
Renate Andersen and Anders I. Mørch, Department of Education, University of Oslo, Norway

Abstract

In this paper we investigate computer-supported collaborative learning (CSCL) and innovation in a large-scale distributed setting. Get Satisfaction (GS), a social media platform for involving customers in product development activities, is our case study. In order to identify how end users contribute to product development, we researched the interactions between end users, champions, and professional developers in this online community as they jointly constructed a shared artifact (a web application). We collected publicly available platform interaction data over a six-month period (N = 229 users). The methods we employed are social network analysis (SNA) and interaction analysis (IA), which we combined in a mixed-methods design. At the network level, we identified key actors according to centrality measures. At the interaction level, we zoomed in on specific interactions. We propose a model of mass collaboration in terms of four interaction patterns: 1) gatekeeping, control of excessive information sharing, 2) bridge building, spreading information across groups in the network, 3) general development, allowing professional developers to create new software functionality and update existing software, and 4) user-user collaboration, facilitating non-centrally organized development activities, ranging from feature requests to local development. We discuss our findings and compare them with related research.

Keywords: collaborative knowledge creation; interaction analysis; mass collaboration; mutual development; social media platform; social network analysis; user-driven innovation

1 Introduction

With the rise of web 2.0 technologies and social media there has been a turn toward users becoming content creators, and web platforms are now largely user-driven. There are many implications to this; for education, the internet is now a platform for collaborating, sharing and connecting people rather than just a source of information. This shift has enabled ordinary people to become active contributors by interacting in distributed settings and in multiple configurations, contributing a range of different skills and expertise. This “bottom-up” approach to knowledge production has challenged the “top-down” approach prevalent in other areas (e.g. Encyclopedia Britannica vs. Wikipedia; predefined curriculum vs. self-driven learning) and is characterized by collective and open-ended production where an anonymous mass constitutes an important stakeholder. A common term for this is mass collaboration (Cress et al., 2013; Tapscott & Williams, 2008).

The essence of mass collaboration resides not only in new technologies and enhanced connectivity but also in the interaction and collaboration of a large number of participants from different places and time zones (Halatchliyski et al., 2014) and in the creation of shared artifacts (Paavola, Lipponen, & Hakkarainen, 2004;
Moenn, Mørch, & Paavola, 2012). New opportunities for collaboration have become possible through social media, where users can contribute to knowledge building in the large, such as on Wikipedia (Cress et al., 2013). A consequence of this for education is that small group collaboration needs to be understood within the framework of mass collaboration. However, the majority of studies in the field of computer-supported collaborative learning (CSCL) investigate collaboration and knowledge building in small groups and classrooms (Stahl, Koschmann, & Suthers, 2006; Scardamalia & Bereiter, 2006). In the present study, we used empirical data from the Get Satisfaction (GS) online community to investigate interaction between end users, champions, and professional developers in their joint effort to improve a software product. According to Hyysalo et al. (2010), there is a need for improved research methods for investigating product development collaboration between professional developers and end users, as previous studies have glossed over moment-to-moment interactions and knowledge sharing in user-developer relations. Our approach is unique in that it combines two traditions: CSCL and innovation studies. To accomplish this, we used two levels of data: fine-grained empirical data of social interactions and a global view of the network and its social structures. We analyzed these two levels in conjunction through a mixed-methods framework to propose interaction patterns of user-developer collaborations.

The research question guiding our inquiry is “What are the patterns of interaction between end users and professional developers in a mass collaboration community, as seen from a mutual development perspective?” The first part of the question addresses the network level (information paths and powerful actors), and the second part incorporates what the participants discuss and what roles they perform. We wanted to investigate both the interaction patterns of the online GS community and focus on some specific social practices in the community during mutual product development in order to extend prior research. In our previous work (Andersen & Mørch, 2009; Mørch & Andersen, 2010), five sub-processes of user-developer interactions were identified at the small-group level, leading to the formulation of the term “mutual development.” Here, we take this one step further by scaling up from small-group collaboration to mass collaboration and integrating SNA and IA.

Several researchers acknowledge social network analysis (SNA) as a highly relevant and necessary research method for describing and understanding interaction patterns in CSCL (Halatchliyski et al., 2014; Cress et al., 2013). For example, Halatchliyski et al. (2014) have underscored how SNA is a unique and largely unexplored method for tackling the large-scale dimensions of mass collaboration within CSCL. De Laat et al. (2007) have used SNA in a mixed-methods approach to provide an analytical framework for understanding message exchanges among mass collaboration participants, capturing a richer, and more accurate, picture of the complexity of such conversations. Our approach differs from previous studies in two ways: 1) it combines SNA with interaction analysis (IA) and 2) it explores patterns of interaction around the development of a software artifact.

The paper is organized as follows: First, we survey related work. Then, we describe our case study and the context for the study. Next, we present and argue for our mixed-methods approach to data collection and analysis. We then present and analyze our empirical data by focusing on topical findings, presenting representative excerpts as instances of mass collaboration in customer-initiated software product development. We compare our findings with results reported in the literature we surveyed. Finally, we summarize our findings and suggest some directions for future research.
2 Literature Review

2.1 Mass Collaboration

Tapscott and Williams (2007) coined the term “mass collaboration” to describe how people can join forces in self-organized communities to dynamically produce new goods and services. Their work has not been without critique. For example, Elliott (2007) claimed that they failed to provide an adequate definition or criteria for discerning collaboration from other collective activities such as cooperation and coordination, making the term a buzzword and stripping it of analytical value.

Cress et al. (2013) differentiated between formal and informal learning when they defined mass collaboration, emphasizing that formal learning involves knowledge building in smaller groups in classroom settings, whereas mass collaboration is about knowledge building “in the wild,” usually outside educational institutions and often in informal or semi-formal contexts of work and leisure activities. They found that these activities induce individual learning while demonstrating collaborative knowledge creation, further developing shared knowledge artifacts (Cress et al., 2013).

Halatchliyski et al. (2014) have discussed the relevance of mass collaboration for CSCL, proposing collective knowledge to be constituted as substance and by participation. By studying article production in Wikipedia, their study shows how collective knowledge is manifested in the structure of artifacts and can be traced back to the collaborative activity of authors with different levels of experience and expertise (Halatchliyski et al., 2014). Wikipedia’s interconnected articles represent a network and were thus analyzed using a network analysis approach. This form of mass collaboration was defined as a knowledge building activity: creating shared knowledge based on existing, openly accessible knowledge in collaboration with many other users.

