# Crème de la crème

## Elite contributors in an online community

Katherine Panciera
Google Inc.
1600 Amphitheatre Pkwy
Mountain View, California, USA
panciera@google.com

Mikhil Masli
IBM
1 New Orchard Road
Armonk, New York, USA
mnmasli@us.ibm.com

Loren Terveen
GroupLens Research
Department of Computer Science & Engineering
University of Minnesota
Minneapolis, Minnesota, USA
terveen@cs.umn.edu

## ABSTRACT

In open content communities like Wikipedia and StackOverflow and in open source software projects, a small proportion of users produce a majority of the content and take on much of the required community maintenance work. Understanding this class of users is crucial to creating and sustaining healthy communities. We carried out a mixed-method study of core contributors to the Cyclopath geographic wiki and bicycle routing web site. We present our findings and organize our discussion using concepts from activity theory. We found that the Cyclopath core contributors aren't the dedicated cyclists and that the characteristics of the community shape the site, the rules, and the tools for contributing. Additionally, we found that numerous aspects about the surrounding ecology of related systems and communities may help to shape how the site functions and views itself. We draw implications for future research and design from these findings.

## Categories and Subject Descriptors

H.5.3 [**Group and Organization Interfaces**]: Collaborative computing, computer-supported cooperative work, web based interaction

## 1. INTRODUCTION

Social computing technologies have revolutionized the way people connect, communicate, and work together. Our research focuses on the use of these technologies for social production [5, 12]: loosely connected individuals freely coming together to produce information and artifacts of value [10]. Open source software systems and Wikipedia are prototypical cases. For example, tens of thousands of people from around the world have written over three million articles on Wikipedia making it the largest encyclopedia in history.

Prior research has shown that much of the content and value in communities like Wikipedia is produced by a small proportion of the community members: highly active *core contributors*. Furthermore, these individuals also take on the majority of community maintenance work, such as welcoming newcomers and ensuring that they follow community norms, policing content contributed by others to verify that it meets quality standards, arbitrating conflict, etc.

We build on and extend prior research on these questions by studying core contributors to Cyclopath (http://cyclopath.org), a bicycle route-finding system and geographic wiki for the Minneapolis / St. Paul metropolitan area. Our results shed light on who the core contributors are, how they use the platform and tools and conduct community maintenance, and how they are embedded in the bicycling community, offline and online, in Cyclopath and other community sites. We also compare our results to findings from studies of other open content systems.

Much research on online communities assumes that all communities are the same or at least similar. By contrasting the results from Cyclopath with data from Wikipedia, MovieLens,[1] and Everything2,[2] we seek to show that detailed characteristics of system and communities help to shape the user experiences within the community.

In the remainder of the paper we briefly survey related work and situate our research with respect to it, describe our methods and datasets, present our findings, and discuss their meaning and implications.

## 2. RELATED WORK

Core contributors are important to the survival of an online community. Work on Wikipedia by Priedhorsky et al. showed that the top .1% of all editors, by number of edits, create 44% of the value of Wikipedia [23]. Prior work on Wikipedia by Panciera et al. found that core contributors (the top 2.5% of editors by number of edits) on that site contribute more edits per editor than the rest of the registered editors, starting on their first day of editing [18]. Similarly, when researching the same issue on Cyclopath, Panciera et al. found that core contributors on Cyclopath started off editing more than the rest of the editors on the first day that they had viewed the site. They also found that editing anything increased retention and time on the site [20].

While these findings look specifically at volume and value of edits, the research presented here attempts to understand more about the core contributors themselves: what skills, knowledge, and experience they bring to Cyclopath and what value that may provide.

There is also a separate body of work that aims to understand who the core contributors in online communities are qualitatively. Antin interviewed Wikipedia readers about who they

[1] http://www.movielens.org
[2] http://www.everything2.com

thought the Wikipedia editors were. He found that people characterized Wikipedia editors in one of three ways. They thought Wikipedia editors were average people, well-educated intellectuals, or the 'solitary techno-geek' [4].

While our work does not aim to understand the perceptions of Cyclopath editors, but rather the editors themselves, these characterizations of Wikipedia editors suggest different types of expertise that may also play a role in other systems. In particular, there is a distinction between content domain expertise and technical domain expertise. In this paper we look specifically at whether the Cyclopath core contributors have content domain or technical domain expertise.

Bryant et al. interviewed the Wikipedia editors themselves to try to understand what they were like as novices and how they changed as they became expert editors. The researchers found that editors went from editing what they knew (single pages) to editing in order to improve Wikipedia as a whole. They also discovered that the interface helped get users started editing, but as editors increased their participation, other tools, like discussion pages, became more useful. New editors also were not likely to view Wikipedia as a community, but as they participated more, they saw the community as the key to the site. Finally, Bryant et al. found that editors become more involved in governance as they become more experienced [7].

This paper aims to take these findings about Wikipedia editors and see whether similar findings hold true on Cyclopath. We have used the same activity theory framework as used in Bryant et al., but will show differing results, due, in part, to the difference in content, community, and the nature of the site. For example, discussion about rules on Wikipedia is intense and people discuss general policy as well as content. In contrast, Cyclopath users rarely converse and if they do have discussions, they are much calmer than debates on Wikipedia and tend to be focused on a specific instance rather than general policies.

One of the most frequently talked about papers in regards to users and editors in online communities is Preece and Shneiderman's Reader to Leader Framework. This framework, built on earlier literature, suggests that there is a ladder that users work their way up as they become more active on a site. First the user is a reader, a consumer who is using the site. Then, a fraction of the readers start editing and become contributors. A portion of those contributors begin to work with others to build the resource, these are the collaborators. Finally, the leaders are the collaborators who begin to work with issues of governance and policy [21].

While our paper defines user groups differently (users, editors, core contributors) than Preece and Shneiderman, we find that the core contributors have difficulties with some of the responsibilities that come with being the leaders. In particular, we discuss rules and how the core contributors interact with the policies of the site and help to create those policies.

