Virtual Community Maintenance with a Collaborative Repository

Derek L. Hansen (corresponding author)
School of Information, University of Michigan, 1075 Beal Ave., Ann Arbor, MI 48109-2112
shakmatt@gmail.com

Mark S. Ackerman
School of Information and Department of EECS, University of Michigan, 1075 Beal Ave., Ann Arbor, MI 48109-2112, ackerm@umich.edu

Paul J. Resnick
School of Information, University of Michigan, 1075 Beal Ave., Ann Arbor, MI 48109-2112
presnick@umich.edu

Sean Munson
School of Information, University of Michigan, 1075 Beal Ave., Ann Arbor, MI 48109-2112
samunson@umich.edu

Virtual communities, like all communities, require ongoing community maintenance activities. This paper presents an empirical study examining how a wiki repository was used to help overcome some of the community maintenance challenges common to help-based email list discussions. Specifically, we found that inclusion of off-topic but related content on the wiki enabled list members to keep the discussion on-topic while still addressing the needs of members. Offloading of repetitive and potentially contentious "holy war" debates to the wiki encouraged list members to summarize their arguments into a meaningful information product. The community's use of the wiki in helping answer frequently asked questions helped attract new members and helped them gain the knowledge they needed to comfortably contribute to the email list. It also helped active participants answer questions more efficiently and effectively by supporting the reuse of information. Finally, the wiki supported peripheral participation by new and former members. This study demonstrates that the architecture of information collections and information flows in an online community has a significant impact on the social processes related to community maintenance.

Introduction

Increasingly, individuals are seeking answers to questions from help-based online communities. To be successful sites for information sharing, such communities require maintenance. New members must be assimilated. Experts must be retained. The conversation must stay on track and consistent with the goals of the community. In an effort to ameliorate some of these community maintenance challenges, many online communities have used FAQs and web pages to supplement their ongoing conversations. However, it is not yet well understood how ephemeral conversations and more stable documents can be harnessed in the maintenance of online communities.

This paper presents the results of an empirical study examining how a community repository (implemented with a wiki) was used to help overcome common social challenges of help-based online communities. Much previous literature has examined the social processes that enable maintenance of a community repository (Ebersbach et al., 2006; Leuf & Cunningham, 2001; Marshall et al., 1994). We focus on the other direction of influence, how a community repository can help a community maintain its social order. The virtual community we examine, css-d, is a large technical support community that relies on voluntary participation and is primarily concerned with helping individuals satisfy their information needs through group discussion. The community is notable because it has augmented threaded conversation with a collaborative authoring tool (i.e., wiki) that is used as a shared knowledge repository. This paper describes how list members and administrators have used the wiki to help overcome some of the challenges common to email list discussions such as staying on-topic, avoiding unproductive repetition of common disputes, helping newcomers join the ongoing conversation, retaining old-timers, and promoting the community to the outside world. Our findings shed light on the

information issues involved in community maintenance and have practical implications for online community administrators, members, and social software designers.

Literature Review

Information scientists have long taken an interest in the role of information resources and their associated community networks in strengthening local communities (e.g., Joan C. Durrance, 1984; Schuler, 1996). Indeed, one of the major functions of (and justifications for) public libraries is their positive impact on communities. Researchers have recently drawn attention to the synergies that exist between information resources (i.e., libraries and librarians) and local communities grappling with social challenges (e.g.,J.C. Durrance et al., 2006). This line of research has led to practical suggestions for strengthening communities through the use of information resources, technologies, and spaces (e.g., Joan C. Durrance & Pettigrew, 2002). Although this work includes discussions of information technology and its role in supporting local communities, it has not yet explored the unique social challenges that arise within virtual communities. This new territory is becoming increasingly important to understand as virtual collaborators have become more involved in the creation and use of information resources.

Benkler coined the phrase *commons-based peer production* to describe processes whereby large groups of individuals voluntarily produce shared information goods and services through the use of technology (2002). It is contrasted with more traditional forms of production including firms, organization, and markets (Benkler, 2002). Numerous successful examples of peer-produced projects (e.g., Wikipedia, Apache, Slashdot, Project Gutenberg) attest to the viability of this new form of production. However, the countless failed projects attest to the difficulty of supporting distributed peer production and the need to better understand the social and technical factors that lead to its success. Ackerman (2000) characterizes this difficulty as a social–technical gap, which is “the divide between what we know we must support socially and what we can support technically”. Both Ackerman (2000) and Benkler (2002) call for research that will provide insights into the social and technical mechanisms that support distributed collaboration. Such research will help us “tap substantially underutilized reserves of human creative effort” (Benkler, 2002).

Commons based peer production often occurs in online communities. One type of online community of particular interest to information scientists is what Fisher et al. (2006) call question and answer communities. In such a community, the primary tasks are understanding and satisfying individuals’ information needs, tasks familiar to information professionals of every shape and size. Yet, the distributed, voluntary, computer-mediated environment of online communities introduces new social challenges and old challenges in new guise. For example, the combination of anonymity and lack of social cues due to text-only conversation has led to an increase in hostile messages, or flames (Dery, 1993). Related “hostile behaviors” include spamming, trolling, and even cyber-rape (Burnett & Buerkle, 2004). Other social challenges include motivating participation, keeping the discussion on-topic (Kollock & Smith, 1996), and avoiding offensive behaviors such as over-quoting of prior messages, asking frequently asked questions, and misusing subject headers (McLaughlin, 1995).