Forte (2005) is critical of the use of the term “knowledge building” in conjunction with mass collaboration, arguing that the discursive processes associated with article creation in Wikipedia cannot be associated with knowledge building. Yet, despite the critique of the collaborative learning potential of Wikipedian discourses during article creation and discussion, Forte (2015) agrees there has been little attention paid to how information is selected, vetted, and verified by learners in this community. Our study differs from these studies in that we address another form of mass collaboration, mutual development, manifest in two ways: 1) there is an asymmetrical relationship between the participants, professionals (software developers) and amateurs (customers and end-user developers) and 2) the goal of their activity is to produce both concrete (software tools) and abstract (knowledge) artifacts.

2.2 SNA Studies in CSCL Research

De Laat et al. (2007) used social network analysis to study patterns of interaction in a networked learning community, investigating how its members share and construct knowledge. The authors used a mixed-methods approach, combining content analysis, critical event recall and SNA.

Siqin and colleagues (2015) investigated synchronous discourses between 27 Chinese undergraduate students collaborating in fixed groups during an introductory research methods course. They used a multifaceted analysis (involving social network analysis and content analysis) to assess online discourses and examine its potential relationship to individual learning throughout the course, as well as to examine different aspects of collaboration.
Martinez et al. (2007) have suggested that SNA can serve as an appropriate method for studying interaction patterns and provided examples of this in three different CSCL scenarios. They demonstrated how effective the method is for supporting the study of participatory aspects of learning at the network level.

Our study differs from previous studies in that we used a mixed-methods approach, combining SNA and IA, and applied this to discourse processes in mass collaboration. Our work builds on the notion of “collaborative knowledge creation” developed in the European KP-Lab (Paavola, Lipponen, & Hakkarainen 2004; Moen, Mørch & Paavola, 2012) and integrates CSCL with innovation studies.

2.3 User-Driven Innovation
User-driven innovation (UDI) refers to innovation by end users, customers, or consumer of products. Eric von Hippel (2005) argued that integrating active users in companies’ product development processes may lead to product innovation and value creation. He introduced a method for identifying sources of innovation by following “lead users.” A lead user is an inventor of something (product, feature, or idea) that is picked up and transformed into an innovation by a company or an early adopter and champion of the product. UDI been studied in open-source development (OSD) and thus is connected with mass collaboration in the sense that open source development was a source of inspiration for the notion of “peering” in mass collaboration (Williams & Tapscott, 2007).

The private-collective model of innovation is a conceptual framework for understanding UDI (Von Hippel & Von Krogh, 2003). It combines the private-investment model and the collective-action innovation model in order to explain the creation of public goods through private initiatives. It is based on the assumption that innovators who privately participate in creating public goods benefit more than free riders who consume those goods. A similar process occurs in mass collaboration on software development when interested users have the skills and tools to locally adapt (e.g., tailor) the software to their own needs (e.g., through hacks, mods, and other local development techniques) (Mørch, Hansen, & Ludvigsen, 2007). Collaborative tailoring of software by end users can bring benefits to a group working together to locally adapt a shared software system (Kahler, 2001). Based on a survey of previous work and an empirical study, Kahler (2001) suggested eight design principles for collaborative tailoring: 1) giving “administrator rights” to users to access, modify, or share tailored files, 2) sharing tailored files, 3) browsing tailored files, 4) providing users with awareness of others’ tailoring activities, 5) making annotations and automatic descriptions possible, 6) allowing for the exploration of tailoring files, 7) making administration and coordination easy, and 8) supporting a tailoring culture. This form of collaboration demonstrates the benefits of involving end users in the development process. Our analysis is informed by the concept of mutual development (Andersen & Mørch, 2009; Mørch & Andersen, 2010), or interaction between two interdependent software development activities: end-user (local) development and professional (in-house) development.

3 Case Description and Participants
The data presented in this paper is from a case study of Get Satisfaction (GS), a company which offers online customer management tools and provides a community platform for facilitating collaboration and interaction between participants, most of whom are customers either of GS itself or a company using GS tools. Get Satisfaction was established in 2007 and empowers more than 63,000 online communities; it has more than
9,600,000 visitors a month. The object of this study was collaboration on the GS platform in order to further co-develop it; we followed one of the communities over a six-month period in 2012-2013.

The support community at Get Satisfaction is continually evolving and is organized around question/answer (discussion) forums, where anyone who is registered can start a new thread or topic and tag it with one of the system’s four thread categories: 1) Ask a question, 2) Share an idea, 3) Report a problem, and 4) Give praise. A screenshot of these categories can be seen in Figure 1.

![Screenshot of the online support community and the four different categories for tagging discussion threads.](image)

**Figure 1:** Screenshot of the online support community and the four different categories for tagging discussion threads.

Table 1 shows the enrollment data for all threads during the time of the study. GS employees are referred to as developers, customers are referred to as end users, and champions belong to both classifications. Champions are customers appointed by GS employees after demonstrating extraordinary skills with GS software. They are given special privileges and may become part-time GS employees, paid for their support to the community.

**Table 1:** Tree snapshots of participation count according to user groups in Get Satisfaction forums over six-month intervals. The vast majority of participants are end users (customers).

<table>
<thead>
<tr>
<th>Overview</th>
<th>March 2012</th>
<th>September 2012</th>
<th>March 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topics posted</td>
<td>13,512</td>
<td>14,850</td>
<td>19,747</td>
</tr>
<tr>
<td>Participants (all)</td>
<td>259,119</td>
<td>262,017</td>
<td>269,280</td>
</tr>
<tr>
<td>Champions</td>
<td>29</td>
<td>29</td>
<td>47</td>
</tr>
<tr>
<td>Employees</td>
<td>43</td>
<td>43</td>
<td>50</td>
</tr>
<tr>
<td>End-users</td>
<td>259,047</td>
<td>261,945</td>
<td>269,183</td>
</tr>
</tbody>
</table>

Participants contribute by posting new messages, replying to previous messages, commenting, identifying moods, “liking” proposals, and rating each other’s contributions by giving “stars.” Figure 2 is a screenshot of a few conversations from the “Share an idea” category.
**Figure 2:** Two screenshots of a discussion thread “Add notification preferences that are product-specific.” The topmost image shows the message that spawned the thread, and some replies are shown below.

The GS online community is reminiscent of a social media platform. We were interested in studying the collaboration spaces in this community during mutual development, with a focus on how new ideas for the further development of existing products are initiated, shared, and spread in the community, captured by talk (text messages) inside the discussion threads, and emerge as recurrent interaction patterns in mass collaboration.