Finally, there is a range of work that looks at the motivation of users to contribute to open content sites. Clary et al. developed a tool called the Volunteer Functions Inventory (VFI) in order to help measure the motivations of people volunteering offline [9]. Nov used a modified VFI to asses the motivations of Wikipedia editors and found that fun and ideology were the top motivations [17]. Fugelstad et al. used a slightly modified VFI to assess people's motivations for using and participating in MovieLens and found that self-oriented motivations were negatively correlated with logins and ratings [11]. Everything2 is a peer production site that, while originally envisioned in 1999 as an open source encyclopedia, is now a writing site. Researchers using Everything2 as a platform found that their users were motivated, in general, by one of four things: increasing their status on the site, building relationships with other users, improving the community as a resource, or increasing their individual skills [32]. Finally, Panciera et al. looked at motivations for editors to participate in Cyclopath from a pragmatic perspective and found that people started editing Cyclopath in order to fix problems [19].

We build on this work, in part by using the VFI to assess the motivations of the Cyclopath core contributors. We also strive to understand how tools, tasks, and the overall experience has changed as the system ages and the users become more experienced. Other work on Cyclopath has studied contributors' on-site behavior [20], how they specialize by type of work [14], their motivations (as compared to those of consumers) [19] and techniques to elicit contributions from non-contributors [15, 24]

This paper uses Cyclopath as a platform, in part because related research has been done on Cyclopath and Wikipedia, and therefore Cyclopath allows us a point of comparison. However Wikipedia differs from Cyclopath in a number of ways. Cyclopath is a local site. Due to the fact that the site only serves a seven county metro area surrounding the Twin Cities, most of the editors are local, and, as such, have greater potential to know each other offline. Cyclopath also requires editors to make geographic edits, something most sites do not require and something that is inherently more difficult than textual editing. The site computes routes that are based on the resource and, as such, the edits. The primary goal of the site is to serve as a route-finder to bicyclists, and the wiki-nature of the map helps to serve that purpose. Finally, Cyclopath is much smaller scale than Wikipedia, and is therefore more typical of open content communities.

While there are many differences between the two, Cyclopath and Wikipedia share the fact that they are both wikis and information resources. They also both require minimal expertise from contributors.

# 3. RESEARCH PLATFORM: CYCLOPATH

Cyclopath is a web-based, bicycle routing system for cyclists in Minneapolis and St. Paul, Minnesota. In addition to computing bicycle routes, Cyclopath also is a geographic wiki–a map that anyone can edit and annotate. There are tools to let users rate the bikeability of road and trail segments, to add tags and long-form textual notes to segments, and tools to edit the map itself, adding new segments and connecting them with existing segments. The authors are part of the team working to develop Cyclopath.

As of May 25, 2013, there were over 5,100 registered Cyclopath users. Registered and anonymous users have requested over 128,000 routes since the site was launched in May 2008, typically requesting around 150 routes per day during bicycling season. Almost 1000 registered users have edited Cyclopath, collectively making over 17,000 revisions. We classify the 22 users (about 2.2%) with the most revisions as core contributors. We chose 22 users to be consistent with previous research on Cyclopath [19, 20].

# 4. EXPERIMENTAL DESIGN & METHODS

We followed a mixed methods research strategy, combining usage log analysis, survey results, and interviews.

**Usage Log Analysis**. Research that required Cyclopath logs comes from the Cyclopath usage history dataset that includes all activity on the site between launch on May 15, 2008 and July 6, 2012.[3] This dataset includes 15,998 edits, 13,123 of which were

[3]The data collection and analysis was done at this time and we believe it is still valid. In the two years since the analysis, fewer than

made by logged in editors. There are 906 editors in this data set. The remaining 2,875 edits were made by anonymous editors.

**Survey**. This research also relies on survey results from the survey also used for [19]. This survey was developed and administered to Cyclopath users in March and April 2010. The survey had 60 questions ranging from standard scales to custom designed items. No survey questions were required to be answered. Five questions (not analyzed here) were previously analyzed in [19].

- A version of the Volunteer Functions Inventory (VFI), originally developed by Clary et al. [9] and modified for use in online community research by Fugelstad et al. [11]. The VFI questions were grouped into six categories, as described by Clary et al.: Values, Understanding, Enhancement, Career, Social, and Protective. Three statements were included for each category, for example: *I feel compassion towards others in need.* Users were asked to rate how accurate or important each of the statements was on a seven point Likert scale.
- Questions about bicycling habits, including: the months a respondent's cycling season began and ended and how often and how far they rode.
- Questions about other websites used, including both bicycling sites (such as mplsbikelove.com and bikely.com) and online/open content communities such as Wikipedia, Facebook, Twitter, and YouTube.
- Questions about Cyclopath, including how respondents learned about Cyclopath features.
- Standard demographic questions.

We had 396 users fill out at least part of the survey. The average age of the users was 42% were male and 21% female (the remainder didn't answer the question). The users averaged four hours a day online and an additional five hours a day on the computer offline.

**Interviews**. To gain more in depth knowledge about core contributors, we conducted seven interviews in spring 2012. These were semi-structured, face-to-face interviews where all participants were classified as core contributors. Details on the participants are shown in 1. All the interviewees were male and they ranged from late twenties to late sixties.

During the interview we asked about cycling habits and preferences, familiarity of the participant with the Twin Cities and the Twin Cities bicycling community, the participant's use of Cyclopath, their perceived role on Cyclopath, editing idiosyncrasies, and their thoughts on the Cyclopath community. In addition, to allow direct parallels to Bryant et al., we also asked questions used in interviews by Bryant et al. in [7]. These questions were used with the permission of PI Andrea Forte and were adapted slightly for the context of Cyclopath.

**Summary of user categories**. For ease of reference in the following discussion, we summarize the different categories of Cyclopath users we have introduced.

> **Cyclopath Users**. All users of the Cyclopath website, registered and anonymous.
>
> **Cyclopath Editors**. All users of the Cyclopath website who have edited or annotated the map.
>
> **Core Contributors**. The 22 Cyclopath editors (top 2.5%) who made the most edits.