Several experienced online community administrators and authors have recognized the need for strategies and technologies that help address these social challenges within online support communities. For example, Preece (2000) argues that the need to design for “sociability” is at least as important as the need to design for “usability”. Other popular books on building online communities focus on fostering appropriate social interactions, rather than the details of the technologies (Kim, 1999; Powazek, 2002). Likewise, the Listserv® manual provides detailed suggestions for email list owners (i.e., administrators) on how to deal with flame wars, promote proper “netiquette”, and welcome newcomers (L-Soft international Inc, 2003). Butler, et al. (2002) used a survey methodology to characterize the current work and motivations of several email list administrators and members. They found that list administrators play a unique role within communities – one that encompasses both technical and social responsibilities. They also found that other members shared in the community maintenance activities through participation, recruitment of new members, and even managing social dynamics. Although this literature paints a fairly detailed picture of current practices, Preece (2004) argues that there is a need for novel social and technical approaches that help foster etiquette in online communities.

In this paper, we examine a community with a novel approach to overcoming the aforementioned social challenges – one that is both practical and that highlights the information flow issues involved in community maintenance.

Site, Data Collection and Methods

This work is part of a larger empirical study of the css-d community that has taken place over a two-year period. We have chosen to study a single site in depth for several reasons. Most importantly, the subject of inquiry is an entire system of action, not the typical behaviors of an individual or even the aggregate behaviors of a group of individuals (Feagin et al., 1991). The ultimate goal of this research is to provide insights that are useful for community designers. A case study design such as ours draws the boundaries of inquiry precisely around the thing (i.e., the system of action) that community designers can influence. It is also appropriate for describing new phenomena (such as the one described in this paper) where current perspectives have little empirical substantiation. As well, in depth empirical examinations such as this one are amenable to a holistic, grounded approach that can lead to insights that are useful to community designers (Button & Dourish, 1996; Hughes et al., 1994).

The particular site, css-d, was chosen primarily because of its successful integration of a community repository with an email list conversation. The continued use of the repository by community and non-community members was the primary evidence for success. Its success was verified by the positive perceptions reported by many community members throughout the project. Although the community has succeeded at creating and maintaining a useful repository on the whole, the result is by no means perfect. The study is intended to learn not only from the successes, but also from the failures of specific practices related to the repository and threaded conversation, although this paper primarily draws attention to its successes in the area of community maintenance. The site was also chosen because it is typical of online help-based communities with its reliance on traditional threaded conversation (via an email list), wide variation in members' levels of expertise, and focus on asking and answering practical questions related to a specific area of interest.

Our dataset includes email list messages, wiki content, and wiki server logs. Over 90,000 email messages were available from the list's inception in January 2002 to the present time (January 2007). Email messages were analyzed from time periods before the wiki was in existence, during its initial implementation, and after it was well established. Messages that reference the wiki or email archives were oversampled, as well as messages that included "ADMIN" in the subject line. Wiki content included several snapshots taken between April 2003 and May 2006 and recorded in the Internet Archive (n.d.) or by the authors. The historical snapshot approach was necessary because the wiki page histories are only stored for two-week intervals. All wiki pages available in May 2006 were read, as well as prior and later versions of more popular pages. We were provided access to several server log reports between May 2004 and January 2007 by Incutio, the company that hosts the wiki. The reports were generated by Advanced Web Statistics 6.5 software package (2006) and included statistics such as the number of unique visitors (excluding robots and spiders), total hits, visit duration, prior origin of visitor (e.g., search engine, other website, direct access), and search terms that led visitors to the site.

We have adopted an approach in the spirit of grounded theory, where the major themes and findings emerge from the data through an iterative process (Glaser & Strauss, 1967). Thus, we have tried to begin "as close as possible to the ideal of no theory under consideration and no hypothesis to test" (Eisenhardt, 1989). Taking this approach allowed us to explore issues most salient to the community we studied and not our own preconceived notions. Indeed, the role of the wiki in social maintenance was not one that we had initially intended to explore. Although we have tried to let the data drive our inquiry, analysis, and findings, it has clearly been filtered through our own theoretical stances and experiences to some extent. Our underlying assumptions are best described in Orlikowski's (2000) application of Gidden's Structuration Theory (1986) to the use of technology in practice. In this approach, the research focus shifts "from a focus on given technologies, embodied structures, and their influence on use – to a focus on human agency and the enactment of emergent structures in the recurrent use of technologies" (Orlikowski, 2000).

We rely on both quantitative and qualitative data, allowing for triangulation. We used an iterative process of exploration. This began with quantitative analysis of participation patterns and qualitative exploration of content. Interviews were then conducted with previously identified members in order to check our understanding and inform future analyses. This led to new quantitative and qualitative analyses of the data, the emphasis always being on understanding the community from an insider's perspective. This cycle continued until we believed the major themes were well enough understood and the major claims supported. A draft version of this paper was presented to the community for comment. Three individuals, including one of the active administrators, responded with comments. They felt that we accurately represented the community and suggested a few changes that were incorporated, such as a greater emphasis on the value of the wiki in retaining members.

Semi-structured interviews were conducted after the bulk of the content analysis was performed so that questions about the interviewees' specific actions could be addressed (e.g., why did you reference the OffTopic wiki page?). Over a 12-month period, we conducted a total of 14 one-hour phone interviews and 7 email interviews. Members of particular interest (as identified by prior analyses) were the four list administrators, 8 active list and wiki contributors, 4 participants that used the email list primarily (or exclusively), 3 that used the wiki primarily (or exclusively), and 4 that rarely contributed to either. Interview questions focused on community roles, activities, perceptions of the wiki and email list, the nature of CSS work, important information resources, motivations for certain observed behaviors, and the social dynamics of the community.