### 4 Methods and Research Design

#### 4.1 A Mixed-Methods Approach

We used a mixed-methods research approach to address the research question; this means that both quantitative and qualitative data were used (Tashakkori & Teddlie 2010). We combined social network analysis and
interaction analysis (Fugelli, Lahn, & Mørch, 2013). Relational aspects of social structure were the unit of analysis for SNA, and conversational turn taking in discussion forums was the unit of analysis for IA. Roughly speaking, we used SNA to describe the “climate” of our data and IA to describe the “weather.” This means we used SNA to get an overview and identify global attributes such as the most powerful participants, how they emerge as such, what subgroups they belong to, and how discussion forums and individuals are connected across all forums. We used IA to go into the details of topic of conversation. IA is not commonly used in conjunction with SNA because SNA largely depends on quantitative data and IA on qualitative data, therefore a mixed-methods approach (Tashakkori & Teddlie 2010) was used to combine them, which we explain below.

4.2 Collecting the Empirical Data

Virtual ethnography, researching internet communities (Hine, 2008; Kozinets, 2010), is the method we employed to collect the data. Virtual ethnography was a plausible method because the interactions are naturally occurring in the GS support community (i.e., they are a publicly available social phenomenon). When gaining access to the discussion threads, no login name or password credentials were needed; the information was open. However, the first author also sent a letter to GS asking for permission to conduct research on their social media platform. Our participant observation approach is referred to as non-obtrusive (Kozinets, 2010) or “fly on the wall” (Hine, 2008), meaning we as researchers were not visible in the online community during field research as we did not post anything. The first author followed the postings of the support community from the beginning of March 2012 to the end of August 2012. Additionally, earlier discussion threads that had been marked with the tags “Share an idea” and “Give praise” were scanned. When collecting the empirical data we as researchers are obliged to reveal as much as possible about the nature and aim of our studies to our informants, as well as providing information about our methods (Boellstorff et al., 2012). However, the notion of informed consent in virtual ethnography is difficult to follow through. Therefore, a passive consent for each discussion thread being analyzed in GS online community was created. The first author gained passive consent by posting a message to the respective discussion threads and informing the participants that the discussion thread was part of a research project and if some persons in the community did not want to participate they could opt out by sending an email to the first author and the person’s data would be removed from the research project. None chose to use this option.

4.3 Selecting and Categorizing the Empirical Data

The selection of the empirical data, participants, and population were done in steps. First, we decided to extract 41 discussion threads from the GS support community. These threads were selected due to their relevance to our research question and our research interest in mutual development (Andersen & Mørch, 2009; Mørch & Andersen, 2010, Andersen & Mørch, 2013). This led us to the top-level categories of “Share an Idea” and “Give Praise” because the process of mutual development starts with end users sharing an idea for developing a product further (Andersen & Mørch, 2009; Mørch & Andersen, 2010. Sharing an idea entails proposing an idea for improvement to (parts of) an existing product (Figure 1). Giving praise means to credit those who do a substantial amount of work for others in the community. To capture the historical record and context for the interactions we decided to go six months back in time to collect those discussion threads that were still ongoing when we started data collection. From there, we chronologically focused on the 41 threads. We ignored the two other sub-communities (“Ask a question” and “Report a problem”), mainly for practical reasons. Due to the large
numbers of participants (269,280) and postings (19,747) (see Table 1), it was necessary to narrow down the data in order to perform the research; from this, the need for new methods emerged, and we chose to conduct SNA in combination with IA. Data selection and categorization followed a two-step zoom-in process: 1) network data selection and 2) interaction data selection.

4.3.1 Network Data Selection
We used a collection of related software tools to perform our network analysis. The first author used a combination of Discourse Network Analyzer (DNA), University of California Irvine NETwork (UCINET), and NetDraw. DNA was used to collect the empirical data by transforming text-based conversational data into a formal network (Leifeld, 2010). The first author manually coded 4,275 lines of forum messages into DNA, generating a set of higher order codes before exporting it to UCINET for SNA analysis. The selected population included 229 different users who were affiliated with 41 discussion threads, involving 546 code statements (utterances) and 31 higher order codes (Table 2). UCINET (Borgatti et al., 2002) represents data as matrices and can perform matrix operations to generate results. The results are presented as tables of measures of nodes and visual displays of networks of nodes and ties using the drawing tool NetDraw (Borgatti et al., 2002).

<table>
<thead>
<tr>
<th>Participants</th>
<th>Discussion threads</th>
<th>Higher order codes</th>
<th>Utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>229</td>
<td>41</td>
<td>31</td>
<td>546</td>
</tr>
</tbody>
</table>

At the network level, we identified the most active discussion threads and participants (nodes) according to centrality measures (Freeman, 1979). We selected the 20 most active threads and narrowed the field further by choosing threads that revealed cross-disciplinary (developer-user) interactions (i.e. informed by our research question). We then selected four empirical excerpts from conversations occurring in the most central and largest discussion threads according to degree centrality and betweenness centrality (Freeman, 1979). In addition we screened the data from a qualitative perspective and looked for relevant data according to our research questions.

4.3.2 Interaction Data Selection
We thematically coded the interaction data (textual conversations) using template analysis. We included all utterances in each of the four empirical excerpts that were chosen. The essence of template analysis is that the researcher produces a set of codes (a “template”) representing themes identified in their textual data (King, 1994). This combines top-down and bottom-up analysis, resulting in a set of intermediate terms (higher order codes in Table 2) to create the templates. King (1994) distinguishes three features of template analysis: defining codes, hierarchical coding, and parallel coding. Examples of defined codes are “end user sharing an idea for feature request,” “user providing solution for suggested feature request,” “feature request implemented into general product,” and “champion providing solution for suggested feature request.” We then went through the empirical data once more in order to cluster it into part-whole groups, as an approach to hierarchical coding (King, 1994). Examples of these higher order codes were “further development of the product,” “general development,” “customer-initiated development,” and “user-user collaboration.” Parallel coding, which refers to
the classification of the same segment within two different codes at the same level, was not relevant to our data analysis (King, 1994).

4.4 Data Analysis

Data were analyzed both quantitatively and qualitatively; first, SNA was used to get an overview of the totality of data (network structure and powerful actors), and then IA was used on individual messages (what is talked about). This was not a strict linear process; we moved between scanning SNA and IA data.

