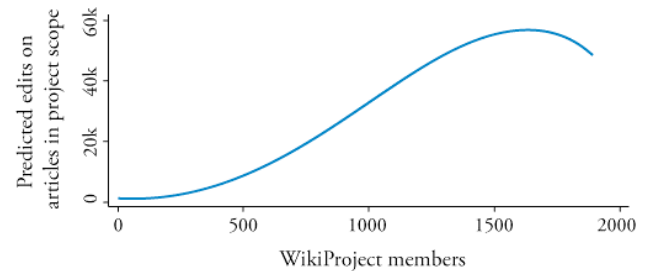


Figure 4: Predicted production-related activity (edits to project-related articles) versus number of project members.

5.4 Protecting and Valuing Editors’ Work

Interviewees suggested that editing as a member of a group can be advantageous because, in the wilds of Wikipedia, it’s nice to know that you have a supportive team protecting your work. Good work can get lost in the whirl of Wikipedia activity for many reasons, ranging from zealotry and vandalism to editors simply not knowing how to explain their efforts if challenged. When editing as part of a WikiProject, Wikipedians feel that their contributions have more staying power.

If you are within the project, you are sort of a little safer. Like if you are just doing articles on your own, you are sort of like freelance, but if you are working with the Military History project than metaphorically it is like working in Chicago Tribune or something like that. Where you have the backing of all these people. (I14)

I will enlist of the aid of other Wikipedia users to help me prevent some pernicious vandal or somebody... from screwing things up. (I12)

Aside from having allies in case of trouble, WikiProject members have a sense that their work will continue to be improved and is likely to lead to high quality content.

I would start an article stub and then find someone to say “Hey, can you expand this?”... in many cases there are professional scholars that get involved and make it very scholarly, they’re well written, well researched... I have seen them transform a lot of articles that way... some people will come in and clean it up and make it look neater and that’s exactly what the project’s designed to do. (I19)

The perception that one’s work will be improved and maintained helps support members who invest their time and effort in the project. As in Wikipedia more broadly, WikiProject members also receive explicit rewards for their contributions to the group: “they have contests with the bronze star award and those kinds of things” (I18). Payoff is an important aspect of member support.

DISCUSSION & OPPORTUNITIES FOR FUTURE WORK

Developing an open content reference work requires retaining knowledgeable volunteers who are willing to learn and continuously improve a body of editorial guidelines and policies that ensure quality. We have demonstrated how local organizational structures help the Wikipedia community meet these goals by engaging in both production- and non-production-oriented member and group support functions.

McGrath’s approach to understanding groups is a reaction to controlled, laboratory studies in which groups are asked to perform tasks by experimenters. In such contexts, anything that distracts groups from the completion of the given task implies a loss of efficiency and performance. McGrath’s typology frees researchers from narrow assumptions about groups’ and group members’ goals and needs. Groups and their members do more than simply get jobs done; they set and revise goals, resolve disputes, bond, support one another, and may have multiple concurrent projects. In a volunteer-based project like Military History, with fluid membership that regularly waxes, wanes, subdivides, and regroups, all of the modes and functions seem to be experienced in an ongoing fashion by different members.

Although many interviewees described experiences with WikiProjects that supported their engagement with Wikipedia in different ways, editors’ experiences vary. For example, some editors suggested that WikiProjects are important for the community:

WikiProjects are good for Wikipedia because it helps keep editors motivated to write and edit. (I13) and even indispensable:

I probably would not have gotten involved [in Wikipedia] without the WikiProjects because they

kind of set the standards for writing and they provide a forum, I guess, for experts to gather and put their work together. So yeah. The WikiProjects are critical to the whole Wikipedia thing. (I18)

One interviewee expressed indifference toward most of the support that projects provide, preferring to “freelance” and avoid the structure of the WikiProject, although he noted in his interview that he respects the advice of Military History editors:

My name is not registered in the Project even though I am probably the most active person in that space... I have my own to-do list that is very very extensive and I don’t need additional ideas where work is required (laughter) (I12).

Our analysis suggests that, for people who choose to get involved, WikiProjects provide many forms of member support that can help make editing Wikipedia an enjoyable experience. Our data suggest that Wikipedia editors who get involved in an active WikiProject experience some benefits that may affect their persistence as an editor. For example, interviewees describe benefits of project membership such as learning how to write well and conform to the standards of Wikipedia; however, we also found evidence that a “critical mass” of activity may be necessary to sustain the project and confer these benefits. The requisite kind and quantity of activity is an open question. We discovered that active WikiProject members view opportunities to interact and socialize with editors who have similar interests as a benefit of editing in a Project and that they maintain slightly higher levels of social activity after joining a Project than before. Finally, WikiProject members observe that editing within a project gives them a strong sense that their work will persist and be improved.

Whether these perceived benefits of WikiProject affiliation have a lasting effect on the persistence of Wikipedia editors and the quality of their contributions poses an interesting question for further study. This work provides preliminary evidence that specialized workgroups play a critical role in peer production environments that go beyond coordination; they confer benefits in the form of member support that may encourage and sustain volunteer participation in the project as a whole.

ACKNOWLEDGEMENTS

NSF grants OCI-09-43148, IIS-0968484, IIS-1111124 and a grant from Carnegie Mellon’s Center for the Future of the Work.

REFERENCES

1. Beschastnikh, I., Kriplean, T., and McDonald, D.W. Wikipedian Self-governance in Action: motivating the policy lens. in *Proc. International Conference on Weblogs and Social Media*. AAAI Press, 2008.

