

Project talk:

Coordination work and group membership in WikiProjects

Jonathan T. Morgan^{*}, Michael Gilbert^{*}, David W. McDonald^{}, Mark Zachry^{*}**

^{*}Human Centered Design & Engineering

^{**}The Information School

University of Washington

Seattle, WA USA

{jmo25, mdg, dwmc, zachry} @uw.edu

ABSTRACT

WikiProjects have contributed to Wikipedia's success in important ways, yet the range of work that WikiProjects perform and the way they coordinate that work remains largely unexplored. In this study, we perform a content analysis of 788 work-related discussions from the talk pages of 138 WikiProjects in order to understand the role WikiProjects play in collaborative work on Wikipedia. We find that the editors use WikiProjects to coordinate a wide variety of work activities beyond content production and that non-members play an active role in that work. Our research suggests that WikiProject collaboration is less structured and more open than that of many virtual teams and that WikiProjects may function more like FLOSS projects than traditional groups.

Author Keywords

Wikipedia; group work; group identity; open collaboration

ACM Classification Keywords

H.5.3. Information Interfaces and Presentation (e.g., HCI): Group and organization interfaces.

General Terms

Human Factors, Theory, Design

INTRODUCTION

New Wikipedia editor Endjinn109 has had a lifelong love of trains. He began editing because he noticed serious gaps in the encyclopedia's train-related coverage. As he begins to rectify this situation, he notices that many of the articles that he is interested in improving are tagged with links to other Wikipedia pages with funny titles like 'WikiProject Trains', and 'WikiProject Transport in Scotland'. From visiting these project pages, he figures out that these so-called WikiProjects have something to do with creating and improving train articles, and that there are many such projects on Wikipedia covering topics from Physics to Feminism. He also notices that each project has a list of members. He wonders whether he should join one of them—and if so which one? What does membership mean? Can only members edit certain articles? Can he request editing help from projects even if he doesn't join up? Will he be obliged to help others? Endjinn109 visits the talk page of WikiProject UK Railways and starts reading through some of the recent discussion threads, trying to figure out what this WikiProject actually does.

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.

OpenSym'13, Aug 05-07 2013, Hong Kong, China.

Copyright 2010 ACM 1-58113-000-0/00/0010 ...\$15.00.

Groups emerge in online collaborations as individuals organize their productive activities around shared goals, interests, tasks and work-spaces. These groups can provide important benefits for their members and perform valuable work for the community they belong to. Lave & Wenger [16] assert that the most effective way to understand working groups like these is to examine the work activities their members engage in. But, as the scenario above illustrates, identifying the members of an online group and the work the group performs can be difficult for an outsider—whether they are a new user, a researcher or a system designer.

Research on the behavior of Wikipedia editors has informed our understanding of group work in open collaboration systems. Despite Wikipedia's reputation as an encyclopedia anyone can edit, participation patterns in Wikipedia are similar to those in other peer production systems in that the majority of the content on Wikipedia is written by a relatively small number of highly active contributors [29]. And although encyclopedia articles are the most visible product of Wikipedia editors' work, they spend a large proportion of their time on work activities that aren't directly related to expanding existing articles or writing new ones [13]. This meta-work is important for supporting and regulating the editor community and maintaining the quality of the encyclopedia [28]. Kriplean et al. [14] identified many valuable meta-work activities by analyzing Barnstars—badges that Wikipedia editors award to one another to acknowledge exceptionally valuable contributions. They found that the majority of Barnstars were awarded for work that was not directly related to writing articles but rather for community maintenance, administrative, and quality assurance activities such as providing mentorship, helping to resolve disputes, reverting vandalism, and welcoming newcomers.

Most of the work on Wikipedia, including meta-work, is coordinated in spaces outside of the articles themselves—on article talk pages, user talk pages, policy pages, and community noticeboards. One common class of group workspace is the WikiProject. WikiProjects are collections of pages that enable persistent group collaboration around particular subject matter domains (such as articles about women scientists) or editing tasks (from categorizing 'stub' articles to promoting editor retention) (**Figure 1**). There are over 2,000 WikiProjects on the English edition of Wikipedia. By 2007, over 20,000 Wikipedia editors had participated in at least one WikiProject. Hundreds of these projects are still active in 2013 [33], and new ones are created every month. WikiProjects also exist within editing communities of many of the hundreds of other Wikimedia wikis, such as the French Wikipedia [25].

Previous research has shown that WikiProjects engage editors in productive editing work [12], and that projects can also provide

Figure 1: main project pages for a large WikiProject (left) and a smaller WikiProject (right)

social support and coordinate meta-work activities. In our study we attempt to provide a more detailed description of the way editors coordinate these work activities through WikiProjects. We build on previous research through a systematic examination of work-related discussions on the talk pages of 138 WikiProjects that differ greatly in their size and scope. We analyze a sample of 788 talk page posts over a one-year period in 2011 and 2012 in order to understand how Wikipedia editors use group workspaces to propose, prioritize, and perform work. We present a typology of the article editing and meta-work activities that projects coordinate, and describe the important role that non-members play in WikiProjects. Our findings suggest that many WikiProjects are less formally organized than the projects examined by previous research studies, and that WikiProject work is in some ways less collaborative than previously thought. We discuss the ramifications of our findings for the nature of WikiProjects as groups and we discuss similarities between the group structure of WikiProjects and open source software projects. We close with set of questions for future research.

RELATED WORK

Voluntary online collaboration can be a powerful method for creating common goods with lasting value. However, creating and sustaining a successful online collaboration can be challenging. Platform designers and community members must provide mechanisms for helping potential contributors find productive and engaging ways to get involved, ensure productivity and project maintenance despite lower levels of member commitment and higher levels of member turnover relative to compensated and co-located teams. Many of these challenges may be effectively ameliorated when participants work together. Groups—large or small, formal or informal, pre-defined or emergent—help match volunteer contributors’ interests and expertise with tasks that need to be accomplished [1]. Well-organized groups can also reduce coordination costs by allowing a relatively small number of core participants to organize the activities of a much larger halo of transient, low-volume contributors [21]. And group collaboration can help motivate volunteers to keep contributing by making the work experience itself more socially engaging and pleasant [18].