The quantitative analysis relied upon the email list corpus, wiki content, and website logs. Email messages were analyzed to determine who sent messages, who referenced the wiki (and how often), what wiki pages were referenced, and how this changed over time. Wiki content was analyzed to determine the number of pages, their growth over time, their size (in bytes), their content features (e.g., number of internal and external links, headers, css code), and the number of times they were edited. Website logs were used to better understand the usage of the wiki and how members arrive at the site. Finally, all messages sent between Jan 2003 and April 2005 that reference the wiki (1,897) or email archive (632) were independently coded by two raters in order to determine the reasons for the reference. Both raters coded the same 500 messages that referenced the wiki (26%) and 200 messages that reference the email archives (32%). This assured that the coding scheme was sufficiently fleshed out and consistently applied. Cohen's kappa values (k) are reported throughout the paper as a measure of inter-rater reliability.

The qualitative analysis included content analysis of email messages, wiki pages, and transcribed interviews. A grounded theory approach was used to identify types of community maintenance work occurring within the community email list and wiki and identify common practices related to that work. Example messages were grouped together, and tentative hypotheses and categories were developed. For example, we suspected that the wiki was used more often than the email archives to shut down off-topic discussions. This initial hypothesis was tested through formal coding of messages from the wiki and archive and better understood through interviews and content analysis of those messages. Early interviews were conducted after this first pass through the data in order to shed light on the preliminary facts and hypotheses, which were abandoned, refined, or strengthened throughout the rest of the data process. They were transcribed on an ongoing basis and common themes addressed by different interviewees were grouped together in order to analyze them side by side.

Community member names and personally identifiable information has been made anonymous throughout this paper. In some cases messages have been slightly edited for presentation purposes. Comments in square brackets are ours. Next we describe the css-d community.

CSS-D

Email List

The css-d community began in January 2002 as a public mailing list devoted to discussions about the applied use of Cascading Style Sheets (CSS), a technology used to add style (e.g., fonts and spacing)

to structured web documents (e.g., HTML or XML documents) (Bos et al., 2006). CSS is intended to simplify the creation and maintenance of websites by separating the styling aspects (controlled by CSS) from the content. The css-d email list has over 7,000 subscribers (as of May 2006). They include both professional and amateur website developers. Membership is open; anyone can join through an automated process that does not require prior approval. The list is unmoderated, meaning that messages sent by members are immediately redistributed to all members, without anyone reviewing and approving them. Members send an average of over 50 messages a day, a number that has remained relatively stable since the list's inception, although the number of distinct posters per month has steadily increased. Many other messages are sent directly to individual authors, especially since the list is configured so that by default replies are only sent to the prior author (i.e., the sender has to deliberately choose to send a reply to the whole list). All email list messages are made available through a public, searchable email archive. Interviewees agree that the list's initial popularity (over 1,000 members in the first 2 days) was a result of the notoriety of the list founder and other regular participants, as well as its unique focus on the practical use of CSS.

Like other help-based communities (Lakhani & von Hippel, 2003; Sproull & Patterson, 2004), most email messages follow a question and answer format, are relatively short, and are grouped into threads (avg. 3.3 messages per thread). Participation in the list is highly skewed, with 4% of all contributors accounting for 50% of all messages. Most members are lurkers. Of those who post, one third only start threads, one sixth only answer them, and the remaining half do both. Significantly, many posters are long-time participants. There were 474 members who posted at least one message a year after their first post. These long-term participants include 45 of the 50 most prolific posters, as well as members all along the participation spectrum (e.g., median number of posts for long-time participants is 19). However, there is a steady stream of newcomers to the list, where membership has steadily increased over time. Content analysis and interviews revealed that members participate to help satisfy an immediate information need, to receive feedback on their websites, to learn from others' mistakes, to keep up on recent developments, and to "give back" to the community.

The primary purpose of css-d is to meet the immediate needs of members, as they relate to the topic at hand. These include the need to find an appropriate webpage layout, achieve a particular effect (e.g., create rounded corners), diagnose an unexpected problem (e.g., extra spacing around an image), find a "hack" or workaround to deal with non-CSS compliant browsers, or receive feedback on a website. Although the CSS specifications are relatively straightforward, in practice the use of CSS can be extremely complex. This complexity derives largely from bugs in web browser software and inconsistencies in the way they interpret the W3C CSS specifications, making it hard to create a page that looks the same in every browser. Problems can be difficult to diagnose and there are often numerous possible solutions to a problem, each with its own side effects and complex sets of contingencies.

Wiki

In the Fall of 2002, more than six months after the email list was created, list administrators allowed a member to set up a web-based collaborative authoring tool based upon the WikkiTikkiTavi wiki engine (n.d.), hereafter referred to as the wiki. Any visitor to the css-d wiki can add or edit content using a simplified markup language that supports structured content (e.g., tables), formatting (e.g., bold text), and hyperlinks. Initial content, in the form of a network of hyperlinked wiki pages, was provided by a handful of members with varying levels of expertise. Wiki content has continued to steadily grow over time, with an average of 2 new pages a week and an average total increase of about 65 words per day. In May 2006 there were over 100,000 words spread across 544 pages. The Wiki has been described by participants as a "shared repository," a "user-defined FAQ," the "collective wisdom of this list," and a "dumping ground for [CSS] tricks." It includes mostly CSS content pages, with some personal biography pages and administrative pages describing the email list policies and how to use the wiki. About 4% of email list messages, or roughly 2 per day, include a URL linking to the wiki. Thus, for most members the wiki serves a subsidiary role to the email list, which is the backbone of the community. However, a

handful of the most active wiki contributors rarely post to the email list and a few wiki contributors aren't even registered on the list.