100 users have registered for Cyclopath and only one significant new feature has been introduced: route saving.

| Participants | Days Active | Revisions |
|---|---|---|
| Participant 1 | 1218 | 500+ |
| Participant 2 | 905 | 0 - 125 |
| Participant 3 | 568 | 250 - 500 |
| Participant 4 | 643 | 125 - 250 |
| Participant 5 | 1148 | 125 - 250 |
| Participant 6 | 1064 | 500+ |
| Participant 7 | 1015 | 500+ |

**Table 1: Interview Participants. Since Cyclopath revisions are public, we bucketed the umber of revisions to preserve participant anonymity. Number of days active is computed as** $date\ of\ last\ view - date\ of\ account\ creation$**.**

> **Survey Respondents**. The 396 Cyclopath users who responded to the 2010 survey. (Note that on any given question we would have 396 or fewer responses.)
>
> **Interview Participants**. The 7 Cyclopath users who the authors interviewed; all the participants were core contributors and four also took the survey.

**Framing our findings with Activity Theory**. Activity Theory is a descriptive theory for studying and interpreting human activity. It has been useful in HCI studies [13] and was the very successful primary framing for Bryant et al.'s Wikipedia study [7].

At an individual level, Activity Theory looks at a *subject* who engages in an activity directed at an *object* mediated by a *tool*. However, Activity Theory also has a framework for community level analysis which introduces the constructs of *rules*, *community*, and *division of labor*.

We have chosen this framework to present our results because it was a useful way to organize and interpret our findings and to contrast our results with [7]. We report on subject, object, community, rules, and tools. Our interviews, surveys, and usage logs do not adequately address division of labor, so we do not present any results about that, but we mention it in our discussion.

# 5. CORE CONTRIBUTORS & ACTIVITY THEORY

## 5.1 Subject

The subject in activity theory is the person engaged in the activity. For our study, we chose to focus our study of subject on the types of expertise that the contributors had. Core contributors may often be viewed as experts, as in Bryant et al. [7]. We wanted to look at and differentiate between two types of expertise that Cyclopath core contributors may have. The first is the content domain, that is, bicycling. In order to evaluate expertise in the content domain, we looked at whether core contributors were dedicated bicyclists. The second type of expertise is in the technical domain, online open content communities. To evaluate expertise in the technical domain we looked at whether core contributors were familiar with and engaged in other online open content communities.

*Content Domain*. Previous work [6, 16] categorizes bicyclists by factors such as distance ridden, frequency of rides, and conditions ridden in. We use the same three factors in our definition. Cyclists are considered dedicated if they meet at least one of the three conditions below.

> Distance: Ride 20 miles or more on average
>
> Frequency: Ride one or more times a day (during riding season)

> Riding conditions: Ride at least nine months of the year (forcing them to ride during at least part of the cold, potentially snowy, months)

These results combine survey data (cycling behavior) with log data (core contributor status). 186 of 299 survey respondents are classified as dedicated cyclists by the above definition There is no statistically significant association between being a core contributor and being a dedicated cyclist ($\chi^2 = 0.0047$, $df = 1$, $p = 0.95$). Even if we look at distance, frequency, and riding conditions independently, there is still no significance.

(A survey by Bike/Walk Twin Cities found that 20–36% of cyclists continued cycling over the winter[8]. This compares to over 50% of our survey takers. We realize that this means that our survey takers, and perhaps Cyclopath participants, aren't representative of the entire cycling population of Minnesota and this shows that our user population consists generally of more dedicated bicyclists.)

*Technical Domain*. We wanted to assess Cyclopath users' familiarity and participation in open content communities. In our survey we asked users whether they were a member of, used, and/or contributed to Wikipedia, Twitter, Facebook, or YouTube. In our user population we found that the only significant differences were with Wikipedia.

A significant number of core contributors used Wikipedia ($\chi^2 = 4.47$, $df = 1$, $p = 0.03$) and contributed to Wikipedia ($\chi^2 = 5.27$, $df = 1$, $p = 0.02$). In addition, core contributors were marginally more likely to be members of Wikipedia ($\chi^2 = 3.46$, $df = 1$, $p = 0.06$).

In our interviews, we were able to discuss this more in depth. One of our interview participants contributed to Wikipedia. Participant 5 edited Open Street Map (OSM), another open content mapping project. He wanted to be able to edit in Wisconsin, because he often bikes there but Cyclopath doesn't cover the area. (Cyclopath actually has measures in place that block users from editing in Wisconsin, which several interviewees mentioned.)

> [While on a trip] I discovered the trail is closed because of some old train wreck... They haven't finished repairing the trail. So I actually went around the tracks and found out that they actually had a detour in place... and I thought, "Well, the maps [OpenStreetMap] ought to reflect the detour." So I did that. (Participant 5)

While they may not be active contributors to other open online communities, core contributors do value knowledge contributed by their peers. This is consistent with prior research [26] which found that bicyclists used information from other bicyclists frequently. The users also valued and trusted information obtained this way.

> In most cases I trust it more because it is people who update it for the most part.... Just knowing that the information has been entered in by people is comforting to me. (Participant 4)

Survey results were consistent with this. All responding core contributors said they used Wikipedia. When asked about how much they trusted services (Cyclopath, Wikipedia, Facebook, and Twitter), all seven core contributor respondents listed Wikipedia or Cyclopath as the most trustworthy site. When asked to rank the value of the same set of websites, five out of the seven said Wikipedia was the most valuable and the remaining two found Cyclopath to be the most valuable.

*Identity*. We were surprised to find that core contributors saw themselves as users, not contributors or elite users. When asked "How would you describe your role on Cyclopath?", four of the interview participants used the term "users". Three also used qualifications like "engaged user", "user that uses Cyclopath a lot", or "occasional contributor". Only one of the seven said "contributor". One participant called himself a "lurker".

*Discussion*. We found that core contributors were not the most dedicated cyclists, but they were similar to each other in that they had belief and engagement in open content knowledge production, both in Cyclopath and beyond.

Our finding that core contributors were not the most dedicated cyclists was somewhat surprising to us as we'd expected core contributors to be more dedicated cyclists than the other survey respondents. However, examining our data, in conjunction with the Bike/Walk Twin Cities data [8], we realize that our sample is likely skewed. If we were to compare our core contributors to the average cyclist in the Twin Cities, it is likely that the core contributors would be much more likely to be dedicated cyclists than non-Cyclopath users.