Wikipedia is a prime example of a successful online collaboration where groups have proven effective in addressing many of these challenges. Groups on Wikipedia help resolve disputes [13], improve the quality of articles [28] and provide scalable, decentralized

mechanisms for community governance [7]. WikiProjects in particular have been shown to shape editor participation in several ways that are beneficial to Wikipedia: editors who join WikiProjects edit more, communicate more with other editors, and are more likely to engage in ‘good citizenship’ activities such as reverting vandalism [12]. Findings from previous research suggest that WikiProjects can also provide a variety of social support functions for their members [6]. Many previous studies have examined the role of WikiProjects in article production work—edits by project members to articles within the project’s area of focus (see for example [12][27][2][25]). Other studies have examined coordination work and social support within large WikiProjects [6][32]. We complement these studies by analyzing how work is coordinated among both self-identified group members and non-members in the project workspaces of a large and diverse set of active WikiProjects.

Joining a project: Declared vs. participatory members

Unlike many groups on Wikipedia, such as the ad hoc groups that emerge in article talk page discussions [17], WikiProjects allow editors to explicitly declare themselves as project members by adding their name to a canonical member list located in the project space (See **Figure 2**). Member lists are a ubiquitous feature of WikiProjects, and many projects list dozens or hundreds of editors as members.

Visible membership can be an important factor in the development of group dynamics: knowledge of who is a member of a group helps shape members’ attitudes and behaviors towards each other, promotes interaction and defines group boundaries. Group members may exhibit in-group favoritism [30], establish group norms and common repertoires [16], share a strong sense of group identity [23], and develop common bonds as they work together on joint tasks towards common goals. The presence of an identifiable ‘out group’ can intensify these tendencies, influencing the behavior and interactions of both members and non-members.

Like Wikipedia as a whole, WikiProjects present few explicit barriers to participation. Declaring oneself a project member by adding one’s name to the project member list is not required in order to contribute to articles within the project’s scope, edit project pages, or participate in project discussions. Previous studies of group dynamics in WikiProjects have operationalized project membership in several ways. Wang et al. [27] and Chen et al. [2] limited their analysis of WikiProject participation to the editing

Clear video game backlog? [edit]

There is a discussion going on at the Video games WikiProject about clearing out the backlog for video game nominations. As you know, there are some month-old GANs that are in need of reviewing them, but there are also second opinion requests for these as well. If anyone is willing to step up and clear out the backlog, please do so. Thanks, Lord_Soren[3] (talk · contributions) 06:19, 14 February 2013 (UTC)

The backlog there is due to one editor who has built a few bridges when it comes to reviewers. The editor in question, Niemi, is currently the subject of an ongoing RFC for disruptive behavior related to video game articles and it seems unlikely he'll be responsive to the RFC. At this point, what is the best possible solution? Lord_Soren[3] (talk · contributions) 06:07, 14 February 2013 (UTC)

I have no idea. A limit on how many nominations an editor could have at once would help with the backlog in these situations at least, but that has failed to get consensus the last few times it has been proposed. It also doesn't solve the underlying issue. I guess we could propose a motion to ban them from submitting articles here, but I don't know whether that would gain consensus either. There is no real ban worthy issues with the nomination that I can think of. It just seems to be a case of an editor that doesn't play well with others.

Part of the RFC is about not banning GANs as detailed here. Niemi (as Hansmattler) was previously banned because of chronic inactivity and inability to cooperate with others back in 2008, was unbanned in 2012 and he has recently avoided that same fate this past November. We should go ask the VD project about this matter. Lord_Soren[3] (talk · contributions) 06:48, 14 February 2013 (UTC)

I would say that any decisions regarding this Good articles should be discussed/made here (or even better at WT:GAN as it has a higher traffic flow). Looking at the VG thread most of the respondents seem to have wiped their hands of this. AFAIK[4] (talk) 07:03, 14 February 2013 (UTC)

I have started a centralized discussion at the GAN talk page and I also feel that this backlog cleanup needs to happen sooner rather than later. The discussion can be found at Wikipedia talk:Good article/nominationalCleaning out the VG backlog. Please direct your comments there. Lord_Soren[3] (talk · contributions) 17:28, 14 February 2013 (UTC)

Really need a qualified 2nd Opinion on Talk:Thomas Traherne/GA1 [edit]

I nominated Thomas Traherne for GA in December, it finally got a reviewer (User:Mohsen) but from his comments he seems to not have read large swaths of the article he's reviewing, admitted he's not qualified to finish the review right from the start, and exhibits being rather oblivious to policies/guidelines. His rigor seems to be demanding the acumen of a FAC, asking for expansion of content that goes beyond the wits limits of WP:SUMMARY and WP:DETAIL, and frankly inappropriate for a simple GA nomination. It's like a second opinion from a qualified reviewer, because his comments are becoming increasingly frustrating. --ColonelHenry (talk) 17:15, 14 March 2013 (UTC)

Blocked reviewer [edit]

I posted this on the help desk, but have been advised to take it here. I'm in the process of having Bradley Wiggins reviewed, but the reviewer (QAnchisus) has been given an indefinite block. Of course, I've been blocked until its lifted, if ever. What do my readers think? Does it need to be closed and uncompromised? Basilio[5] 17:47, 14 March 2013 (UTC)

Honestly, I don't know whether you should close & remove him or not. He looks like the GAF failed. It seems that QAnchisus is the sock of a anonymous user, not just simply blocked. So, why don't you go and remove QAnchisus as the reviewer? What if that is a new reviewer? Someone from this project might patch in, or you could post notices on the talk pages of the various WikiProjects which pertain to your article. It looks like you've already addressed the editing concerns, so the review should be quick. (Sorry, I'm not in a WP:DY mood right now.) — S. Rich (talk) 18:27, 14 March 2013 (UTC)

GAN discussion [edit]

After a Wikipedia:Reference source failed, J. Valentine due to licensing concerns as explained at IRON, a discussion about GANs has started over at WT:GA/N#Problem: reviewers often not

If you would like to join the project, please add your name to the list in alphabetical order.

Active Members

[edit]

- Alexis Paladin
- Alison
- Anirishwoman
- Arverniking
- Aughrafin
- Autarch
- Backbeatlistener
- BalticPat22
- Bardcom
- Barry O'Brien

Figure 2: talk page for WikiProject Good Articles (left). Member list for WikiProject Ireland (right)

activities of editors whose names appeared in the member rolls. Ung & Dalle [32] and Zhu et al. [25] counted editors as project members if they had previously made at least one edit to any project page. However, the degree to which a project’s list of declared members reflects the actual number of project participants at a given point in time has not been determined. Kittur et al. [12] assert that non-members rarely edit project pages. However at least one highly active WikiProject Military History participant interviewed by Forte et al. [6] was not a declared member of that project.