Administration

Like many technical support groups, css-d includes a "list chaperone" and a handful of other volunteer administrators (i.e., admins) who help keep the discussion running smoothly. The admins are all long-time, active community members who are generally well respected CSS experts as evidenced by the many times members refer to their work in the wiki and list. In addition to managing email list software (e.g., keeping the membership list current), they deal with the typical social maintenance issues that come up on unmoderated email discussions. Admins post public messages at least weekly to remind members of list policies (e.g., trimming quoted material) and to keep the discussion friendly and on-topic. These messages often include the word "ADMINS:" in the subject line and are typically sent in reply to the offending message(s). They often include an appeal to the list rules and policies (which are found on a static community website and on the wiki), state the authority of the administrator, and sometimes offer arguments about the rationale for the policy. Admins also send messages in private, especially for more minor or private offenses (e.g., asking a new question in an existing thread). Although the threat of removing individuals from the email list is mentioned regularly, only the most egregious of sins (e.g., posting job solicitations or automatic replies) actually prompts action. This act is tempered by the fact that members can re-subscribe and in many cases do (presumably having learned their lesson).

Challenges to the admins' authority are rare and are encouraged to occur offline. Several members have complained in public messages posted to the list and personal communication with the authors that the admins are too authoritarian, particularly in regards to keeping the discussion on-topic. However, others have defended the admins' approach in public messages and interviews. Many other members of the community help discourage misbehavior and enforce list policies through snide remarks, gentle reminders, and on occasion use of the silent treatment (i.e., ignoring a message that blatantly violates community norms).

Nearly all interviewees described the general tone of the css-d email list as friendly and professional, in contrast to other website development email lists that they participated in. The community leaders are adamant about welcoming newcomers and assuring them that there are "no stupid questions." As the list chaperone has stated in the policies: "css-d is meant for beginning and experienced authors both, but I'm actually more interested in helping out the beginners." A few long-time members voiced frustration with this approach in interviews, while others defended it. Quantitative analysis reveals that many experts have stuck around and the number of novices has grown substantially, suggesting that the current approach is having the intended effect.

Community Maintenance with the Aid of a Wiki Repository

This section describes how the work of community maintenance is performed at css-d, highlighting the role of the wiki repository in this work. Many factors have led to the success of css-d. Interviews suggest that the notoriety of the list founder and other active members, its friendly and welcoming tone, and its unique focus on the practical issues surrounding CSS are among the most important. In this paper we focus attention on the role of the wiki in supporting these factors. Namely, we examine how the community uses the wiki to help keep the discussion on-topic (e.g., practical) and friendly, attract and retain new members, and keep old-timers around. This focus is justified for two reasons. First, it draws our attention to the ways in which information environments can influence, and be influenced by, social dynamics. Second, all of the administrators mentioned in interviews that the wiki is very important to the work of community maintenance.

Managing Social Dynamics

Although there are many issues related to the management of social dynamics, we focus our attention on the two most common ones at css-d: keeping the conversation on-topic and avoiding "holy wars"

that repeat past disagreements on controversial topics. The wiki has been used by list administrators (and other core members) to help manage these social challenges occurring on the email list. Coding revealed 76 cases where the wiki was used to help offload the discussion because it was off-topic or considered a “holy war” (k = 0.96). Interviews suggested that these are only a small fraction of the total times the wiki was referenced to help offload the discussion, since the majority of times these messages were sent directly to individuals rather than to the list as a whole. In contrast, only 3 messages during the same 28 month time period were identified that referenced the email list archives for a similar reason. The following discussion helps explain why the wiki has been particularly useful in managing these issues.

Staying On-topic

One of the most significant challenges faced by many help-based communities is the need to stay on-topic. If the community conversation drifts too far from its stated focus, members who came to learn about the topic (e.g., CSS) can easily become frustrated by the need to filter out off-topic messages. Kollock & Smith (1996) argue that staying on-topic is critical to the coordination of community knowledge sharing, but unfortunately, is challenged by the free-rider problem. In other words, while the collective good is best served by people staying on-topic, there is an individual temptation to post off-topic messages in order to reach more people. Other more innocent posters may simply not understand what constitutes an on- or off-topic post. For example, at css-d many well-intentioned individuals post off-topic messages because it is not obvious to them what constitutes a practical CSS question (which is on-topic) from a theoretical one (which is not). In addition, there are many gray areas at the intersection of CSS and related technologies (e.g., using a scripting language to serve different CSS pages to different browsers). Finally, some people post off-topic messages because they have formed friendships with community members and want to share meaningful experiences with them, even if they aren’t directly related to the topic at hand. This can be helpful in strengthening ties among those members, but can frustrate everyone else. The challenge is to negotiate the boundary of what is on- and off-topic so that both the posters and those overhearing the conversation are satisfied.

The css-d community has used the wiki to help successfully negotiate this boundary by providing a new method for dealing with off-topic content. To understand how this plays out, we present a typical off-topic example that began when a “newbie” asked for recommendations of Content Management Systems (CMS software) that were CSS compliant. Within a few hours he had received 7 replies from community members, which suggests that this topic was of interest to several members. It was, however, technically not on-topic. The final message in the thread was sent by a list admin who shut down the thread as shown below:

> **Email 9 of 9 in Thread (3.5 hours after original post)**
> **Subject: CMS Question**
>
> Unfortunately, I'm going to have to call this thread to a close for being off-topic. Yes, it's true that choosing a standards-friendly CMS has bearing on CSS authoring... but so does writing standards-friendly PHP code, JavaScript, .Net code, etc. etc. If we walk down that road, traffic could quite easily double. We can't cover everything standards-related here; there simply isn't room.
>
> A great place to ask your question would be Webdesign-L. There's also evolt's thelist, and I'm sure there must be other venues where CMS discussions are on topic. For a short list of forums, see [URL for OffTopic page]. On the topic of CMS software, see the page others in the thread have brought up: [URL for CssFriendlyCms page].
>
> Thanks, and sorry to have to end the thread.

There are several things to note from this example. First, the admin is more courteous than he is in some other comparable posts, most likely because the offender introduced himself as a “newbie.” This

is evidenced by the apology for ending the thread, the pointers on where to go, and the explanation justifying the list policy. Second, the admin references the wiki's OffTopic page, which is an annotated list of links to related discussion groups. It also includes a brief description of what is on- and off-topic at css-d. Finally, the admin links to another wiki page (CssFriendlyCms) that discusses the questioner's very topic and had already been pointed out by other list members. The page includes details on how well various CMS products interact with CSS.