Belief and engagement in open content is a characteristic we had expected to see in our core contributors and we were not surprised by these results. There may be ways to leverage this trait when building systems relying on user contributions.

Previous work on Wikipedia showed that editors often began in their areas of personal expertise and later branched out to new content [7]. Other research showed that while 92% of Wikipedia editors declare themselves to be "proficient in computers", only 36% of editors are programmers [3]. This indicates a different pattern than we've witnessed on Cyclopath. In addition, Wikipedia editors see themselves as contributors who transition to caretaker or creator roles instead of simply "users" [7]. More work is needed to tease out why these differences occur.

## 5.2 Object

In our case, the Cyclopath system is the object of the activity (editing) being performed by the core contributor. However Cyclopath has two main interaction paradigms. Like Wikipedia, Cyclopath is an open content system based on end users editing content. But Cyclopath is also a computational service: computing bicycle routes [25]. These two paradigms are linked. User input influences routes that are computed. The paradigms have implications for users as well. Like a recommender system, e.g. MovieLens, users benefit personally from data they input: they receive routes that better match their preferences, especially when they input bikeability ratings. Note that Cyclopath users don't have to input any data in order to get a route, unlike MovieLens. Without personal ratings or data, Cyclopath uses aggregate ratings and objective properties of trails and roads. In MovieLens, all movie ratings are private and only displayed in aggregate while edits to the movies or posts in the Q&A sections are public. However, unlike MovieLens, the public effects of user edits are more prominent in Cyclopath. Many types of user edits are public: geographic edits, notes, discussions, and tags.

We were interested in how the user input influencing the routing engine may affect attitudes and goals of core contributors.

*Building a Better System*. Core contributors participate because they want to build a better computational system. Contributing gives them power to improve the output (e.g. routes) of Cyclopath.

> I just liked the idea of this routing tool that was actually pretty broken, but I could make it work by fixing things. (Participant 1)

> It is not designed to be usable at the outset: it is designed to give powers to users to make it work. (Participant 3)

In some cases, users added information to Cyclopath so that the route-finder (here referred to as Cyclopath), not others, could use it.

> In my neighborhood, there was a little sidewalk that actually gets you to Target and without that there, Cyclopath wouldn't know it existed... So I made that little that path to say it actually goes there so that Cyclopath can actually see it. (Participant 3)

The brokenness of the computation of Cyclopath (its ability, or inability, to compute routes with the information it had) served as a motivator to core contributors to "fix it" by contributing. This was especially true early on. Half of the core contributors started using Cyclopath before the public release on August 1, 2008. The state of the map at that point was poor. There was a lot to be done. A number of interviewed core contributors mentioned that they have reduced their editing over time because there was less to do (or to fix).

> I still think I edit more frequently than anybody else, but I think I probably edit less, a fair bit less, than I did, largely because the area that I know well is already in pretty good shape. (Participant 1)

The amount of information and quality of the information in the map varies from location to location. Parts of the Twin Cities metro area where Cyclopath users are the most interested in and familiar with have great coverage. Outer suburbs and rural areas have had much less attention paid them, but are also routed through less frequently.

The interviews supplement the results presented in [19] which show that 37.04% of surveyed editors began contributing to fix a problem, although in those results, no editors have said that they continued contributing solely to fix problems.

*Discussion*. Users were motivated by creating an effective computational resource and fixing problems. Previous research also found fixing problems to be a key motivation for people editing Wikipedia [7, 17]. In Cyclopath this is definitely compelling. In addition, the Cyclopath routing engine (e.g. computational aspect) is important and sometimes leads to cycles of requesting routes, editing, and then re-requesting the original route.

Many systems requiring user input rely on problem-fixing to draw in users. However this requires a delicate balance between problems and functionality. Without functionality, users may not see the value in contributing or participating in the system. However, if the content is fully formed, users may not realize where the opportunities for contribution lie. Also, keep in mind that as systems differ, some projects may benefit more from this sort of approach than others.

## 5.3 Communities

In this section we focus on community, an interesting concept in Cyclopath. The community is strongly tied to the rules and the division of labor, but here we look in depth at the ties between the community and the core contributor.

Cyclopath provides features for community interaction in addition to the map and route finder features. When Cyclopath launched in May 2008, a Google Group (mailing list) was also launched. This was used for discussion among users and developers. In particular it was used for bug reporting, troubleshooting, and discussions about governance (see more in 5.4). In addition to the Google Group, there was also a user editable text wiki that allowed for user and developer documentation. Finally, in April 2010, an in-application discussion system was launched. This system functions as a forum (although the Google Group is still used) and allows users to tag discussions with words or locations (users can tag a certain segment of a certain route).

However, in comparison with Wikipedia, these community resources are rarely used. Every Wikipedia article has a talk page where discussions regarding content and organization of the page happen. In addition, every user can create a user page with an associated user talk page. Talk pages are "the most frequently used communication channel on the Wikipedia" [7]. There is also the Village Pump, an area for Wikipedia core contributors to discuss "technical issues, policies, and operations of Wikipedia".[4]

Cyclopath has a potential advantage over Wikipedia in that Cyclopath is closely tied to an offline, geographically local community. From August 1, 2011 to July 31, 2012, Google Analytics shows that Cyclopath had 32,011 views. 27,447 of those were from inside the state of Minnesota, 14,709 from Minneapolis and 4,705 from St. Paul. This means that the person who bikes past you on your commute could be the person whose edits helped you find your route. In Wikipedia the primary page editor could be halfway across the world.

*Online Bicycling Communities*. We wanted to know if Cyclopath core contributors participated in other online bicycling communities. In our 2010 survey we asked users if they participated in Minneapolis Bike Love (a local bicycling forum), Bikely (a route sharing site), Bikemap (another route sharing site), Gmaps-Pedometer (a route distance calculator), MocaTrails (a local off-road club) , or other online bicycling communities.

A chi-squared analysis showed that core contributors weren't more likely than other users to participate in these online communities.