The possibility of active participation by both declared members and non-members in WikiProjects suggests an opportunity to examine group dynamics related to interpersonal interaction, group identification and the formation of common bonds in open collaborations. The setting for these interactions is the group workspace itself, where the distinction between declared members and participatory members is most salient and any in-group/out-group behaviors are most likely to be more pronounced. Chen et al. note [2] that many editor characteristics, such as their interests or their edit count, are not readily apparent to the other editors they interact with on a day-to-day basis. In most contexts on Wikipedia, an editor’s membership status with a particular WikiProject is similarly difficult to ascertain unless the editor chooses to advertise it explicitly on their own user page. The visibility and ubiquity of project member list suggests that behaviors related to group identification such as in-group favoritism or behavioral similarity may be more evident in interactions on project pages. We leverage this observation in our investigation, which contributes to the existing body of research on group dynamics in WikiProjects and other open online teams in two ways: we analyze the degree to which membership status reflects real differences in how editors use project talk pages to coordinate work, and we analyze whether declared members display behaviors related to common bonds and common identity in their interactions with other members and non-members.

WikiProject work: Content production vs. coordination

Many WikiProjects have a topic focus, and several previous studies have operationalized the production work of WikiProjects by measuring edits by a project’s members to articles within that project’s scope [27][2][6]. Several studies have shown a strong

relationship between project membership and edits to project-related articles, and between the level of activity on project pages and article-editing activity [12][25]. However, the work of WikiProjects that are organized around editing tasks that span topical boundaries, or activities that are not directly related to writing and editing articles, cannot be measured in this way. The work these projects perform, and the way they coordinate that work has not yet been examined. Furthermore, as Wang et al. [27] note, using member-edits to project-related articles as the sole criterion for project participation may limit the explanatory power of the analysis even for topic-focused projects because it does not capture other ways editors contribute to WikiProjects.

Several previous studies have presented evidence that coordinating the work of editing articles in topic-focused constitutes an important form of meta-work. Zhu et al. [32] found that Collaborations of the Week (COTWs), a type of structured editing event organized on the project pages of some WikiProjects, were effective at getting project participants to edit articles. And an in depth case study by Forte et al. [6] based on interviews with 15 members of WikiProject Military History identified a variety of social functions that project performs for its members, many of which were explicitly coordinated within the group workspace: providing opportunities for social interaction and networking, and forums for editors to find help and collaborators. A separate research study based on these interviews [7] described some of the sophisticated coordination tools that Military History had developed to improve and evaluate the quality of articles within the projects scope: project newsletters, topic-specific formatting guidelines, specialized task forces and article assessment noticeboards.

A project that can host weekly editing events, develop and maintain guidelines and publish regular newsletters reflects an exceptionally high level of investment for a voluntary collaboration. Many smaller WikiProjects may lack the critical mass of involved participants necessary to sustain structured collaboration mechanisms like these. The authors of [32] acknowledge that COTWs were only used consistently in 13 of the largest WikiProjects, and suggested that other projects may have stopped running COTWs or decided not to adopt them because of their high coordination cost. WikiProject Military History is one of the largest, oldest and

most organized projects on Wikipedia, with 1170 active members and 1000 monthly edits in 2007.

In order to examine how work is coordinated in both large and small projects, we focus our analysis on project talk pages, a feature that all WikiProjects share. Talk pages in other parts of Wikipedia have been shown to play an important role in collaboration and conflict resolution. Kittur et al. [11] demonstrated a link between a high level of talk page activity and article quality in the formative stages of article creation. Viegas et al. [26] and Schneider et al. [24] found article talk page discussions were primarily focused on coordinating editing activities around individual articles. WikiProject talk pages have not been systematically examined in this way, but findings from multiple studies suggest that editors use talk pages to coordinate a broader range of activities across many different articles: Lam et al. [4] found evidence that project members used talk pages to welcome new members, suggest tasks for them to perform and provide constructive criticism; and interview subjects in Forte et al. [6] listed project talk pages as places where members go for help with articles.

Following many previous researchers on Wikipedia (see for example [26][24][14]), we use content analysis to characterize a class of communication activity within a particular interaction context. We categorize the coordination work implicated in messages that editors post to talk pages. This approach allows us to both *qualify* a greater range of work activities and systematically *quantify* the prevalence of those activities across a diverse set of active projects and between declared members and non-member participants.

Hypotheses: Membership status and coordination work

The definition of WikiProjects provided by Wikipedia, “*a group of editors that want to work together as a team to improve Wikipedia*”¹ suggests that WikiProjects are closely-knit groups of close collaborators, and findings from previous studies generally support this characterization of WikiProjects. However substantial gaps in our knowledge about the role of non-members in WikiProjects and the way smaller projects coordinate work introduces a degree of uncertainty as to whether WikiProjects exhibit a degree of collaborative activity and group dynamics similar to teams in other settings—from bridge clubs and office workgroups to football teams and virtual teams. Previous studies of WikiProjects have drawn on theories of group identification to explain the production work and social support mechanisms of WikiProjects. For our own investigation we draw on common bond theory and group identity theory to frame several hypotheses related to how work is coordinated among declared members and non-member participants in WikiProject talk page discussions.

H1: Members and non-members will post different kinds of messages on project talk pages

People who identify with a group are more likely to be influenced by the group's norms and values, and make decisions favorable to the group [3]. WikiProject members are more likely to engage in content production work relevant to the project than other Wikipedia editors [12]. We posit that the work activities that members coordinate on project talk pages will show similar regularities. To investigate this, we analyze the whether messages posted on talk pages by members reflect different kinds of content production and meta-work activities than messages posted by non-members.

H2: Members will respond more to other members than to non-members.

Group members often manifest a sense of common identity and purpose by helping to plan group activities and participating in group discussions with other members [22]. WikiProject members are likely to be aware of the member status of the editors they interact with on project talk pages. We investigate whether common identity influences member participation by analyzing the rate at which project members respond to members and non-members who ask questions or propose new collaborative activities on talk pages.

H3: Members will perform work activities requested by other members more often than activities requested by non-members.

Group members may be motivated to provide individual support to other members because they develop common bonds through a shared history of interaction within the group context, increasing their likelihood of assisting each other and their willingness to work together [22]. To investigate the influence of common bonds on project work, we compare the number of members who perform work requested by other members and non-members on project talk pages.

METHODS

We gathered our data from toolserver.org², a public data repository hosted by the Wikimedia Foundation that maintains a nearly live slave database of many Wikipedia editions. We conducted queries against the Toolserver database parsing results to create our own metadata about WikiProjects and cached results to offload subsequent processing.