In this example (and others like it), the wiki complements the email list discussion. While the email list is a push technology that is broadcast to all members of the list, the wiki is a pull technology that requires members to actively seek it out. Members at css-d have used this combination to their advantage. Wiki content is allowed (and even encouraged) to cover topics that are "on the margins of on-topicness" but are technically off-topic (e.g., CssFriendlyCms, Javascript Hacks). These pages do not force themselves into anyone's email box; hence, they bother people less. However, members that do care about the topic can use the less invasive wiki technology to educate one another and collaborate if necessary. Likewise, when newcomers post off-topic messages, it is possible to send them to a useful, community-created page rather than blowing them off altogether or taking the time to answer their question in a private message.

The existence of the OffTopic wiki page is also helpful in reducing the admins' workload, while being friendly to newcomers. One admin put it this way: "if you just tell them that [a post is] off-topic, they will often email back and say, 'well, where else can I learn about this?' And so then you gotta point them to Webdesign-L's list or Evolt's thelist or a bunch of other common mailing lists, so instead it's just all there. So we don't have to type the same email out to them several times." The OffTopic page also helps non-admins keep the discussion on-topic by lending credibility. In fact, one-third of the references in emails to the OffTopic page were posted by non-admins.

Although the examples discussed so far occur after-the-fact, the wiki has been used to preemptively offload discussion on certain hot topics. For example, in order to reduce some of the list traffic about the Internet Explorer 7 beta release, one admin created an IE7 wiki page and encouraged members to use it as a place to collect and share testing results. One of the primary reasons this approach was taken was to "avoid having the list drowned by IE7 testing results and related traffic."

Avoiding Holy Wars

Another common social challenge in online technical support communities is the outbreak of holy wars, or un-resolvable debates with little practical value. The term holy war, as applied to threaded discussions is not unique to css-d. The Netlingo Internet Dictionary defines it as perpetual discussions "that never die, the arguments never change, and no one's opinions ever budge one iota" (n.d.). Administrators try to avoid holy wars because, like off-topic discussion, holy wars are generally only of interest to a few, but demand the attention of many. Furthermore, when they lead to flaming, relationships can be damaged, members may become more hesitant to ask questions of their own, and members that are sensitive to conflict may even leave the community. Early cases of people leaving the list after holy wars broke out prompted the administrators to take a hard stance against them as evidenced by their revision of the policies and messages sent to the list to shut down (or prevent) holy wars from breaking out. Their view is captured by one administrator's comments: "Such discussions will at best clog up the list with fruitless back-and-forth, and at worst ignite a massive flame war. Neither is acceptable." When the wiki was created, the administrators immediately recognized it as a tool that could be used to help address this problem, as explained below.

The following quotes are taken from two messages (out of 17 in the entire thread) where debate broke out over the appropriate use of font size on a webpage.

> **Email 1 of 17 in Thread**
> **Subject: Comments Requested on website.com**

I have recently launched the *beta* version of my new blog at website.com. There are no graphics yet, but the layout is in place. I'd like to get some public feedback on the display. Do you notice any browser-specific quirks? Is the text too small?

David Meyers
ABC Inc.
Kalamazoo, MI

Email 16 of 17 in Thread (23 hours later)

If there is one thing this list will never, ever make possible, it's a final resolution to the Font Sizing Holy War. This is why we established a Wiki page that gives people a good jumping off point for reading up on various perspectives: [URL for FontSize].

The next time the war breaks out, let's just refer each other to that page and move on to discussions of a less religious nature. If there are non font sizing comments to make about website.com, let's get back to making those— otherwise, I think it's time to move on. Thank you.

--

John Jones [personal URL], List Administrator
[CSS related tagline]

The thread begins when David asks the community for feedback on the "beta" version of a new website he's developed. This type of request (known as a "Site Check" by community members) is common at css-d. Because the same CSS code is rendered differently by each browser, Site Checks often include requests for members to view the site in a browser not immediately accessible to the poster. This explains David's question about "browser-specific quirks."

Many replies to Site Checks are not sent to the entire list because the suggestions are not on-topic (e.g., comments on the photography, not the CSS code) or of general interest. However, on occasion Site Checks will initiate a debate about best practices, as occurred in this example. David's request for feedback on the size of his text served as fodder for a holy war. In CSS there are a number of ways to specify the font size including using pixels, percentages, ems, keywords, and points. Each method has different implications for different browsers and user groups. In this case, the discussion degraded into a fruitless debate about whether or not users should be expected to change the browser's default text size settings. As often happens in these cases the debate was also becoming personal as evidenced by members use of strong language, numerous exclamation points and all-caps statements for emphasis, and comments directed at one another rather than the issue at hand. A list administrator ended the holy war by pointing list members to a wiki page, in this case the FontSize page. That page and related pages describe how to use the various font sizing techniques and summarize their potential benefits and drawbacks.

Diverting holy wars by pointing to an FAQ on the topic is not new. However, the differences in who can edit the page and the format of the page are significant. Because the wiki is editable by those who are being diverted, they have the option of contributing their knowledge rather than being completely shut down. However, their contributions are not forced upon anyone since the wiki is a pull technology. As stated on a particularly contentious wiki page, "Since HolyWars are by their nature long winded and boring let's keep this one tucked away on its own page where it won't bother anyone who doesn't want to be bothered."