4.4.1 Social Network Analysis

As a quantitative approach to studying large online communities, SNA provides a set of techniques and operations for analyzing the relational aspects of social structures (Scott, 2001), drawing on a library of algorithms and computational techniques (Freeman, 1979; Borgatti, Everett, & Johnson, 2013). We computed the degree centrality and betweenness centrality for each actor (Freeman, 1979) using UCINET (Borgatti, Everett, & Freeman, 2002). These measures were chosen because they were relevant to our research focus and allowed us to identify powerful actors in the network. Degree centrality is defined as the number of ties/links incident upon a node (Scott, 2001). In our case, this means the number of messages a person has posted to a discussion thread, and for discussion thread nodes the number of participants who post in a thread. A person with a high degree centrality has “selection power” (i.e., the ability to choose among alternatives) and it makes this person less dependent on any specific actor (Hanneman & Riddle, 2005).

Betweenness centrality, on the other hand, is defined as the number of times a node is on the shortest path between two other nodes in a network (Freeman, 1977). This means that the more persons that are dependent on a user to make connections with others in the network, the more power that user has, which is a form of “control power” (Hanneman & Riddle, 2005). In our case, betweenness centrality can tell us the probability of a user being in a position to share or withhold information or otherwise exercise control over who receives information in the community. A person with high betweenness centrality in the GS community contributes to several discussion threads but has few postings in each thread, contributing mainly to those threads that are likely to be decisive (e.g., succeed in or, alternatively, fail to implement a new idea). Degree and betweenness centrality were calculated for the 20 most active participants and discussion threads in the network.

4.4.2 Interaction Analysis

In order to analyze the content of messages in the different discussion threads, we used a version of IA (Jordan & Henderson, 1995), where we looked at log data in discussion threads. Thus, we applied the idea of IA to focus on turn taking in connection with exchanged textual messages in the online community, along the lines suggested by Arnseth (2004). The online setting precluded an analysis of language work in terms of linguistic utterances in f2f discussions and the separate analysis of artifacts through deictic references (Jordan & Henderson, 1995; Arnseth, 2004). By “interaction” we refer to one participant’s response to another utterance in the discussion thread. Artifacts are illustrated by end results (screen images) produced by the online community.

The following central concepts of interaction analysis were useful for our analysis: turn taking, micro-level analysis, and structure of events. Turn taking occurs when a new post in a discussion thread responds to a previous message or initiates a new discussion, and it encompasses the whole range of behaviors through which people can contribute and participate in an “interactional exchange system” (Jordan & Henderson, 1995). Each
utterance was numbered and studied in detail and interpreted in the context of the utterances that preceded it and to what it could mean or refer to. This is a process that Jordan and Henderson (1995) refer to as “micro-level analysis.” IA also examines the temporal organization of moment-to-moment interaction, which provides context for an event with its high and low points (Jordan & Henderson, 1995). “Structure of events” is defined as “stretches of interaction that cohere in some manner that is meaningful to the participants”, which through “a chronological time provides analysts with a standardized time line for the activities” and “events always have a structure (minimally beginnings and endings)” (Jordan & Henderson, 1995), p.57. In the reproduction of discussions below, the participants’ contributions and replies are organized chronologically in order to follow the discussion as a series of events.

5 Empirical Findings and Discussion

In this section we present and analyze our data. First, we present the network-level data to provide an overview of all collected data. Second, our main findings are presented as a combination of the network data and the interaction data. It should be noted that the network data gives both a bird’s eye view of the whole data set and a rationale for the subsequent interaction data selection. The interaction data is presented as a set of excerpts, numbered 1-4, reproduced from the discussion threads and representing the more detailed analysis.

5.1 General Network Characteristics

Two types of networks can be analyzed with SNA: single-mode (actor-to-actor) networks and dual-mode (affiliation) networks (Borgatti, Everett, & Johnson, 2013). In our case, it was necessary to construct an affiliation network of the GS online community. An affiliation network (Breiger, 1974; Faust, 1997), defines two non-overlapping sets of nodes (actor nodes and affiliation nodes), in our case the participants and the discussions they take part in. The affiliation network is thus our most comprehensive model of the network. Figure 3 visualizes participants and threads as differently shaped (four different shapes) colored nodes (rainbow color range) and represents the entire population of 229 users and 41 discussion threads; it also shows both degree and betweenness centrality of all network nodes. Links between two nodes do not reflect direct conversation between two persons but rather shared participation in a thread (indirect links). This implies that we assume those who contribute in the thread have read the postings to which they respond, and on that basis we create an indirect connection between the two persons (e.g. questioner and answerer).

In Figure 3, node size indicates degree centrality, which in our case means the number of threads a person has posted to. The degree centrality of a thread is the number of participants who participate in it (with the number of postings of each contributor as a numeric qualifier). The data obtained from Figure 3 helped us to address our research question and informed our subsequent interaction analysis by extracting the following information: 1) the most active discussion threads (thread degree centrality) and 2) the most active contributors in terms of number of postings (actor degree centrality).
Figure 3: Sociogram of the affiliation network of all contributors (N = 229) of “Share an idea” and “Give Praise” threads. Node size reflects degree centrality; shape reflects node type (end user: circle; developer: triangle; champion: diamond; thread: square); rainbow color range reflects betweenness centrality (red = lowest; violet = highest).
We used SNA to compute degree centrality and identify the 20 most active discussion threads in the community (Table 3). Degree centrality was normalized and reported as a percentage of the maximum degree centrality a node can have (UCINET output). We included the degree centrality of a node (actor and discussion thread) and the network, the latter meaning the network’s “distribution of centrality” (Hanneman & Riddle 2005; Borgatti, Everett, & Freeman, 2013).