The first step in our analysis is to identify a large sample of active WikiProjects. Because WikiProjects may have functioned differently at different points in Wikipedia's history, we sampled WikiProjects that were most active during the one-year period: July 2011 to July 2012. To ensure that all the projects in our sample were active, all WikiProjects in our sample averaged at least 1 edit per day to its project pages (the main WikiProject page, the talk page and any project subpages), excluding edits by automated bot user accounts. This sampling resulted in a set of 138 WikiProjects with a median of 2.6 edits per project per day over the course of the year. For each of these projects, we gathered a random sample of 20 talk page posts that started a new discussion thread. 32 of the 138 projects (23%) sampled had fewer than 20 total thread-starting posts during the sample period. For these projects, we gathered all thread-starting posts. This sampling yielded 2,465 thread-starting messages.

To determine whether someone was a declared member of a WikiProject at the time they posted to the talk page, we first identified the member lists for each WikiProject by looking for sections of the main page and/or project subpages with titles like ‘Member’ or ‘Participant’. We then parsed the text of every revision to that page or section, capturing the date at which each editor added their username to the list. If an editor's username appeared on the member list before they posted to the talk page, we considered them a member at the time of that post. The median number of declared members in the projects we sampled was 111. WikiProject Military History was the most populous project in our sample, with 1,955 declared members as of July 2011.

¹ <http://enwp.org/w/index.php?title=WP:WikiProject&oldid=545662745>

² <http://toolserver.org/>

Table 1: Categories of work-related requests and reports

Code	Definition	Example	# Messages	% Messages
REQ-OPIINION	Requests advice, opinion or informal feedback, or attempts to gauge local consensus around an idea	<i>I would like articles like Solar eclipse of July 16, 2186 to have their own subcategory in the eclipse WikiProject. Is this possible? Am I duplicating a structure that already exists?</i>	303	38%
FYI	Post is only a statement or announcement - contains no requests or other directives	<i>The popular pages list has been updated for January. The only surprises are Jane Fonda and Yoko Ono shot into the top 10. Other than that, I didn't notice any major changes.</i>	141	18%
REQ-DISCUSSION	Requests or suggests that others join a discussion on another wiki page	<i>Greetings. A discussion related to this WikiProject has been opened at Wikipedia talk:Articles for deletion#Deletion sorting idea. Your thoughts are most welcome.</i>	130	16%
REQ-OTHER-People	Request that other people perform edits to content pages, but does not offer to help out	<i>Hello there. If you look carefully at the info box for The Dungeonmaster, you'll see that it's a little screwed up. I'd like to fix it, but I don't know how. Help would be appreciated.</i>	60	8%
REQ-COORD-ART	Requests or proposes coordinated editing of articles	<i>Is everyone still happy to keep bashing away at episode articles for the time being, or should we maybe see about working towards something major together?</i>	30	4%
REQ-PEER-REVIEW	Requests that someone perform official peer review of encyclopedia content	<i>Please join the discussion on whether List of members of Stortinget 2005–2009 meets the featured list criteria. Articles are typically reviewed for two weeks; editors may declare to "Keep" or "Delist" the article's featured status.</i>	29	4%
REQ-INFO	Requests information that is not related to editing Wikipedia or being an editor	<i>I am trying to find out what happen to KRCR-TV anchor, Tim Mapes. How can I find out.</i>	28	4%
REQ-COORD-NONART	Requests or proposes coordinated editing of content pages that are not articles	<i>I have created this category and categorized it with a few articles. I bring attention it here for all interested and knowledge to add more.</i>	23	3%
INVITATION	Request to join another project, take part in a named initiative, or attend an event	<i>For anyone who is interested I will be giving a talk on my recently-published book "Jewry in Music: Entry to the Profession from the Enlightenment to Richard Wagner" at the Gustave Tuck Lecture Theatre on 22nd February.</i>	21	3%
REQ-MONITOR	Request to keep an eye on a page or user, or perform an administrator action	<i>Keeping an eye on Dream Chaser might be a good idea - a clearly COI/promotional account tried to turn it into a spammy puff piece. The account has now been blocked but they could always try again</i>	14	2%
REQ-OTHER	Post contains a request that does not fit into any of the request categories	<i>Hi everyone, I sadly had to change my status as a project mentor to "busy" because that is the state of my life right now. I would appreciate other members who would like to come forward to help out here.</i>	7	1%
REQ-TASKS	Requests that others suggest tasks for the poster to perform	<i>I am new to Wikipedia and new to this WikiProject, but I would like to help in other ways. Please let me know if there's anything I can do to help improve conservatism-related articles on Wikipedia!</i>	2	0%
Total			788	100%

Codebook development

We based our coding scheme on the codebook developed by Viegas et al. [26] and subsequently refined and expanded by Schneider et al. [24], and supplemented their categories with new codes based on work activity identified in previous studies of WikiProjects. The first author performed an initial open coding of a subsample of our dataset—refining the definitions of existing categories to reflect their presentation in WikiProjects, dropping categories that were not attested in our data, and noting edge cases and unclassifiable posts for discussion with the research team. Our final codebook contained 12 post categories (**Table 1**). Because our focus is on coordination work, we follow Viegas and Schneider in discriminating between talk page posts that contain explicit directive cues such as *requests*, *suggestions* or *proposals* (hereafter, ‘requests’) and posts primarily intended to convey information that do not contain explicit requests for responses or follow-up actions (hereafter, ‘reports’). 11 of our categories describe different types of request; one category, FYI, was assigned to all reports. Messages coded as FYI contained information relevant to editing Wikipedia, but no explicit request. 7 posts that were self-

evidently off-topic or unclassifiable—obvious vandalism, nonsensical remarks or parser errors—were marked as INVALID. These posts were removed from our final dataset and not included in analysis.