> "I do participate in the Minneapolis Bike Love message board... Mostly just reading it, I usually don't post... I just like to stay up to date and it's always nice to hear what people are up to, interested in.... I tend to focus on threads involving safety or conditions.... It's purely functional use." (Participant 4)

*Offline Bicycling Communities*. In contrast to the above, core contributors were active in the local cycling community offline. Participant 3 said the following when he was asked if he was attached to the local cycling community:

> Not at all... To me it's just not feasible in one day to spend three hours traveling to an event. (Participant 3)

But as he continued, it was obvious that he was fairly active in the local cycling community.

> I go to events and I am participating in the cycling community that way... but there's no attachment... I went to a bike summit... I have an LCI [League Cycling Instructor meet up] down at Peace Coffee... (Participant 3)

Interviewees who did say they were attached to the local cycling community were often involved in cycling advocacy groups, not just riding groups. Participant 2 talked about what he got out of participating in a local bicycle coalition:

[4] http://en.wikipedia.org/wiki/Wikipedia:Village_pump

> I think I get more opportunity for community input affecting the city and the neighborhood and where I live and then also kind of get to work to foster change to benefit the community as a whole and provide more options for cycling. (Participant 2)

Participant 1 was the only user we talked to who was active in both online and offline cycling communities.

> Well, besides Cyclopath, Minneapolis Bike Love, the Minneapolis Bicycle Coalition, the bike alliance of Minnesota... (Participant 1)

*Reasons for Volunteering*. In trying to understand more about core contributors and other contributors, we wanted to understand why Cyclopath editors were editing and volunteering. To do this, we administered a version of the Volunteer Functions Inventory (VFI) [9] as modified by Fugelstad et al. [11].

In the VFI, we were particularly interested in two factors. The first of these was value. This is a factor that would come up if you were volunteering because you are acting on personal values. This would likely be a factor if you were helping people less fortunate than yourself or doing humanitarian work. The second factor was social. This factor is strong if you are volunteering to strengthen social ties or because your friends volunteer. The higher the number for the VFI factor, the more the survey participant felt that the factor was an important reason for them to volunteer or an accurate description of why they volunteered

We calculated correlation coefficients using the VFI factors against the total number of revisions that the user had made. The number of revisions an editor made was negatively correlated with the value factor ($r = -0.20$, $p = 0.002$) as well as the social factor ($r = -0.15$, $p = 0.032$).[5]

This result shows that the more revisions a user has, the less likely that that user is volunteering due to values or to receive social benefits. This is similar to the results found on MovieLens by Fugelstad et al.[11].

*Awareness of Others*. Perhaps because core contributors are not participating on the site for social reasons, they know few contributors. When asked whether they know anyone of Cyclopath or recognize the usernames of other editors, four interviewees said one to "a handful" of usernames were familiar.

> Sure, I recognize some of the names in Cyclopath. I don't know any of them though. (Participant 1)

Some of the users used this name recognition to identify users in the recent changes list. This allowed them to focus their effort on changes that might be suspicious, instead of on changes that were likely well done.

This differs from the findings of Bryant et al. of core contributors [7] as well as findings of Velasquez et al.[32] in interviews of Everything2 users. That research found much higher awareness of other editors. On Everything2, some users were "status builders" or "personal relationship builders" and were very conscious of their audience and had personal relationships with other users on the site.

*View of System*. We asked people how they would describe Cyclopath to others. The answers we got were all similar to the quotes below.

> Cyclopath is an alternative to Google Maps with editable functions where locals can edit to show its accuracy. (Participant 3)

> Cyclopath is a site that lets you find bike routes. It is also a wiki, community-supported. (Participant 6)

> It is a cycling geowiki... [To non tech-savvy people] it is a way of finding a bike route that you can modify and set parameters. (Participant 7)

We were surprised that none of the participants explicitly described Cyclopath as a community, but the core contributors did not see Cyclopath as a medium for social interaction, but rather as a resource that is modifiable by its users.

*Discussion*. Despite the way Cyclopath researchers think about the site, core contributors don't view Cyclopath as a community and don't seek out social interaction. In addition, they are not motivated by the prospect of social interaction. To contrast this, in the early days of Wikipedia, Bryant et al. found that initially users didn't see Wikipedia as a community, but that changed as people became experts [7].

When Cyclopath was introduced, Minneapolis Bike Love was an existing online discussion community for Twin Cities bicyclists. Cyclopath was an information resource, in contrast to Minneapolis Bike Love's discussion forum and social interaction. In talking to users, it seems like this dichotomy still exists. This is consistent with prior research on MovieLens. When social interaction was introduced on MovieLens experienced users, who saw MovieLens as a film recommendation site, did not respond positively to those features [28]. Wikipedia had few rivals in early days and also had the infrastructure in place to support social interaction such as User pages and Talk pages.

While many sites today are trying to figure out how to build community and incorporate social interaction, users may not want those features or the features may be conflicting with other systems. The complete ecosystem that a site exists in can be an important factor when designing or changing functionality.

## 5.4 Rules

Rules in the context of Activity Theory refer to socially defined and enforced norms, conventions, and social relations [1]. They are influenced directly by the community and the individual.

Some of the rules in Cyclopath were derived from other open content communities. For example, Wikipedia has a rule that no private data should exist in a public resource. While there is difficulty in enforcing this type of rule from a system perspective (simply not allowing said edits), this type of rule is enforced, in general, by other users or the users themselves after the fact. This is the reason for 147 of 628 (23.4%) reverts done in Cyclopath. (These are reverts that contain the word "private" in the comment accompanying the revert.)

*Establishing Rules and Norms*. Cyclopath core contributors want rules to help create a uniform resource and to handle controversies, questions, or problems incurred while editing. One of the recurring discussions regarding rules and norms revolves around terminology used in tags. Tags are used for routing and a uniform vocabulary mask using tags, applying tags, and routing with tags easier. For more information on tags in Cyclopath, see [31].

The following is a note that appeared in the Cyclopath discussion section in an attempt to try to determine whether the correct tag should be "bikelane" or "bike lane".