The format of the wiki and the shared understanding of its purpose also influenced how things play out at css-d. One list administrator described his approach this way, "What I stressed was to present [the contentious topic] with as little heat as possible on the wiki. If there were two sides to an issue they should be presented fairly and sort of neutrally. Document the pros and cons. And if somebody couldn't

think of any cons for their preferred method, they could write up the pros and someone else could think up the cons and can add them." Several pages have done just that. For example, UsingFontSize lists a distilled version of the pros and cons of the various font sizing techniques in a comparison table, many of which were contributed by members that regularly contribute to holy wars. The structured format and established neutral tone of the wiki has encouraged the useful distillation of even the most problematic discussions. Indeed, a previous holy war participant, when interviewed, stated that she did not feel put off by this approach because she felt like she could still voice her opinion on the wiki. The result is that holy war participants are encouraged to turn their "weapons into plowshares," by converting their arguments into a useful information product instead of using it to endlessly debate or belittle others.

Attracting and Retaining Members

Unlike physical communities, the cost of joining and leaving online communities is relatively low. In order for communities to be sustainable, they must retain a healthy percentage of members while constantly adding new members to the mix (Butler, 2001). Although the ongoing sustainability and growth of the community is attributable to many factors, the css-d wiki has significantly impacted the ways in which new members enter the community, as well as the incentives for old-timers to stick around. Furthermore, it has provided a new form of peripheral participation for new and former members (Lave & Wenger, 1991).

Attracting New Members

The css-d community has several possible online entry points including the wiki, a static community website, the public email archives, and the founder's website. Google searches for "css discuss" and "css discussion list" made on January 27, 2007 brought up the wiki's home page as the first result, suggesting its importance in relation to other pages. While each of these, along with word of mouth, acts as a public face to the outside world, the wiki plays a significant and unique role in the promotion of css-d.

Many members are initially introduced to the community through the wiki. In January 2007, there were approximately 10 times more visitors to the wiki than members subscribed to the list. Most connections to the wiki site are from search engines (44%) or websites not affiliated with css-d (18%), avenues likely to attract individuals new to the community. In comparison, 33.7% connections were direct access (i.e., from bookmarks or following links in email clients) and 4% were from other css-d websites such as the archives, avenues likely to be used by existing members. Analysis of the search terms suggests that newcomers find the wiki by using search phrases on specific CSS related topics like "three column layout," "css hacks," and "css font size" and are taken to the specific wiki pages on these sub-topics.

The technical structure of the wiki and its use by the community have contributed to its ability to reach new members. The persistence of wiki page URLs allows individuals to link to a page with confidence that it will be there in the future. Although it is technically possible to change the URL of a wiki page, it has rarely been done. Because content on a wiki page can be updated, it is more likely that pages will stay up-to-date than links to unchangeable, archived email messages. Thus, the combination of persistent URLs and changeable content makes it more likely that individuals will link to specific wiki pages. This increases the chance that others will stumble upon the wiki (and thus the community) and increases the likelihood that individual wiki pages will show up high in search engine results (i.e., it increases the wiki's Page Rank) (Brin & Page, 1998). Furthermore, this process makes it more likely that newcomers are exposed to the most popular (and presumably most useful) wiki pages first. For example, the IE7 page was created to document how well the new version of Internet Explorer supports CSS. Dozens of community experts posted CSS test results and bugs to the page. Because of its usefulness and timeliness, many members linked to the IE7 page from elsewhere on the web. The result was that many individuals (including the Microsoft IE 7 development team) were introduced to the community through the page which showcased the expertise of its members.

Not only does the wiki attract people to the community, it also helps newcomers integrate into the community. Because the wiki content is heavily influenced by the email list discussion, it represents the core values, knowledge, and interests of the community. Potential members can browse through the wiki pages indexed on the front page and get a quick overview of topics most salient to the community. Although exploring the public email list archive can also provide this information, its lack of organization, repetition, and length (of over 85,000 messages) make it hard to quickly digest. The wiki also includes pages describing list policies (e.g., PostingGuidelines), as well as helpful suggestions for newcomers (e.g., a page explaining how to reduce excessive quoting when using Gmail).

New members can also benefit from the CSS content pages that describe some of the basics related to a given topic. Like an FAQ document, many of the wiki pages (e.g., BoxModelHack, RoundedCorners, CenteringBlockElement, FontSize) are written primarily by experts so that they can use them to answer novice questions in the future. The result is that the wiki includes topics that are suggested by novices, along with summary write-ups that are intended for novices – an important source of new members. Content analysis and interviews revealed that many members read these pages and the email list archives prior to posting messages. Around 14% of all messages that referenced the wiki in a 28 month period did so when asking a question ($k = 0.94$). Analysis of these 262 messages suggests that members use the wiki to help know what to ask and how to ask questions, refine their questions, ask questions more efficiently (by referencing a particular technique described on the wiki), and justify that they have done their homework (and thus deserver an answer). During the same time period only 60 messages referenced the email archive ($k = 0.82$). These messages were used for similar reasons as those that reference the wiki. However, they differ in that individuals that referenced the wiki were more likely to mention specific techniques that were borrowed from the resource and to link to specific wiki pages. Individuals that referenced the email archive were more likely to simply justify their question by stating that they had searched the archives without finding the answer. This analysis suggests that the wiki is more frequently used to help ask questions and more useful when questioners want to justify or explain the use of a particular technique.

One newcomer described how he regularly searched the wiki before posting in order to “get more know-how to ask a question,” especially since it was “going to a lot of people – CSS experts.”

> So I don’t want to sound like I don’t know what I’m talking about or I’m a stupid newbie or something like that. So, yah, I try to find the answer myself first, if I can, then at least I try to be fairly knowledgeable on how to ask the question without looking too stupid.

This ability to become knowledgeable before acting enables peripheral participation (Lave & Wenger, 1991), which is an important factor in helping people move from a peripheral to a central role. Likewise, members can use a technique from the wiki (and post questions about it) with confidence that it is an accepted technique recognized by the community. Although this can be done with techniques from the email archive, those on the wiki have increased credibility because of its selective nature.