Data annotation

We coded 788 posts randomly-selected from our initial dataset according to our coding scheme. Each post was categorized by two independent coders. All coding disagreements were adjudicated through discussion among two or more members of the research team in order to finalize our coded dataset. While the application of message-level categories involves a greater degree of subjective interpretation than would have been necessary for a content analyses of lower-level discursive phenomena (e.g. grammatical structure), we computed inter-coder agreement statistics (Krippendorff’s α) on samples of data at various points during our coding process. This additional validation step proved useful for identifying and refining categories that were difficult to distinguish from one another, helped us train our individual mental classifiers, and functioned as a basic sanity check on the viability

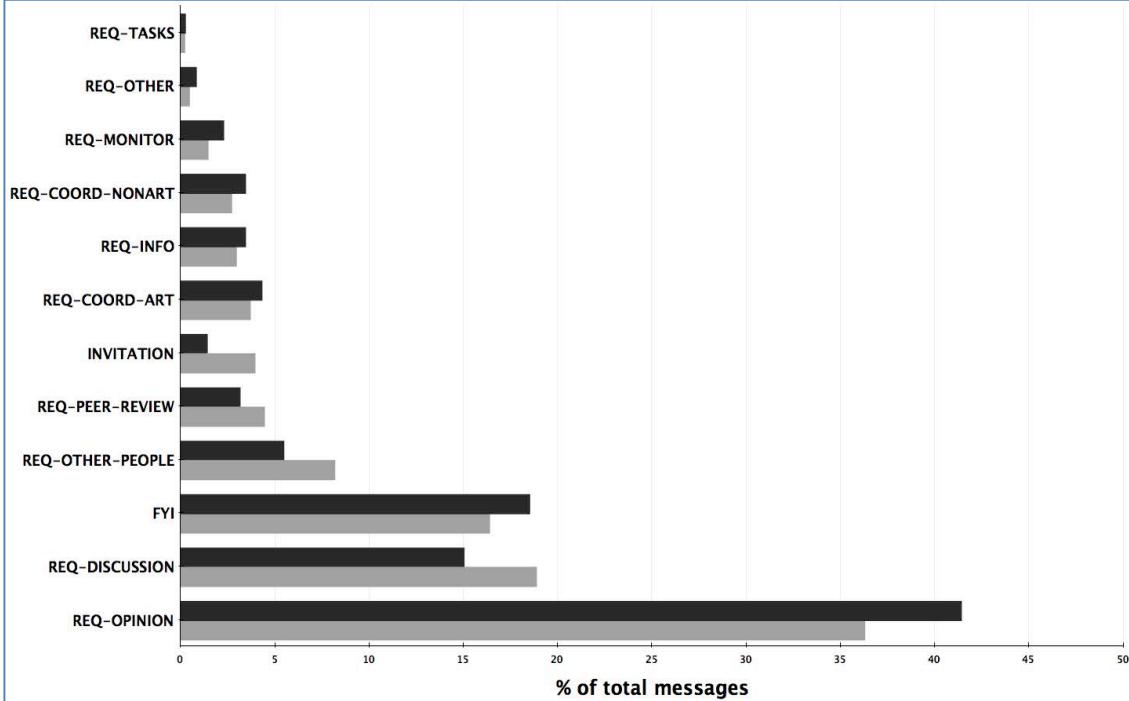


Figure 3: Proportion of message types by project members (black) and non-members (grey)

of our categories. Our average agreement *before* adjudication across all categories was $\alpha = 0.58$. This is well within the range (0.4 – 0.6) considered ‘moderate’ agreement by Landis & Koch [15].

We also gathered a set of 2,047 replies to the posts in our sample, which we used to analyze the subsequent activity within the coded threads thread.

FINDINGS

Our full coded dataset is presented in **Table 1**. Requests of all types comprise 82% of our data, with 18% of posts falling into our report category, FYI. A breakdown posts by category between members and non-members is presented in **Figure 3**.

Types of coordination work

Most of the requests in our dataset were not directly related to collaborative editing of Wikipedia articles. Only 4% of messages contained explicit requests to edit specific articles together (REQ-COORD-ART). A similarly small proportion (3%) contained requests to collaboratively edit content on pages that are not articles such as templates, category pages, and task lists. Non-collaborative content editing requests (REQ-OTHER-People) were more common than both of the collaborative editing categories combined.

This result suggests that the way WikiProjects have been framed, as groups of editors who edit articles together, may be somewhat misleading. While project members may collaborate more directly in other forums on Wikipedia, such as user talk pages, our data suggest that the coordination of collaborative article editing organized at the project level is relatively rare. Instead, most requests reflect a more lightweight approach to collaboration. The largest request categories—REQ-OPTION and REQ-DISCUSSION, which do not contain explicit requests to edit content collabora-

tively—account for 54% of all messages in our dataset. Many REQ-OPTION messages implicate collaborative activities of a short-term and transactional nature on the article talk page itself.

“What are the opinions of those here on whether [The Formula 1 Blog](#) should be regarded as a reliable source for facts related to Formula One. It is a self-proclaimed blog site after all, and [WP:BLOGS](#) is fairly clear about the reliability of such self-published sources. Is there an overriding reason why this one should be exempt?”
[WikiProject Formula One]

The frequency of this type suggests that many editors work autonomously and only turn to the WikiProject when they get stuck, need advice, or just want to bounce some ideas off someone. REQ-DISCUSSION messages also reflect coordination of work that does not involve content production.

“I have just proposed a merge from [The Schoolmaster’s Progress](#) to [Caroline Kirkland](#), and I would be grateful if any of you could come and comment. The merge discussion can be found at [Talk:Caroline Kirkland#Merge proposal](#).” [WikiProject Novels]

In this message, the requester is asking for others to add their voice to an ongoing discussion on another talk page. Anecdotally, we observed that many of these requests involved some sort of dispute resolution, with the requestor intent on recruiting allies to bolster their position. In other cases, requests for discussion seemed to serve a similar purpose to FYI messages: the requestor wanted to make the group aware of activities elsewhere that may be relevant to their interests or expertise, and to which they could contribute if they saw fit.

“A proposal to merge [Scientific law](#) into [Laws of science](#) is being discussed here.” [WikiProject Philosophy]

Simple reports of activity with no associated requests were also common. Our second largest category, FYI, was attested in 18% of all posts. The prevalence of reports on WikiProject talk pages shows that editors frequently use these forums in order to alert other editors of relevant goings on, rather than to elicit a direct response. This finding suggests that these forums function as mechanisms for group awareness in addition to group coordination, similarly to email lists in open source software projects [9].

"I made a new userbox, if anyone is interested in it. This one can be used by people who aren't a part of the WikiProject as well, just to show their support."
[WikiProject Cooperation]

Although we designed our coding scheme to account for concrete types of work-related requests, rather than abstract types of collaboration, we see parallels between our codes and the three broad, roughly hierarchical types of collaborative activity—*coordination*, *cooperation* and *co-construction*—outlined by Kaptelinin & Nardi in [10]. In their typology, *coordination* “refers to cases in which people work towards a common goal, but carry out their activities basically independently.” Many of our FYI messages could fall into this category. *Cooperation* by their definition requires individuals to “relate their goals to the overall objective of a collective activity, be aware of the actions of other participating individuals, and adjust their actions to the actions of other people.” REQ-OPINION messages often fit this definition well. The third, and most direct, type of collaboration described by Kaptelinin & Nardi, *co-construction*, involves not only the collective pursuit of a common object, but the potential for collective redefinition “of the object—and the collective activity—itself.” Both article editing collaboration requests and the deliberative discussions implicated in REQ-DISCUSSION messages reflect this type of intensive, open-ended coordination—involving multiple participants who weigh alternative, experiment and debate with one another in order to achieve a resolution the shape of which is initially undetermined.