Table 3: Freeman’s degree centrality of the 20 most active discussion threads. Numbers are in percentages

<table>
<thead>
<tr>
<th>Discussion thread</th>
<th>NrmDegree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Include referring URL and OS browser details in GS interface</td>
<td>20.137</td>
</tr>
<tr>
<td>Customers should be able to opt in to notifications for Company updates</td>
<td>16.705</td>
</tr>
<tr>
<td>Offer an easy way to import content to GS</td>
<td>15.789</td>
</tr>
<tr>
<td>Offer sticky or featured topics</td>
<td>14.645</td>
</tr>
<tr>
<td>Clean up the topic pages</td>
<td>14.188</td>
</tr>
<tr>
<td>Add notification preferences that are product specific</td>
<td>11.556</td>
</tr>
<tr>
<td>Preview before posting a topic</td>
<td>11.442</td>
</tr>
<tr>
<td>Create a how to/FAQ/tutorial topic type</td>
<td>11.327</td>
</tr>
<tr>
<td>Close replies</td>
<td>9.039</td>
</tr>
<tr>
<td>Don't accept feedback then demand signup</td>
<td>8.810</td>
</tr>
<tr>
<td>I’d like to export the topic title and details to Excel</td>
<td>8.810</td>
</tr>
<tr>
<td>Highlight the last post made</td>
<td>8.467</td>
</tr>
<tr>
<td>Company favicon on the company page</td>
<td>8.467</td>
</tr>
<tr>
<td>Privacy settings</td>
<td>8.238</td>
</tr>
<tr>
<td>Feedback Tab widgets should include company name</td>
<td>7.437</td>
</tr>
<tr>
<td>Better rate-limiting to help prevent spam</td>
<td>7.323</td>
</tr>
<tr>
<td>Using GS for feature request and voting – not support</td>
<td>6.407</td>
</tr>
<tr>
<td>Remove nofollow attributes from the links to the company domain</td>
<td>5.950</td>
</tr>
<tr>
<td>Topics in need of attention Community Tab</td>
<td>5.721</td>
</tr>
<tr>
<td>Link to an image in a post</td>
<td>5.492</td>
</tr>
</tbody>
</table>

The network centralization of all discussion threads is 14.033 (Table 4), which means there are many discussion threads in the GS online community (a number closer to 100 would indicate fewer and larger threads) (Scott, 2001). We selected the following three threads for subsequent interaction analysis: 1) “Create a How To/FAQ/Tutorial topic type,” 2) “Offer sticky or featured topics,” and 3) “Add notification preferences that are product-specific.” The rationale for this selection is that the three threads all have higher degree centrality (Table 3) than the average/mean degree centrality in the network (6.824, Table 4). The chosen threads are therefore among the most active discussion threads in the online community. Furthermore, the network degree centralization variance, which is 21.773, indicates that some discussion threads have more activity than others. A low variance would indicate an equal distribution of degree centrality in the network (Hanneman & Riddle, 2005). In addition to centrality measures, another criterion we used for selecting discussion threads was their relevance to our research question (i.e. is do they relate to mutual development, user-driven innovation, and developer-user interactions).
Table 4: Descriptive statistics of all discussion threads. Numbers are in percentages.

<table>
<thead>
<tr>
<th>All discussion threads</th>
<th>NrmDegree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average/Mean</td>
<td>6.824</td>
</tr>
<tr>
<td>Variance</td>
<td>21.773</td>
</tr>
<tr>
<td>Network centralization</td>
<td>14.033</td>
</tr>
</tbody>
</table>

Next, we converted the two-mode (affiliation) network into a single-mode actor-by-actor matrix using UCINET (Borgatti et al., 2002). This allowed us to calculate degree centrality and betweenness centrality for each actor (Freeman, 1979) and identify the most active contributors and information brokers in the network. The resulting UCINET outputs are placed next to each other in Table 5; the numbers shown are percentages. The “NrmDegree” and “NrmBetweenness” columns in Table 5 are normalized values, a percentage of the maximum centrality value one can have (a number closer to 100 in the first column would indicate a user had contributed to almost all of the discussions).

Table 5: (Left) normalized degree centrality for the 20 most active participants in terms of selection power, and (right) normalized betweenness centrality for the most active participant in terms of control power. Numbers are in percentages

<table>
<thead>
<tr>
<th>Participant</th>
<th>NrmDegree</th>
<th>Participant</th>
<th>NrmBetweenness</th>
</tr>
</thead>
<tbody>
<tr>
<td>User 9</td>
<td>11.307</td>
<td>User 9</td>
<td>18.399</td>
</tr>
<tr>
<td>End-user_12</td>
<td>10.475</td>
<td>Developer_13</td>
<td>15.434</td>
</tr>
<tr>
<td>Developer_2</td>
<td>8.957</td>
<td>End-user_12</td>
<td>14.870</td>
</tr>
<tr>
<td>Developer_5</td>
<td>8.517</td>
<td>Developer_2</td>
<td>13.820</td>
</tr>
<tr>
<td>Developer_13</td>
<td>8.174</td>
<td>Developer_7</td>
<td>9.157</td>
</tr>
<tr>
<td>Champion_3</td>
<td>7.660</td>
<td>Champion_2</td>
<td>9.112</td>
</tr>
<tr>
<td>Champion_2</td>
<td>6.828</td>
<td>Developer_5</td>
<td>8.009</td>
</tr>
<tr>
<td>End-user_13</td>
<td>6.461</td>
<td>Developer_11</td>
<td>7.085</td>
</tr>
<tr>
<td>Champion_9</td>
<td>5.751</td>
<td>Champion_3</td>
<td>3.998</td>
</tr>
<tr>
<td>End-user_11</td>
<td>5.702</td>
<td>Developer_9</td>
<td>2.750</td>
</tr>
<tr>
<td>Developer_11</td>
<td>5.702</td>
<td>End-user_29</td>
<td>2.681</td>
</tr>
<tr>
<td>Champion_7</td>
<td>5.042</td>
<td>Champion_7</td>
<td>2.188</td>
</tr>
<tr>
<td>End-user_119</td>
<td>4.797</td>
<td>End-user_152</td>
<td>1.769</td>
</tr>
<tr>
<td>End-user_129</td>
<td>4.552</td>
<td>End-user_119</td>
<td>1.651</td>
</tr>
<tr>
<td>End-user_128</td>
<td>4.552</td>
<td>End-user_11</td>
<td>1.617</td>
</tr>
<tr>
<td>End-user_126</td>
<td>4.552</td>
<td>Champion_1</td>
<td>1.484</td>
</tr>
<tr>
<td>End-user_141</td>
<td>4.552</td>
<td>End-user_50</td>
<td>1.408</td>
</tr>
<tr>
<td>End-user_51</td>
<td>4.479</td>
<td>End-user_17</td>
<td>1.379</td>
</tr>
<tr>
<td>End-user_15</td>
<td>4.405</td>
<td>End-user_13</td>
<td>1.379</td>
</tr>
<tr>
<td>End-user_40</td>
<td>4.332</td>
<td>End-user_15</td>
<td>1.379</td>
</tr>
</tbody>
</table>

While SNA data provides useful information about the overall structure of a social network, it does not tell us anything about the quality of interactions or the content of postings. This limitation could lead to misguided interpretations of who is powerful in the network (Borgatti et al., 2002). Descending to a more detailed level of analysis and studying the content of participants’ messages is one way to address this limitation. In the next section, we describe the combination of SNA and IA in order to analyze the content of participants’ conversations in a mass collaboration context.
5.2 Analyzing Interaction Data and Network Data in Combination

In this section we present four excerpts of interaction data extracted from the selected discussion threads. The labels emerged as intermediate terms while screening and classifying the data. The first excerpt illustrates the phenomenon of “gatekeeping,” the second “bridge building,” the third “general development,” and the fourth “user-user collaboration.”