Retaining Long-Time Members

Many online communities have difficulty retaining members for long periods of time. While some of this is due to people changing jobs or losing interest in the subject-matter, other times it is caused by the tension between supporting both newcomers and old-timers. Newcomers often have “newbie” questions that wear on the patience of old-timers. Conversely, old-timers often want to discuss more advanced topics or continue conversations without having to bring everyone up to speed. At css-d, where the primary focus has always been on supporting newcomers, there is a stronger possibility that old-timers will not stick around. Although some core members no longer actively participate in the list, there are many who have stuck around for extended periods of time. The wiki and the social practices around its use have played an important role in retaining these old-timers while still meeting the needs of the newcomers as described below.

As in other technical support communities, newcomers are encouraged to look for an answer to their question in the wiki and email list archives before posting to the list. Several long-time members mentioned that they believe the wiki has led to fewer basic, repeat questions since many newcomers go there before posting to the list (as discussed in the prior section). For example, one interviewee said that the wiki was useful in “encouraging long-term members like me to stick around, since we can move FAQs to the wiki and don’t have to be bored or frustrated by those threads as often.” Although it is hard to definitively prove that fewer frequently asked questions arise on the list, in this case the members’ perception is more important than the reality because it is their perception that influences them to stay or leave.

The wiki has also helped efficiently answer frequently asked questions when they do come up on the list. At css-d the encouragement to search the wiki and archives before posting to the list is tempered by the strong emphasis on welcoming newcomers who are told in the list policies that the group is “usually pretty tolerant of repeated topics as long as they're spread out over time.” Unlike many technical support lists where members are consistently told to RTFM (Read The Flipping Manual – polite form), css-d rarely explicitly tells newcomers to search the archives or wiki. Of the 331 messages that referred to the email archives only 9 (2.7%) explicitly stated that they should check the archives before posting ($k = 0.78$). Likewise, only 20 of the 1,308 messages (1.5%) that referenced the wiki explicitly stated that the wiki should have been consulted first ($k = 1$). In one telling example an administrator severely reprimanded another core member who scolded a newcomer for posting a basic question. Instead of chiding newcomers for not having looked in these resources, the typical response has been to follow the suggested list policy:

> Simply posting a URL as an "answer" is also discouraged. Back up that URL with a little explanation of what the reference is about, why you posted it, and some keys to understanding the resource you're referencing. It doesn't have to be a novel; a line or two will usually suffice. But that line or two will be of enormous help to people reading your message, who may not be as expert as you are.

While this approach is helpful to newcomers, it also places an extra burden on old-timers. This burden has been largely reduced through the use of the wiki. As described before, many experienced members have created pages on frequently asked topics. When these topics arise in the list, members can easily answer them by linking to the wiki along with a couple sentences customizing the message to the member’s particular needs. In fact, the primary reason for referencing the wiki (60% of all wiki references; $k = 0.96$) is to help answer questions and educate others. The email archive is used for a similar purpose, but is only referenced $1/6^{th}$ as often as the wiki (165 times compared to 973 in the same time period). This suggests that the wiki more effectively supports the reuse of information than the archive. The specific reasons for this are not within the scope of this paper, but are being examined by the authors.

In addition to dealing with FAQs, the wiki has also helped some active members who cannot spend the time reading all of the email messages to stay at least partially connected to the community. Two members mentioned in interviews that they continue to review and edit wiki pages even though they don’t have time to subscribe to the email list any longer. By providing a new form of lower cost participation, the wiki has enabled experts to continue to share their knowledge through the wiki even if they don’t through the list.

Conclusions

Virtual communities, like all communities, require ongoing community maintenance activities. Paramount among these are the need to maintain social norms around interaction and the need to attract and retain members. This paper reports on a specific online community, css-d. As a singe site study, it has the standard generalization problems. However, the study shows the potential of a repository for this site to help important social maintenance activities.

Other communities may differ, but all must address these social maintenance needs. We found that the css-d community was able to use the wiki repository to help meet these needs in several specific ways:

- Inclusion of off-topic, but related content on the wiki enabled list members to keep the discussion on-topic while still addressing the needs of members.
- Offloading of repetitive and often contentious “holy war” debates to the wiki encouraged list members to summarize their arguments into a meaningful information product.
- Creating and maintaining distilled wiki pages on frequently discussed topics helped efficiently provide new members with the knowledge necessary to comfortably contribute to the email list. In addition, new members were attracted to the community through these more findable and refined wiki pages because of their persistent URLs and changeable content.
- Referring to wiki pages in reply messages helped lower the cost and improve the efficiency of question answering, helping retain core participants. Former members also were able to stay connected as peripheral participants through contributing to the wiki.

Community maintenance is, of course, a social process. It is about attracting people, making them feel good about participating, and channeling that individual participation so that it makes others also want to be part of the community. This study highlights, however, that the art of community maintenance in help-based communities is, in part, a problem of designing information collections and information flows. It is easier to gently keep people on-topic in email-based push communication if their off-topic interests can be acknowledged and partially satisfied through external, pull sources. Similarly, the socially destructive recapitulation of inconclusive discussions can be channeled into the collaborative creation of documents where authors express many positions and jointly agree to disagree. A community repository reduces the costs of providing useful contributions in conversations because the conversations can reference particular documents in the collection and assume that other members are familiar with them. A repository can help attract members and provide them with a useful information scent that accurately conveys the kinds of information and communication opportunities that the community offers, without requiring them to sift through the raw email list archives. It can also help members get up to speed so that they are comfortable participating. Thus, the design of information environments can heavily influence the ways in which community maintenance occurs.