We believe that a natural extension of our study might be to attempt to categorize these data according to the *degree* of collaboration implicated in WikiProject discussions. Such an analysis could be instructive for understanding the way lean media, as well as differing community norms and work objectives, influence how people coordinate with one another in open collaborations.

Non-member participation

Non-members participated in WikiProjects at a much higher rate than we expected. 403 message threads in our sample (54%) were initiated by non-members and we captured 642 responses by non-members. Overall, non-members posted 37% of all messages in our dataset. Because we did not anticipate the high degree of non-member participation, we did not frame any hypotheses around the relative likelihood that members and non-members would start new threads vs. responding to them. However, a post-hoc CHI square test shows that the difference in thread-initiation vs. thread-response rates is significant ($\chi^2 = 115.5$, df = 1, p < 0.01). One consequence of the fact that non-members initiated roughly half of all talk page discussions is that declared members frequently participated in discussions started by non-members. This suggests that non-members may exercise a degree of control over what gets discussed on the project talk page, and by extension help set the agenda of the *de facto* group that uses the project space despite their lack of official member status or declared affiliation.

H1: member and non-member posting behavior

WikiProjects ostensibly exist to support collaboration, and messages that contain requests are more likely to reflect a desire to collaborate than those that do not. Since members are more invested in the project, more likely to edit project related articles and more likely to have collaborated with other members in the past, we posited that members would be more likely to post request messages than non-members. To investigate whether members used the group workspace differently from non-members, we also analyzed whether members made different kinds of requests than non-members.

The hypothesis that members would post different kinds of messages was not supported. A Chi Square test for independence showed no significant difference between the proportion of requests and reports made by members and non-members ($\chi^2 = 0.4$, df = 1, p = 0.52). A second Chi Square test performed across all message categories with an expected count > 5 (excluding REQ-TASKS and REQ-OTHER) also showed no significant difference between the types of requests made by members and non-members ($\chi^2 = 11.03$, df = 8, p = 0.19).

Our findings indicate that project members are no more likely to request work from a WikiProject than report it, and that they do not tend to coordinate different types of work than non-members (**Figure 3**). We do note that non-members post certain types of requests—such as invitations to join other projects, requests for other people to perform edits for them and requests to participate in an external discussions—at a slightly higher rate than members do, behaviors that would be consistent with out-group status. However the lack of overall significance shows that even if these patterns are valid, they only reflect minor behavioral differences between in-group and out-group.

H2: member responses to requests

Previous research also suggests that members will reply more to other members than to non-members because they are more likely to recognize a shared history or a common bond. This hypothesis was supported by our data: 59% of posts by members (across all message types) received at least 1 reply from another project member, but only 45% of posts by non-members received a response from a member. To validate the difference in responding behavior, we performed two unpaired t-tests: one to measure the average number of member responses to member and non-member messages, and one measuring the average number of members who responded in the subsequent thread. In both cases, we excluded messages from the initial poster when they were a member. We found that posts by members received a significantly higher number of responses from other members than non-member posts did ($\text{mean}_{\text{mem}} = 2.09$, $\text{sd}_{\text{mem}} = 4.48$ vs. $\text{mean}_{\text{non}} = 1.53$, $\text{sd}_{\text{non}} = 3.08$; $t = 2.07$, df=788, p=0.04). We also found that more members responded to posts by other members than to posts by non-members ($\text{mean}_{\text{mem}} = 1.07$, $\text{sd}_{\text{mem}} = 1.21$ versus $\text{mean}_{\text{non}} = 0.86$, $\text{sd}_{\text{non}} = 1.29$; $t = 2.28$, df = 788, p = 0.02).

H3: member follow-up to editing requests

Studies of the editing and communication behavior of WikiProject members, as well as theories of group identification, suggest that a request to perform work from a fellow group member will be more likely to result in a follow-up action than a similar request from a non-member. To determine whether this finding held true for requests on WikiProject talk pages, we first identified a sub-set of coded messages that were likely to contain requests to perform editing work beyond the message thread: REQ-DISCUSSION, REQ-COORD-ART, REQ-COORD-NONART, REQ-OTHER-

PEOPLE, REQ-PEER-REVIEW, and REQ-OTHER. We wrote a Python script that followed all wikilinks in the text of these requests that pointed to pages in Wikipedia's primary content namespaces (Article, Wikipedia, Template and Category) and their respective talk namespaces. We then counted how many project members (besides the original requester, if they were a member) edited one of those pages within the next 30 days.

The hypothesis that members are more likely to perform work proposed by other members was not supported. The number of requests that resulted in at least one follow-up edit by a project member was roughly equivalent between the two groups (35%_{non}/36%_{mem}). Pages linked from non-member requests actually received slightly more member attention, on average, than pages linked from member requests ($\text{mean}_{\text{mem}}=2.95$, $\text{sd}_{\text{mem}}=9.41$ versus $\text{mean}_{\text{non}}=3.64$, $\text{sd}_{\text{non}}=12.3$), although this difference was not significant ($t=0.5$, $\text{df}=305$, $p=0.6$).

This result appears to be at odds with our finding from H2, which showed that members exhibited a bias towards responding to threads started by other members. Why would project members talk more with other members, but perform more work for non-members? We speculate that rather than reflecting a favorable bias towards other members, the higher response rate found in H2 may simply indicate that members were more likely on average to monitor the project talk page and therefore participate in discussions they initiated, offering other members more opportunities to respond.

To explore this new hypothesis, we performed a post-hoc t-test comparing the average number of times members and non-members posted replies in the threads they themselves started. We found that members did post more replies to their own initial posts on average ($\text{mean}_{\text{mem}}=1.9$, $\text{sd}_{\text{mem}}=8.49$ versus $\text{mean}_{\text{non}}=0.83$, $\text{sd}_{\text{non}}=2.01$; $t=2.56$, $\text{df}=788$, $p=0.01$). This result contextualizes our findings from H2: it suggests that one reason members respond to one another more often is that they are more actively engaged in the conversation. Over time, these member-to-member interactions may contribute to bond-based affinity or a sense of common identity. But our findings from H3 suggests that such affinity does not necessarily extend to all project members or result in collaboration outside the project workspace: it is based primarily on a history of interaction rather than social categorization based on explicit markers of group identity. Importantly, our findings suggest that such bonds may form through the performance of coordination work within the project workspace rather than in the context of collaborative content editing, even though facilitating content production is the ostensible purpose of many WikiProjects.