The empirical excerpts are presented in sequentially enumerated tabular format (Jordan & Henderson, 1995) and extended with SNA measures regarding the participants’ positions in the online community. The structure of the presentation of the empirical data is as follows: First, we have a column identifying the turn in the conversation. Then, the utterance made by the participant is presented. In order to integrate the SNA data, we added the following two columns: normalized degree centrality (abbreviated NrmDegree) and normalized betweenness centrality (abbreviated NrmBetweenness).

5.2.1 Excerpt 1: Gatekeeping

The first excerpt is from the discussion thread “Create a How To/FAQ/Tutorial topic type,” which was the eighth most active discussion thread in our data (Table 3). A champion moderates the conversation of two customers (End user 168 and End user 169), who propose slightly different solutions for a requested feature and disagree about which is the more generally applicable.

<table>
<thead>
<tr>
<th>Turn</th>
<th>Participant</th>
<th>Text from discussion thread</th>
<th>NrmDegree</th>
<th>NrmBetweenness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>End user 150</td>
<td>Create a How To/FAQ/Tutorial topic type. Could really use a 'knowledge base' topic where the company could post simple how-to’s, screencasts etc.</td>
<td>2.839</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>End user 151</td>
<td>I suppose the fact that this thread is almost 2 years old means I shouldn't hold my breath :( right now users can do one of the following: 1) ask a question 2) share an idea 3) report a problem 4) give praise. i'd like to suggest a fifth type 5) tutorial / wiki</td>
<td>2.839</td>
<td>0.00</td>
</tr>
<tr>
<td>3</td>
<td>Champion 8</td>
<td>We do plan to introduce this sort of topic type, but in the mean time if you have a topic that needs to be edited, let me know and I'd be happy to make the change for you.</td>
<td>1.371</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>Developer 19</td>
<td>Unfortunately I don't have an update for this Idea. I also really want this feature and will continue to bring it up to the team.</td>
<td>0.489</td>
<td>0.00</td>
</tr>
<tr>
<td>5</td>
<td>End user 168</td>
<td>Guys, I've created a solution, which may be helpful to you, based on some of the ideas in this thread. Take a look at my community at [name of website redacted for anonymity] Note on the right under Additional support links I have an FAQ page for each of my products. B) If you click on the first link, named Events-Ticketing.com FAQs, it will take you to Community Products Events-Ticketing.com FAQs. C). Note that each entry on</td>
<td>4.185</td>
<td>0.00</td>
</tr>
</tbody>
</table>
that page is a single question by an employee named FAQ with a little FAQ button as their photo. D) Each answer is an Official Response, again by employee FAQ. And there you have it - A dedicated FAQ page for each of my products that my customers love. It's not a perfect solution, but it does work well.

6 Developer 13 This is great, end-user 168! Thank you for sharing! 8.174 15.434

7 End user 168 My pleasure! 4.185 0.00

8 End user 169 Instead of creating a new product for each FAQ, why not just tag each FAQ topic with a PRODUCTNAME FAQ tag? To show your FAQ you just link to a page of all tagged topics. 2.839 0.00

9 End user 168 Because my products are actually separate businesses. I really don't want my software customers looking at items for my apparel customers, so simply tagging will just not cut it for me. Your mileage may vary, though. 4.185 0.00

10 End user 169 I get that. We have the same problem using GS with multiple products. I replied because your approach separates your FAQs from your businesses by defining them as individual products. To avoid that I'm suggesting, a specific FAQ tag for each of your businesses/products. You can then click on a tag to show all topics with that tag and show that page as your FAQ. In the end though, it's just another way of hacking round the issue until the GS folk make proper FAQs a feature. :-) 2.839 0.00

11 Champion 9 Good stuff, End user 168! I like it. Another way this could work, and one that has been discussed over time here and there on Get Satisfaction, would be to maybe promote a question asked by a customer to a FAQ — with the option of maybe stripping out unneeded conversation so that you could provide a really great answer to an authentically asked question. A true FAQ in a sense. 5.751 1.379

12 End user 168 Hi Champion 9, yes I was doing that for a while, but it's a lot of messing around. It's easier just to copy & paste as employee FAQ, and looks more consistent that way. 4.185 0.00

We consider this a representative instance of the interaction pattern we call “gatekeeping” in mass collaboration. It shows how champions and developers act as moderators or gatekeepers of information (and correspondingly may be unwilling to develop a requested feature). This excerpt starts out with End user 150 suggesting an idea to improve the product (a GS tool). End user 150 has a degree centrality of 2.839, indicating that the user is a very active and central participant in the thread (average degree centrality is 1.9). This person
has many contributions and/or responses concentrated in a few discussion threads. At the same time, End user 150 has low betweenness centrality (0), indicating that this person is not an influential actor in the network because no other people depend on this user to access information from other discussion threads (in which case End user 150 would have had higher betweenness centrality). The following is an illustrative comparison: In an infectious disease activation network, End user 150 would not be a powerful node because he cannot spread the disease to other people (because he has a betweenness degree of 0).

Responding to turn 1, End user 151 (turn 2) underlines that the development process has taken two years, revealing that gatekeeping can span a long time. Looking at the SNA data, we see that End user 151 has exactly the same degree and betweenness centrality as End user 150.

Champion 8 in turn 3 replies that GS is planning to introduce the possibility for including tutorials as topic type and by reassuring that he can help edit a topic, underscoring his assigned role as a champion, possibly with the ambition to become a gatekeeper and moderator of this process. The SNA data tells us that Champion 8 has a degree centrality of 1.371, which is below the average of 1.9, but he reveals that he has skills in making modifications (turn 3: “I'd be happy to make the change for you.”).

Six months later, in turn 4, Developer 19 tells the community that he does not have an update for the idea proposal because the product team has not yet approved it. This utterance shows that Developer 19 is in a position to act as an intermediary between the user community and the developer team, or in other words as a gatekeeper or moderator of information, deciding what information to pass on and how to act on it. However, Developer 19 could be new in this role because his centrality measures are low. The statement issued in turn 4, stating he “will continue to bring it up to the team,” indicates that he may not yet be on the developer team and does not have power to make any decisions. As a result, End user 168 (turn 5) replies that he has created a solution on his own and provides a web link to it.