> 1) Is one term of other preferred? 2) If so is it documented anywhere? 3) If not, should it be? 4) If so, where? In the [text] wiki? 5) If so, would you expect users like me to just jump in and start proposing (in the wiki) some standards? (Cyclopath Discussions)

[5]Regarding correlation coefficients, Rosnow and Rosenthal have argued that small effects can have a powerful impact on outcomes over time, especially in the aggregate. [29]

As a result of the subsequent discussion, users developed conventions for when the "bike lane" tag should be applied.

> It's been our convention to mark a two-way road as having bike lanes only if it has lanes in both directions of travel. Pleasant has only a general travel lane (marked with poorly placed sharrows) in the southbound direction and a counter-flow bike lane northbound so it doesn't seem a good candidate for the bike-lane tag. (Cyclopath Discussions)

One user we interviewed gave another example. He had spent hours simplifying roundabouts, distilling the eight or more segments to a simpler four way intersection. In retrospect though, he wondered if roundabouts needed to appear in detail or if his simplification was acceptable. This is something that is very rare, so the community has not developed any visible guidelines. That said, the user didn't want to overstep or do something that would be considered "wrong" by others.

> I never know if that's acceptable, what the proper way to deal with situations like that is. (Participant 4)

*Benefit of Lack of Rules*. A lack of system enforced rules allowed for some innovation on Cyclopath. One interview participant added links to pictures of complicated intersections to points near those intersections. This is discussed more in detail in 5.5.

While some users may find the enforced rules useful, environments with more (and more strongly enforced) rules and policies (e.g. Wikipedia) may not have allowed for some of this innovative behavior, instead squashing such behavior.

*Opacity of Editing Rules*. Cyclopath rules are, in general, difficult to find and perhaps to follow. This makes it both hard and nerve wracking for users to begin editing.

> I was terrified to edit and then I tried a few things, and my terrors were justified... Well, make a mistake, do something that is wrong that is going to inconvenience someone else, that somebody else needs to clean up your mistakes... There's so many discussions about reversions and tags...I read this and I'm terrified to start, absolutely. (Participant 5)

Another participant had an edit reverted because he didn't know what the community norms were.

> Those [mistakes] are scary. You don't want to edit again for a while. It's like I don't know what I'm doing anymore. (Participant 4)

*Dangerous Consequences*. Cyclopath is geographic and is used for finding routes, this allows for editing situations that are likely unique to this system. Interviewees were worried about getting things on the map wrong. In essence, editors wondered, "what happens if due to my screw up, someone goes on the wrong road and gets hurt". To our knowledge, this has never happened, but core contributors are worried about it and think about it while editing.

> I hesitated for a long time to rate the Washington Ave. Bridge. I know that it's not legal to ride there but I am sometimes willing to do it so I would rate it something better than "impassable". However, when I did that the default rating change [sic] from "impassable" to "poor". Does that make me in a way responsible when Cyclopath routes an unfamiliar rider over that road? (Google Group)

There are places (Google Group, Cyclopath discussion forum) where rules can be discussed. Core contributors were frustrated that rules were often hidden and they weren't able to learn the norms until people (in discussions or revert comments) came and told them explicitly.

Despite these problems and issues, Cyclopath is young and the rules and available resources are evolving as the community ages.

*Discussion*. Some Cyclopath rules are based on those from other open content systems, but Cyclopaths unique features led to unique rules and a desire among core contributors for more rules.

This is an area where communities may differ significantly. For communities like Wikipedia and Cyclopath, letting the community make their own rules has, for the most part, worked. In other communities, such as 4chan or reddit, it could be argued that, while they don't have many rules, the lack of rules has led to a much more volatile community.

We feel that the organic community-creation of rules in the Cyclopath community has benefited the community and the system. As the community has developed rules, they have also, as shown in this work, begun enforcing their set of rules. This is a necessary part of almost every community, so, as with many of these factors, it is one part of the puzzle.

We heard from our participants that the potential for repercussions in the physical world made them pause and sometimes delayed their editing. While we think it likely that this is a significant differentiator of a geographic wiki, we don't have similar data for OpenStreetMap, Wikipedia, or other contribution sites.

## 5.5 Tools

While traditional Activity Theory looks at the tool in relation to the individual and the site, in the previous rules section, some tools were mentioned. The consensus building tools (discussion, Google Group) will not be discussed here as here we focus on the individual experience.

The Cyclopath web interface has a number of tools for different functions. Users can rate bikeability of roads and trails, add tags, add text notes, participate in discussions, save routes, edit map geography and topography, create new segments, and connect segments with existing segments. Yet the number of tools and complexity of the tools make editing in Cyclopath harder than other systems such as Wikipedia.

*Learning the Tools*. Our survey asked editors how they learned to use the interface and editing tools. (Note that users could list multiple responses.)

321 (81%) taught themselves, 104 (26%) learned from the text wiki, and 43 (11%) learned from existing video tutorials. Of the nine core contributors who took the survey, eight taught themselves (while sometimes referring to the text wiki or video tutorials if they couldn't figure something out).

*Using the Tools*. When they started using Cyclopath, even the core contributors thought the editing tools were hidden and awkward. Half of the core contributors were early adopters, joining prior to the public launch and at that point the interface was still in flux.

> Earlier on, there were problems. It was quite clunky, especially doing a lot of edits... Like breaking and merging intersections or moving a whole segment and things like that... (Participant 6)

Since this time, core contributors agree that the site has gotten easier to use and the process of editing is less clunky and more robust than it was initially However, this shows the potential importance of a good initial user experience. Had the map itself

or the tools been harder to use, adoption may not have been as widespread.

*Wanting More Tools*. We found that core contributors are opportunistic in nature. They are interested in additional tools and often think in terms of Cyclopath when riding. In particular, they think about how they can enter data relevant to where they are at the time and how they can enter information about problems they encounter. One interviewee drew an entire area of the map freehand. He drove around the newly developed area often for other reasons.

> The... area had no satellite imagery and I drew it out freehand just by driving around because I went there for work anyway. Get back home at the end of the day and draw it out. When the satellite imagery actually came out, someone even commented that whoever did the drawing did a really great job. (Participant 7)

Participant 5 edited OpenStreetMap, in part because OpenStreetMap is able to import GPS traces and use these traces to automatically make map edits.