DISCUSSION

We set out to provide a more nuanced and comprehensive typology of the coordination work WikiProjects perform and the impact of group dynamics such as common identity, common bonds and explicit group membership on the performance of that work. Projects facilitate a wide range of work activities that extends beyond encouraging collaborative content production, but large-scale, highly structured and intensive editing collaborations are relatively infrequent in most WikiProjects. The majority of the collaborative activities implicated in our sample of talk page posts were short-term and lightweight: requests for informal feedback, for participation in a deliberative discussion (in the project workspace and elsewhere on Wikipedia), or for answers to work-related questions. Non-members are active participants in this coordination work, as both requesters and respondents, and they tend to use WikiProjects to coordinate the same kinds of work activities as

declared project members. This suggests that project member lists are not a good proxy for *de facto* project membership. The purpose of these lists, and the reasons project members chose to add their names to a project's member list would be a productive topic for future research.

Our finding from H2 that members respond more to posts by other members suggests that they may exhibit some degree of in-group favoritism. However, post-hoc analysis suggests that this affinity may be due to the development of common bonds through repeated interactions on the talk page, rather than a shared sense of identity as group members. The lack of significant difference between the rate at which members respond to editing requests by members and non-members (H3) suggests that correlations previously observed between project membership and article editing work [12] are not necessarily reflected in coordination work: members do not exhibit a bias towards performing work requested by other members.

WikiProjects as groups

WikiProjects share some features of traditional groups, such as defined membership boundaries. Previous studies have considered the presence of these features as evidence of group identification and studied its impact on member behaviors such as the kind of content production work they chose to perform. Our findings indicate that the relationship between declared membership and one type of meta-work, coordination work, is equivocal. This suggests that in WikiProjects, coordination work and content production are loosely coupled: editors perform majority of their work independently of one another. Loosely coupled work, which is common in computer-mediated groups, may not promote the strong group bonds found in traditional groups [19]. Further evidence that content production and coordination work in WikiProjects are loosely coupled is provided by Ung & Dalle [25]. They examined the relationship between discussion activity on the project talk page (coordination work) and editing activity on project-related articles (content production) among 644 French WikiProjects, and found that in most cases the most active discussion participants were not the most active editors of project-related articles. They inferred two distinct group structures among the WikiProjects they examined: some projects (a minority) function as closely-knit groups of editors who coordinate frequently, edit collaboratively and exhibit more traditional group dynamics; however in the majority of projects a small group of active discussion participants were instrumental in coordinating the work activities of a larger group of peripheral participants, but engaged in relatively less content production themselves. In these groups, the project pages may function as hubs maintained by a small group of (usually) project members who play the role akin to forum moderators or help desk personnel: answering questions, providing feedback and occasionally responding to work requests that come in. This may be the case in English WikiProjects as well. Larger projects such as Military History, which exhibit many group-like features, may use sophisticated coordination tools that allow them to orchestrate large-scale collaborative activity. Smaller projects, lacking the critical mass to maintain such infrastructure, may adapt their work practices to require less intense collaboration.

WikiProjects and FLOSS projects

While the loose coupling of work and the equivocal role of explicit group identification in WikiProjects sets them apart from many traditional groups, this model has many parallels in other peer production systems. One class of working group that often shares these characteristics is Free/Libre Open Source Software (FLOSS)

projects. Like WikiProjects, many FLOSS projects are egalitarian groups in which both declared members and non-members participate and which exhibit porous group boundaries and possess few defined roles. In a study of message threads in open bug tracking systems of three successful FLOSS projects, Crowston et al. [5] observed group structures similar to those found in many WikiProjects: each project featured a large group of peripheral users and small group of core developers. Yamauchi et al. [31] also observed this strong core/periphery dynamic in two other FLOSS projects, with a small group of developers maintaining project resource pages, participating in coordination discussions in support of a larger group of peripheral participants. The project spaces examined both in Crowston and Yamauchi's studies also exhibited structural similarities to WikiProject pages: most had official developer lists (some of which were open, like WikiProject member lists), and each project provided centralized communication channels, documentation, and other relevant production resources. Furthermore, declared members ('developers') and non-members ('users') often behaved in ways that blurred this official distinction. In general, users tended to post more bug reports and ask questions and developers tended to reply more, but users also frequently responded directly to other users' questions, and even submitted software patches in response to developers' requests. Both studies also noted that core developers sometimes shifted their work activities from content production to coordination work, a similar pattern to the one observed by [25] and suggested (but not investigated) by Wang [27] to explain their finding that long-term WikiProject members tended to edit articles at a lower rate.

While previous studies of WikiProjects have noted their similarity to FLOSS projects, we believe the parallels between these types of group are strong enough to warrant additional investigation. Designing tools to support technologically-mediated groups presents many challenges [8] and doing so effectively requires a detailed understanding of the group's structure and work activities. Designers and community managers who wish to support WikiProjects or foster similar self-organized volunteer projects in other open collaboration systems must provide mechanisms for coordination that not only work for large, well-organized, close-knit groups but also for groups that are smaller, more open, and less formal because many active and successful work groups may reflect one or more of those tendencies.

CONCLUSION

Previous research on WikiProjects has examined them primarily as groups, and drawn on group theories developed from empirical studies of offline groups to explain the behavior of project members and the way the projects themselves function. Both quantitative and qualitative findings from several of previous studies clearly indicate that the group lens is an appropriate one for analyzing some projects—particularly larger, more established projects—and some features of many smaller projects. However our findings suggest that the group lens may not be the most productive way of describing the work most WikiProjects perform, the role of declared members and non-members in the performance of that work, the way participants coordinate work, and the degree to which they collaborate.

Highly organized projects like WikiProject Military History may function as close-knit collaborating groups, at least for their most active and senior members. Highly active projects may be able to afford the heavy coordination cost of maintaining assessment departments, organizing weekly collaborations and distributing

monthly newsletters. However our findings, drawn from a sample of discussion threads from the 138 most active projects on English Wikipedia, suggest that many WikiProjects are less formally structured. The 'typical' WikiProject described by our findings contains a few participants (mostly, but not necessarily, declared members) monitoring and maintaining a project communication hub that supports a wide variety of work activities of a large number of peripheral participants (non-members, more often than not). The work requestors bring to the project to coordinate are largely self-assigned and pursued independently, or possibly a small piece of a more intensive collaboration organized through other channels. When members do engage in collaborative work requested through the talk page, their choice is not necessarily guided by considerations of group status or mutual obligation born out of common bonds or a shared group identity.