Developer 13 acknowledges the solution and replies with gratitude, noting that this is great work. What is interesting here is that Developer 13 has very high betweenness centrality (15.434), the second highest in the GS community, which means that this person is a very central actor and contributes in many discussion threads. When such a prominent person is positive about the idea proposed by End user 168, we can interpret from our SNA data that this is important information and an effective solution. However, in SNA, “a high betweenness node” reflects the ability to threaten or stop transmitting information across the network, making communicating nodes use less efficient paths to reach one another (Borgatti, Everett, & Johnson, 2013). Looking at the IA data in isolation would not show this, as the textual message exchange demonstrates a positive response. The SNA data thus brings additional information about how to interpret the message (in this case, the power to control information by choosing to share or not share an effective solution with other developers).

As the conversation continues, the structure of events revolves around two end users (End user 168 and End user 169) exchanging ideas about two different solutions for implementing the originally suggested feature (turns 8-10). This exchange is interrupted by Champion 9, who selectively chooses End user 168’s proposal and suggests another way for it to be developed (turn 11). End user 168 later responds that Champion 9’s solution led to “a lot of messing around,” and thus he found his own “brute force” solution easier to implement.

Comparison to related work. From an IA perspective, gatekeeping is exemplified as a bracket of interaction where several champions and developers participate in a discussion thread, suggesting ideas and improvements for how to develop a feature requested by one or more end users, without resulting in the
implementation of a developer-suggested solution. The information provided by the champions and developers sometimes indicates control over (through moderating or hindering) the free flow of information in the network. For example, Champion 9 writes, “Good stuff, End user 168! I like it. Another way this could work is …” (turn 11) without bringing the proposal forward to the developer team.

Our claim that sometimes champions and developers are controlling information in the community is substantiated by our SNA data through degree and betweenness centrality measures. The betweenness centrality values are especially pertinent here (especially of Developer 13 and Champion 9, to a lesser extent), implying the control of information flows through the network, referred to in the SNA literature as a toll-taking role (Borgatti, Everett, & Johnson, 2013). Developer 13’s utterance in turn 6 is significant in this regard. This person is strategically positioned in the network (central for information flow by participating in multiple threads with potentially interesting ideas) and contributes here with a short remark, which turns out to have big impact on the conversation. Borgatti, Everett, and Johnson (2013, p 175) explain this behavior as follows: “Nodes with high betweenness are in a position to threaten the network with disruption to operations or in a position to distort data or information.”

5.2.2 Excerpt 2: Bridge Building
The extract below is from the discussion thread “Offer sticky or featured topics,” which in its full length contains 57 postings by 44 participants (the longest thread in the community). The extract includes the beginning of the thread and involves two end users, three champions, and one developer. “Sticky” refers to threads that are considered important and appear before others in discussion forums.

<table>
<thead>
<tr>
<th>Turn</th>
<th>Participant</th>
<th>Text from discussion thread</th>
<th>NrmDegree</th>
<th>NrmBetweenness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>End user 125</td>
<td>Offer sticky or featured topics</td>
<td>3.084</td>
<td>0.00</td>
</tr>
<tr>
<td>2</td>
<td>Developer 5</td>
<td>Hi, End user 125, You can make a reply &quot;sticky&quot; but we don't currently have a mechanism for making a post sticky. If you're a company rep you can use the &quot;Company Update&quot; topic type to post that topic on your company home page, which might partially solve the issue for you. Can you describe your need a bit more?</td>
<td>8.157</td>
<td>8.009</td>
</tr>
<tr>
<td>3</td>
<td>End user 125</td>
<td>I am a company rep in GS and we got this question from our users a couple of times. They see a post (be it a question or an idea shared), and they suggest making the thread/post sticky. And I just wanted to see if there is a way in GS to do so. Thanks for your reply. I will look into your suggestion.</td>
<td>3.084</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>Champion 1</td>
<td>Just got a similar request from one of our users. <a href="http://getsatisfaction.com/izea/topic">http://getsatisfaction.com/izea/topic</a>...</td>
<td>1.713</td>
<td>1.484</td>
</tr>
<tr>
<td>5</td>
<td>Champion 7</td>
<td>I've shared this with the product team - I'm working on pulling together a community-manager focused release to help get some of these ideas and bugs all bundled together for maximum awesomeness. Stay tuned. ;)</td>
<td>5.042</td>
<td>2.188</td>
</tr>
<tr>
<td>6</td>
<td>End user 131</td>
<td>Any progress on sticky topics?</td>
<td>3.084</td>
<td>0.00</td>
</tr>
</tbody>
</table>
In excerpt 2, champions (and sometimes professional developers) participate in an interaction pattern we call “bridge building.” Bridge building at the interaction level can be characterized by stating something in the discussion thread that can be interpreted as providing links or bridges to previous or later utterance; it thus serves to integrate participants in the discussion. The SNA data shows that End user 125, who initiated the thread, is a very active participant in the GS community. End user 125’s degree centrality is 3.084, which is higher than the average (1.9, but not among the top 20, see Table 5). Developer 5 responds by suggesting a solution that “might partially solve the issue” (turn 2). In this way Developer 5 takes on a bridge building role, mediating between the customer (End user 125) and the developer team. The response is centered around answering whether or not the requested feature is planned or not. In this capacity, Developer 5 acts as a broker of information and tries to help End user 125 find a solution. The claim that developer 5 acts as a bridge builder is supported by the SNA data by showing a very high degree centrality (8.157), ranking this person among the top four most active participants in the community (see Table 5). In addition, Developer 5 also has high betweenness centrality (8.009), among the top seven participants in the network (Table 5), meaning this person is strategically positioned to make decisions and contributes to important discussion threads (see Figure 3).

End user 125, who initiated the “bridge building activity” responds to Developer 5’s question “Can you describe your need a bit more?” (Turn 3) by describing the problem and how he would like the feature to be developed. Then, Champion 1 writes that other users have requested the same feature in the past, thus providing additional justification for the request and putting some pressure on the developers to find a solution. Champion 1’s degree centrality is slightly lower than the average (1.713), and his betweenness degree is above average (1.484, average of 0.55). Thus, Champion 1 is strategically located to spread or alternatively curtail information flow.