> If there's a detour on the trail, I like to ride it so that I get a GPS trace so that I can edit OpenStreetMap or Cyclopath. Snipping GPS traces and adding info to OpenStreetMap is easy. (Participant 5)

(Since the interviews, Cyclopath has launched a mobile app that allows users to upload GPS tracks.)

Other participants also mentioned that while riding they would discover edits that they wanted to make, but they weren't always able to recall these when they got to their computers to edit.

> On really long rides, it is really hard to remember sometimes and I don't have any fancy technology that allows me to update it while I'm riding, which would be really nice. (Participant 4)

Participant 3 solved this problem by carrying a notepad or camera around to record problems as he encountered them.

> That's why I take pictures... so that I can take a picture of what I see and then come back and [record edits on Cyclopath]. (Participant 3)

One core contributor wanted keyboard shortcuts to make editing faster. When Cyclopath developers said they didn't have the time, he downloaded the open source Cyclopath software and wrote the code himself. The developers merged the code into the main code base and the shortcuts are still used.

While the Cyclopath tool set is not ideal for all users yet, the group of core contributors has learned and embraced the tools and started to request additional new tools.

*Discussion*. Cyclopath has various tools ranging from easy to use (rating a block) to quite hard (dividing blocks into multiple segments). These geographic editing tools are unique to Cyclopath and OpenStreetMap. (However the two systems use different tools.) These tools allow users to correct and extend the map. Editing has quantitatively improved the routes Cyclopath generates [24, 25]. However the complexity of tools creates an entry barrier. The discussion forums and route library offer a way for users to contribute more easily, but it is not clear that lowering the barrier to entry in this way would attract many more contributors.

The core contributors suggest that there's a need for additional tools, in particular tools that would be of use for recording while riding. An example of a tool that may be of use would be Biketastic [27]. Previous work on a citizen science river monitoring application showed a similar need for additional tools [30]. Now that Cyclopath has a mobile app with GPS, it would be possible to incorporate note-taking functionality into the app though this has not been done yet.

On Wikipedia, making initial edits is relatively easy, but expert users go beyond the editing UI to additional tools such as the watch list [7], bots, user scripts, and automated tools [3]. Like the rules of Wikipedia, the tools have also evolved as needed over the history of the site. One of the problems with Cyclopath is that even making the initial edits is difficult. Allowing users to make simple, non-geographic edits more easily may hook more users and elicit more contributions.

## 5.6 Division of Labor

People familiar with Activity Theory may notice that we have not analyzed core contributors with respect to Division of Labor. We found during our research that while we think of, and refer to, Cyclopath as a community, the people in the community tend to be fairly independent and not interconnected. The core contributors were barely aware of each other and definitely not negotiating anything in regards to division of labor. (Note that, as mentioned in the Rules section, they did participate in consensus building and rule enforcement.) This is in contrast to a site like Wikipedia, where editors often have defined roles and spend time negotiating changes to Wikipedia pages, among other things [7].

# 6. DESIGN IMPLICATIONS

We found that Cyclopath core contributors aren't the most dedicated bicyclists, but they are engaged in open content systems and believe in open content knowledge. Our research also showed that community characteristics matter and help to shape the site and draw (or repel) certain users. Finally, we discovered that while we, the researchers, view and think of Cyclopath as a community, users think of it as a resource and not a destination for social interaction.

This demonstrates yet another reason for people building online communities to understand their users at a deeper level. By understanding primary motivations of core contributors, builders can create communities that better suit the needs of users and also attract core contributors. For example, if users are attracted more to open content than the specific content of that site, make sure that all marketing materials (including the community itself) promote the open content concept. Open content wouldn't have to be promoted to the exclusion of other material, but it should be front and center.

This work also argues for talking to users, especially core contributors, early and often. When Cyclopath was founded, research was done to try and understand the user base [22]. Those initial interviews were incredibly useful and helped form the site that exists today, but they were done before the site was created. Subsequent interviews weren't conducted until spring 2012, when the site had been live for four years. If we'd talked to users in between, we may have discovered some patterns or learned more about core contributors that would have helped form the site or inform some of the potential design changes.

One other idea to highlight is that of sandboxing. Wikipedia and other wikis have pages called sandboxes [2] that allow users to experiment and try edits out before posting them to the actual websites. This is a concept regularly used in software testing as well to test things before pushing them live. Many of the interview participants we talked to mentioned being nervous about making edits or actually breaking things. For sites like Wikipedia and Cyclopath, providing users with a safe place to experiment may encourage more editing and editing that is more risky, but incredibly valuable for the site.

## 6.1 Generalization

This work led us to think, in depth, about generalization. As researchers, we want to generalize results and give guidelines to practitioners, however findings from this research suggest several cautions. There are a number of factors that can affect online communities and ways communities can affect users. The specific details of a system and an online community matter. In particular, how the system fits into an ecology of related tools and communities will likely play a role in community dynamics. Another factor is how users view the tool, the community, and their role or interactions with both.

# 7. CONCLUSION

**Cyclopath core contributors aren't who you might think**. They aren't the most dedicated bicyclists (riding long distances, year round, or riding multiple times a day), but they are engaged in open content systems and believe in open content as a form of knowledge production. In addition, the Cyclopath core contributors that we interviewed were active in the local cycling community, not just as members or leaders of riding groups but also as participants in advocacy groups.

**Characteristics of the community matter**. In the case of Cyclopath, the site was launched with lots of issues on the map. (Note that this was not intentional, but a result of messy data.) Many users were motivated to begin contributing to the site because there were problems that needed to be fixed. As the site has evolved, there are fewer problems, but the tools provided by the site have matured and now allow for additional interactions and forms of contribution. In the same way, when Cyclopath started, we didn't launch the site with rules for contributors, instead those rules sprung up from the community itself. Analyzing with respect to Object, Rules, and Tools and contrasting with other existing systems, it is clear that these features and aspects of a site evolve and grow, but also help differentiate between communities.