Future work

One productive area of future work involves characterizing the range of variation between projects. Most studies, including our own, have lumped dozens or hundreds of projects together and examined small numbers of common features and functions across projects. But such 'averaging' may mask critical variations in the way projects are structured (e.g. more group-like, more FLOSS-like, or an entirely different structure), the work they perform for Wikipedia and in what they provide for their members. Forte's case study [7], Ung & Dalle's [25] study and stray findings from other investigations hint at a world of variation hiding behind our averages and assumptions.

Comparative case studies of FLOSS teams like those performed by Crowston et al., Yamauchi et al., by researchers of other online groups and described in Elinor Ostrom's [20] work on self-organized offline groups can be effective for identifying critical differences and common features associated with group outcomes which can be distilled into design principles and patterns to inform both technological development and group management strategies. Case studies may also facilitate the identification of additional or more complex variables related to project structures and activities, allowing the temporal dynamics of WikiProjects—their formation, development and dissolution—and the role of ecological factors in project success to be modeled computationally.

ACKNOWLEDGEMENTS

This work is supported in part by the National Science Foundation under Grants IIS-0811210 and IIS-1162114. The authors also wish to acknowledge the Wikimedia Foundation's 2011 Summer of Research for supporting the initial development of this work.

REFERENCES

1. Benkler, Y. Coase's Penguin, or Linux and the Nature of the Firm. *Yale Law Journal* 112, 3 (2002).
2. Chen, J., Ren, Y., and Riedl, J. The effects of diversity on group productivity and member withdrawal in online volunteer groups. *CHI 2010*.
3. Cheney, G. On the various and changing meanings of organizational membership: A field study of organizational identification. *Communications Monographs* 50, 4 (1983), 342–362.
4. Choi, B., Alexander, K., Kraut, R.E., and Levine, J.M. Socialization tactics in wikipedia and their effects. *CSCW 2010*.
5. Crowston, K., Li, Q., Wei, K., Eseryel, U.Y., and Howison, J. Self-organization of teams for free/libre open source software development. *Information and software technology* 49, 6 (2007), 564–575.

6. Forte, A., Kittur, N., Larco, V., Zhu, H., Bruckman, A., and Kraut, R.E. Coordination and beyond: social functions of groups in open content production. *CSCW 2012*.
7. Forte, A., Larco, V., and Bruckman, A. Decentralization in Wikipedia governance. *Journal of Management Information Systems* 26, 1 (2009), 49–72.
8. Grudin, J. Groupware and Social Dynamics: eight challenges for developers. *Communications of the ACM* 37, 1 (1994), 92–105.
9. Gutwin, C., Penner, R., and Schneider, K. Group awareness in distributed software development. *CSCW 2004*.
10. Kaptelinin, V. and Nardi, B.A. *Acting with technology : activity theory and interaction design*. MIT Press, Cambridge, Mass., 2006.
11. Kittur, A. and Kraut, R.E. Harnessing the wisdom of crowds in wikipedia: quality through coordination. *CSCW 2008*.
12. Kittur, A., Pendleton, B., and Kraut, R.E. Herding the cats: the influence of groups in coordinating peer production. *WikiSym 2009*.
13. Kittur, A., Suh, B., Pendleton, B.A., Chi, E.H., Angeles, L., and Alto, P. He says, she says: conflict and coordination in Wikipedia. *CHI 2007*.
14. Kriplean, T., Beschastnikh, I., and McDonald, D.W. Articulations of wikiwork: uncovering valued work in wikipedia through barnstars. *CSCW 2008*.
15. Landis, J.R. and Koch, G.G. Measurement of Observer Agreement for Categorical Data. *Biometrics* 33, 1 (1977), 159–174.
16. Lave, J. and Wenger, E. *Situated learning: Legitimate peripheral participation*. Cambridge Univ Pr, 1991.
17. Morgan, J.T., Mason, R.M., and Nahon, K. Lifting the veil. *iConference 2011*.
18. Nov, O. What motivates wikipedians? *Communications of the ACM* 50, 11 (2007), 64.
19. Olson, J.S. and Teasley, S. Groupware in the wild: lessons learned from a year of virtual collocation. *CSCW 1996*.
20. Ostrom, E. Collective action and the evolution of social norms. *The Journal of Economic Perspectives*, (2000), 137–158.
21. Raymond, E.S. The Cathedral and the Bazaar-Musings on Linux and Open Source by an Accidental Revolutionary (Revised edition). 2001.
22. Ren, Y., Kraut, R., Kiesler, S., and Resnick, P. Encouraging Commitment in Online Communities. In *Evidence-based Social Design: Mining the Social Sciences to Build Online Communities*. Cambridge, MA: MIT Press, 1–43.
23. Ren, Y., Kraut, R., and Kiesler, S. Applying common identity and bond theory to design of online communities. *Organization Studies* 28, 3 (2007), 377.
24. Schneider, J., Passant, A., and Breslin, J.G. A Content Analysis: How Wikipedia Talk Pages Are Used. *WebSci 2010*.
25. Ung, H. and Dalle, J.-M. Project management in the Wikipedia community. *WikiSym 2010*.
26. Viegas, F.B., Wattenberg, M., Kriss, J., and Van Ham, F. Talk Before You Type: Coordination in Wikipedia. *HICSS 2007*.
27. Wang, L.S., Chen, J., Ren, Y., and Riedl, J. Searching for the goldilocks zone: trade-offs in managing online volunteer groups. *CSCW 2012*.
28. Wilkinson, D.M. and Huberman, B.A. Cooperation and quality in Wikipedia. *WikiSym 2007*.
29. Wilkinson, D.M. Strong regularities in online peer production. *Proceedings of the 9th ACM conference on Electronic Commerce*. ACM (2008), 302–309.
30. Williams, K.Y. and O'Reilly, C.A. Demography and diversity in organizations: A review of 40 years of research. *Research in organizational behavior* 20, (1998), 77–140.
31. Yamauchi, Y., Yokozawa, M., Shinohara, T., and Ishida, T. Collaboration with Lean Media: how open-source software succeeds. *CSCW 2000*.
32. Zhu, H., Kraut, R., and Kittur, A. Organizing without formal organization: group identification, goal setting and social modeling in directing online production. *CSCW 2012*.
33. Wikipedia Database _reports.
http://en.wikipedia.org/wiki/Wikipedia:Database_reports/WikiProjects_by_changes.