2. Blumer, H., What is wrong with social theory. *American Sociological Review*, 19, 1, (1954), 3-10.
3. Bruckman, A., Studying the Amateur Artist: a perspective on disguising data collected in human subjects research on the internet. *Ethics and Information Technology*, 4, 3, (2002), 217-231.
4. Bryant, S., Forte, A., and Bruckman, A. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. in *Proc. Group: 2005*, 1-10.
5. Cosley, D., Frankowski, D., et al. SuggestBot: using intelligent task routing to help people find work in wikipedia. in *12th International Conference on Intelligent user interfaces*. ACM, 2007, 32-41.
6. Crowston, K., Wei K., L.Q., et al., Self-organization of teams in free/libre open source software development. *Information and Software Technology Journal: Special issue on Understanding the Social Side of Software Engineering*, 49, 6, (2007), 564-575.
7. Dorogovtsev, S.N. and Mendes, J.F.F., Evolution of networks. *Advances in Physics*, 51, 4, (2002), 1079–1187.
8. Forte, A., Larco, V., and Bruckman, A., Decentralization in Wikipedia Governance. *Journal of Management Information Systems*, 26, 1, (2009), 49-72.
9. Grudin, J., McGrath and the Behavior of Groups (BOGS), in *HCI Remixed: Reflections on works that have influenced the HCI community*, T. Erickson and D.W. McDonald, Eds., MIT Press, Cambridge, 2008, 105-110.
10. Kim, A.J., *Community Building on the Web : Secret Strategies for Successful Online Communities* Peachpit Press, 2000.
11. Kittur, A., Suh, B., et al. He says, she says: Conflict and coordination in Wikipedia. in *Proc. CHI*, 2007, 453-462.
12. Kittur, A. and Kraut, R. Harnessing the wisdom of crowds in Wikipedia: quality through coordination. in *Proc. CSCW*, 2008, 37-46.
13. Kittur, A., Pendleton, B., and Kraut, R. Herding the cats: the influence of groups in coordinating peer production. in *Proc. WikiSym*. 2009.
14. Kollock, P. and Smith, M., The economies of online cooperation: gifts and public goods in cyberspace, in *Communities in Cyberspace*, M. Smith and P. Kollock, Eds., Routledge, New York, 1999, 220-239.
15. Krieger, M., Stark, E., and Klemmer, S. Coordinating Tasks on the Commons: Designing for Personal Goals, Expertise and Serendipity. in *Proc. CHI*, 2009, 1485-1494.
16. Kriplean, T., Beschastnikh, I., et al. Community, Consensus, Coercion, Control: CS*W or how policy mediates mass participation. in *Proc. Group*, 2007.
17. Kriplean, T., Beschastnikh, I., and McDonald, D.W. Articulations of wikiwork: uncovering valued work in Wikipedia through barnstars. in *Proc. Computer Supported Cooperative Work (CSCW)*, 2008, 47-56.
18. Lampe, C., Wash, R., et al. Motivations to participate in online communities. *Proc. CHI*. 2010, 1927-1936.
19. Lave, J. and Wenger, E., *Situated Learning: legitimate peripheral participation* Cambridge University Press, Cambridge, 1991.
20. McGrath, J., Time, Interaction, and Performance (TIP): A Theory of Groups. *Small Group Research*, 22, 2, (1991), 147-174.
21. Mossa, S., Barthélémy, M., et al., Truncation of power law behavior in “scale-free” network models due to information filtering. *Physical Review Letters*, 88, 13, (2002).
22. Newman, M.E.J., Power laws, Pareto distributions and Zipf’s law. *Contemporary Physics*, 46, 5, (2005), 323–351.
23. Nov, O., What motivates Wikipedians? *Commun. ACM*, 50, 11, (2007), 60-64.
24. Ostrom, E., Collective action and the evolution of the social norms. *Journal of Economic Perspectives*, 14, 3, (2000), 137-158.
25. Panciera, K., Masli, M., and Terveen, L. 'How should I go from __ to __ without getting killed?' Motivation and Benefits in Open Collaboration. *International Symposium on Wikis and Open Collaboration*, 2011.
26. Raymond, E., *The Cathedral and the Bazaar: Musings on Linux and Open Source by an Accidental Revolutionary* O'Reilly & Associates, Sebastopol, CA, 2001.
27. Resnick, P., Konstan, J., et al., Starting New Online Communities, In *Building successful online communities: Evidence-based social design*. R.E. Kraut and P. Resnick, Eds., MIT Press, Cambridge, in press.
28. Schutz, A., *The Phenomenology of the Social World* Northwestern University Press, 1967.
29. Stewart, K.J. and Gosain, S., The Impact of Ideology on Effectiveness in Open Source Software Development Teams. *MIS Quarterly*, 30, 2, (2006), 291-314.
30. Strauss, A. and Corbin, J., *Basics of Qualitative Research: techniques and procedures for developing grounded theory* Sage, London, 1998.
31. Viegas, F., Wattenberg, M., et al. Talk before you type: coordination in Wikipedia. in *Proc. Hawai'ian International Conference on System Sciences*, 2007.
32. Weber, M., *Basic Concepts in Sociology* Citadel Press, 1966.
33. Wikipedia. *WikiProject Military History Members*. http://en.wikipedia.org/wiki/Wikipedia:WikiProject_Military_history/Members.
34. Zachte, E. *WikiStats: Wikimedia Statistics*. [cited 06/03/2011]; <http://stats.wikimedia.org/>.
35. Zehavi, I., Weinberg, D.H., et al., On departures from a power law in the galaxy correlation function. *The Astrophysical Journal*, 608, 1, Part 1, (2004), 16 - 24.