**What is a community?** In all our work on Cyclopath, we discuss the site as a community of users. While we as researchers may believe that Cyclopath is a community, the users do not. Users also don't go to Cyclopath seeking social interaction. This could be for several reasons. First, according to the Volunteer Functions Inventory, the core contributors were not motivated by the prospect of social interaction. Second, the Twin Cities had a well-established online cycling community before Cyclopath started. These likely both contribute to Cyclopath not being seen as a community.

# 8. ACKNOWLEDGMENTS

ACKNOWLEDGEMENTS We thank the members of GroupLens Research, the Cyclopath team, and the Cyclopath user community. This work was supported in part by the National Science Foundation grant IIS 08-08692.

# 9. REFERENCES

[1] Wikipedia:activity theory - wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Activity_theory.

[2] Wikipedia:Sandbox - Wikipedia, the free encyclopedia. http://en.wikipedia.org/wiki/Wikipedia%3ASandbox.

[3] Wikipedia Editor Survey. http://commons.wikimedia.org/wiki/File:Editor_Survey_Report_-_April_2011.pdf, June 2011.

[4] J. Antin. My kind of people?: perceptions about wikipedia contributors and their motivations. In *Proc. CHI*, 2011.

[5] Y. Benkler. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press, 2006.

[6] A. Bergström and R. Magnusson. Potential of transferring car trips to bicycle during winter. *Transportation Research Part A: Policy and practice*, 37(8), 2003.

[7] S. Bryant, A. Forte, and A. Bruckman. Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia. In *Proc. GROUP*, 2005.

[8] B. W. T. Cities. Bike/Walk Pilot Profile: Minneapolis, Minnesota. http://www.railstotrails.org/resources/documents/whatwedo/policy/Pilot%20Projects_Minneapolis_final_ltr_lr.pdf.

[9] G. Clary, M. Snyder, R. Ridge, J. Copeland, A. Stukas, J. Haugen, and P. Miene. Understanding and assessing the motivations of volunteers: A functional approach. *Journal of Personality and Social Psychology*, 74(6), 1998.

[10] D. Cosley, D. Frankowski, L. Terveen, and J. Riedl. Using intelligent task routing and contribution review to help communities build artifacts of lasting value. In *Proc. CHI*, 2006.

[11] P. Fugelstad, P. Dwyer, J. F. Moses, J. Kim, C. A. Mannino, L. Terveen, and M. Snyder. What makes users rate (share, tag, edit...)?: predicting patterns of participation in online communities. In *Proc. CSCW*, 2012.

[12] E. Gilbert and K. Karahalios. Using Social Visualization to Motivate Social Production. *IEEE Multimedia*, 11(3), 2009.

[13] K. Kuutti. Activity theory as a potential framework for human-computer interaction research. In B. Nardi, editor, *Context and consciousness: Activity theory and human-computer interaction*, pages 17–44. MIT Press, 1996.

[14] M. Masli and L. Terveen. Task Specialization in Social Production Communities: The Case of Geographic Volunteer Work. In *Proc. ICWSM*, 2011.

[15] M. Masli and L. Terveen. Evaluating compliance-without-pressure techniques for increasing participation in online communities. In *Proc. CHI*, 2012.

[16] NJDOT. Introduction to Bicycle Facilities - State of New Jersey. http://www.state.nj.us/transportation/publicat/pdf/BikeComp/introtofac.pdf.

[17] O. Nov. What motivates Wikipedians? *CACM*, 50(11), 2007.

[18] K. A. Panciera, A. Halfaker, and L. Terveen. Wikipedians are born, not made: a study of power editors on Wikipedia. In *Proc. GROUP*, 2009.

[19] K. A. Panciera, M. Masli, and L. Terveen. "How should I go from ___ to ___ without getting killed?": motivation and benefits in open collaboration. In *Proc. Wikisym*, 2011.

[20] K. A. Panciera, R. Priedhorsky, T. Erickson, and L. Terveen. Lurking? cyclopaths?: a quantitative lifecycle analysis of user behavior in a geowiki. In *Proc. CHI*, 2010.

[21] J. Preece and B. Shneiderman. The reader-to-leader framework: Motivating technology-mediated social participation. *AIS Transactions on Human-Computer Interaction*, 2009.

[22] R. Priedhorsky, B. Jordan, and L. Terveen. How a personalized geowiki can help bicyclists share information more effectively. In *Proc. Wikisym*, 2007.

[23] R. Priedhorsky, S. T. K. Lam, K. A. Panciera, L. Terveen, and J. Riedl. Creating, destroying, and restoring value in wikipedia. In *Proc. GROUP*, 2007.

[24] R. Priedhorsky, M. Masli, and L. Terveen. Eliciting and focusing geographic volunteer work. In *Proc. CSCW*, 2010.

[25] R. Priedhorsky, D. Pitchford, S. Sen, and L. Terveen. Recommending routes in the context of bicycling: algorithms, evaluation, and the value of personalization. In *Proc. CSCW*, 2012.

[26] R. Priedhorsky and L. Terveen. The computational geowiki: What, why, and how. In *Proc. CSCW*, 2008.

[27] S. Reddy, K. Shilton, G. Denisov, C. Cenizal, D. Estrin, and M. Srivastava. Biketastic: Sensing and Mapping for Better Biking. In *Proc. CHI*, 2010.

[28] Y. Ren, F. M. Haper, S. Drenner, L. Terveen, S. Kiesler, J. Riedl, and R. E. Kraut. Building Member Attachment in Online Communities: Applying Theories of Group Identity and Interpersonal Bonds. *Management Information Systems Quarterly*, 36(3), 2012.

[29] R. Rosenthal and R. L. Rosnow. *Essentials of behavioral research: Methods and data analysis*. McGraw-Hill, 1991.

[30] S. A. Sheppard and L. Terveen. Quality is a Verb: The Operationalization of Data Quality in a Citizen Science Community. In *Proc. Wikisym*, 2011.

[31] F. Torre, S. A. Sheppard, and L. Terveen. bumpy, caution with merging: an exploration of tagging in a geowiki. In *Proc. GROUP*, 2010.

[32] A. Velasquez, R. Wash, C. Lampe, and T. Bjornrud. Latent users in an online user-generated content community. *Computer Supported Cooperative Work (CSCW)*, 2013.