Acknowledgements

This work has been funded, in part, by the following National Science Foundation grants: IRI-9702904, IIS-0325837, IIS-0308006, and 0114368. Special thanks to Incutio for providing access to the wiki log reports and to the extraordinarily helpful administrators and members of css-d for the many hours of time they devoted to this project.

References

Ackerman, M. (2000). The intellectual challenge of cscw: The gap between social requirements and technical feasibility. *Human-Computer Interaction, 15*(2-3), 179-204.

Advanced web statistics 6.5. (2006). Available at http://awstats.sourceforge.net/

Benkler, Y. (2002). Coase's penguin, or, linux and the nature of the firm. *Yale Law Journal, 112*(3), 369.

Bos, B., Çelik, T., Hickson, I., & Lie, H. W. (2006). *Cascading style sheets, level 2 revision 1 (working draft)*: W3C.

Brin, S., & Page, L. (1998). *The anatomy of a large-scale hypertextual web search engine.* Paper presented at the Proceedings of the seventh international conference on World Wide Web 7, Brisbane, Australia.

Burnett, G., & Buerkle, H. (2004). Information exchange in virtual communities: A comparative study. *Journal of Computer Mediated Communication, 9*(2).

Butler, B., Sproull, L., Kiesler, S., & Kraut, R. (2002). Community effort in online groups: Who does the work and why? In W. S. & L. Atwater (Eds.), *Leadership at a distance*. Mahwah, NJ: Erlbaum.

Butler, B. S. (2001). Membership size, communication activity, and sustainability: A resource-based model of online social structures. *Information Systems Research, 12*(4), 346-362.

Button, G., & Dourish, P. (1996). *Technomethodology: Paradoxes and possibilities.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems: common ground, Vancouver, British Columbia, Canada.

Dery, M. (1993). *Flame wars: The discourse of cyberculture*. Durham, NC: Duke University Press.

Durrance, J. C. (1984). *Armed for action: Library response to citizen information needs*. New York, NY: Neal-Schuman.

Durrance, J. C., & Pettigrew, K. E. (2002). *Online community information: Creating a nexus at your library*. Chicago, IL: American Library Association.

Durrance, J. C., Souden, M., Walker, D., & Fisher, K. E. (2006). Community problem-solving framed as a distributed information use environment: Bridging research and practice. *Information Research, 11*(4), paper 262.

Ebersbach, A., Glaser, M., Heigl, R., & Dueck, G. (2006). *Wiki: Web collaboration*. Berlin; New York: Springer.

Eisenhardt, K. M. (1989). Building theories from case study research. *The Academy of Management Review, 14*(4), 532-550.

Feagin, J. R., Orum, A. M., & Sjoberg, G. (Eds.). (1991). *A case for the case study*. Chapel Hill, NC: The University of North Carolina Press.

Fisher, D., Smith, M., & Welser, H. T. (2006). *You are who you talk to: Detecting roles in usenet newsgroups.* Paper presented at the 39th Annual Hawaii International Conference on System Sciences, 2006, Manoa, Hawaii.

Giddens, A. (1986). *The constitution of society: Outline of the theory of structuration* (Reprint edition). Berkeley and Los Angeles: University of California Press.

Glaser, B. G., & Strauss, A. (1967). *The discovery of grounded theory: Strategies for qualitative research*. New York: Aldine Transaction.

Hughes, J., King, V., Rodden, T., & Andersen, H. (1994). *Moving out from the control room: Ethnography in system design.* Paper presented at the Proceedings of the 1994 ACM conference on Computer supported cooperative work, Chapel Hill, North Carolina.

Internet Archive. (n.d.). Available at http://www.archive.org

Kim, A. J. (1999). *Community building on the web: Secret strategies for successful online communities*. Berkeley, CA: Peachpit Press.

Kollock, P., & Smith, M. A. (1996). Managing the virtual commons: Cooperation and conflict in computer communities. In S. Herring (Ed.), *Computer-mediated communication: Linguistic, social, and cross-cultural perspectives* (pp. 109-128). Amsterdam: John Benjamins.

L-Soft international, Inc. (2003). Listserv introductory list owner's guide.

Lakhani, K. R., & von Hippel, E. (2003). How open source software works: "free" user-to-user assistance. *Research policy., 32*(6), 923.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge [England]; New York: Cambridge University Press.

Leuf, B., & Cunningham, W. (2001). *The wiki way: Quick collaboration on the web*. Boston: Addison-Wesley.

Marshall, C., Shipman, F., & McCall, R. (1994). *Putting digital libraries to work: Issues from experience with community memories*. Proceedings of Digital Libraries '94: the first annual Conference on the Theory and Practice of Digital Libraries. College Station, TX.

McLaughlin, M. L., Osborne, K. K., & Smith, C. B. (1995). Standards of conduct on Usenet. In S. Jones (Ed.), *Cybersociety* (pp. 90-111). Thousand Oaks, CA: Sage.

Net lingo: The internet dictionary. (n.d.). Holy War. Retrieved May 25, 2007, from http://www.netlingo.com/lookup.cfm?term=holy%20war

Orlikowski, W. J. (2000). Using technology and constituting structures: A practice lens for studying technology in organizations. *Organization Science, 11*(4), 404-428.

Powazek, D. M. (2002). *Design for community: The art of connecting real people in virtual places*. Indianapolis, Ind: New Riders.

Preece, J. (2000). *Online communities: Designing usability, supporting sociability*. New York: John Wiley.

Preece, J. (2004). Etiquette online: From nice to necessary. *Communications of the ACM, 47*(4), 56-61.

Schuler, D. S. (1996). *New community networks: Wired for change*. Reading, MA: Addison-Wesley.

Sproull, L., & Patterson, J. (2004). Making information cities livable. *Communications of the ACM, 47*(2), 33-37.

WikkiTikkiTavi. (n.d.). Available at http://tavi.sourceforge.net/