In turn 5, Champion 7 states that he has shared the idea with the product team, which indicates this may be an issue developers are working on. Champion 7 is an influential person in the community with both high degree and betweenness values (5.042 and 2.188, respectively). Champion 7 is thus both very active as poster and/or responder and can reach many participants in the network in few steps. Champion 7 is therefore in a position to pass valuable information to the developer team.

In turn 6, End user 131 asks for progress on “sticky topics.” There are two responses, one by Champion 7 and one by Champion 2. Champion 7 writes that it is a “tough change” and reiterates the sentiment in turn 5, stating, “I'll update over here once we've rolled it out,” which indicates it could take some time. Then, Champion 2 enters the discussion. Champion 2 is the most powerful champion in the network according to centrality measures. He has very high degree centrality (6.828) and also very high betweenness centrality (9.112). In turn 8, Champion 2 proposes a solution for how to incorporate the feature in either the left or right sidebar of the user interface (Figure 4). In this manner, Champion 2 acts as bridge builder by providing a technical solution that responds to user needs and advances the software.
Comparison to related work. There are four champions among the top 20 participants with the highest betweenness centrality (see Table 5), implying they play an important role in spreading information through the GS community. Some champions also have high degree centrality, implying they are active participants in many discussion threads. However, some of them (e.g., Champion 9) have relatively higher degree than betweenness centrality, implying their control/relay ability may not keep pace with their number of postings. By looking into the IA data for champions in excerpt 2, we can see that they relay more than they control. For example, Champion 7 states, “I've shared this with the product team …” (turn 5). This reveals a much more active response than that of Developer 19 in excerpt 1 (turn 4) on a similar issue (“Unfortunately I don't have an update for this Idea”).

Another example of positivity and outreach in excerpt 2 is Champion 2 (turn 8) stating, “I do think there is room for a "sticky" if we just arrange things a little…” This person’s bridge-building and gatekeeping capacity is supported by the SNA data, which shows that Champion 2 has very high betweenness centrality (9.992). According to Borgatti, Everett, and Johnson (2013), “a node with high betweenness centrality means that many nodes need that node to reach other nodes via efficient paths.” Because many other participants need Champion 2 to reaching one another (e.g., end users to developers), Champion 2 has a lot of power to spread or block information quickly. However, whether or not Champion 2 does this cannot be gleaned by the SNA data. We needed IA, corroborated by the first author’s reading the entire list of message exchanges, including those that preceded and followed the utterances presented here.

5.2.3 Excerpt 3: General Development
Excerpt 3 is an extract of the discussion thread “Add notification preferences that are product specific,” where a champion shares an idea for how to improve the product. This excerpt suggests a “path” from a user-requested feature to general (developer-organized) implementation of the feature by the company and the shared integration of the feature into the product (the online GS community software).
specific products.

3 End user 41 You know what else would be pretty awesome and not hard to do? If the emails had the product name in the subject so I could at least set up mail filters. 0.710 0.00

4 End user 40 Please please please... this has been in progress for over 3 years! Come on guys... this is seriously limiting our organizations adoption of Get Satisfaction. 4.332 0.864

5 End user 42 I like End user 40s idea. That feature would really make my life easier!!! 1.371 0.00

6 End user 43 Yes please! 0.000 0.000

7 Champion 7 I’m grateful for the suggestion! I’ve got this idea loaded up into our feature request queue, and I’ll update all y’all once I know a bit more. 5.042 2.188

8 End user 40 Hey, Champion 7, is there any update on getting product specific notifications into the product? 4.332 0.864

9 Champion 7 I’m checking in on the status of this one - I like this idea, too, but I'm not sure how complex/expensive it gets when it comes to our email system. I'll let y'all know when I know more. 5.042 2.188

10 End user 44 We have numerous products and they all arrive as emails to ”Highway Records” - it'd be great if there was some differential based on what product the feedback pertained to. Even changing the line within the body of the email where it says: ”End user 44 just shared this idea in Highway Records:” for instance to say ”End user 44 just shared this idea in Highway Records about this product: [name of product redacted for anonymity]” I could then easily set up filtering rules in Gmail (where I catch all these emails) to forward them through to other people in my organization. It's worth noting that you already do this with respect to private feedback (i.e. it states which product the feedback was in relation to) so it would seem like a super simple thing for you guys to add to the regular, non-private email alerts 0.710 0.00

11 End user 45 In our organization, different people are responsible for each product and/or service listed in our Get Satisfaction installation. As the local GS admin, I'm happy to receive all notifications but the people who look after each product have often asked me if they can limit their notifications to just the product they are responsible for. I eagerly await the development! 0.710 0.00

12 End user 40 Developer 11 and Developer 12; any update on the progress of this? 4.332 0.864

13 Champion 8 It looks like this feature is bundled into our notification improvements, but sadly I don't have an exact time frame for when this gets released. 1.371 0.00

14 Developer 7 I’m thankful that everyone has been so patient while we worked on other stuff): Hi everyone, I'm really glad to say 3.206 9.157
that we finally launched a Product follow feature. Learn more about it on our blog: url: http://product.getsatisfaction.com/2011/07/follow-products-on-get-satisfaction

The proposal for a new feature by Champion 1 (turn 1) is responded to by Developer 11 in turn 2, who says that the feature is already planned. Developer 11 has the highest degree and betweenness centrality in this discussion thread (5.702 and 7.085, respectively), which means that this is the most active and influential participant in the thread and that this person is also involved in several other discussion threads. This person sponsoring your idea would normally be a good sign for an end user or champion. Multiple requests are also made by other end users (turns 4-9). Champion 7 intervenes in turn 7 reporting the status of a feature (“I've got this idea loaded up into our feature request queue”), but in turn 9 this champion indicates uncertainly around the time it will take to implement it based on its complexity (“I'll update all y'all once I know a bit more”). The thread spans several years. Multiple end users ask about the status of the promised implementation (End users 40, 44, 45). End user 40 complains in turn 4 that the feature request has been in progress for over three years, implying a long-term development process. This end user is 20th in terms of degree centrality (4.322, Table 5, throughout the whole network, indicating that the user is very active with many posts in the GS community. Other end users continue to post messages (turns 10-12) emphasizing how important is for them to get the feature developed (End user 45 in turn 11 writes, “I eagerly await the development!”). Champion 8 (turn 17) notes that the idea is in the developer team’s pipeline but cannot provide a fixed date for its realization. Finally, Developer 7 (turn 14) announces that the requested feature has been developed and launched as part of the product. A screenshot of the new feature in a beta release of the product is shown in Figure 5. This version derives from the product blog empowered by Get Satisfaction and consists of two pictures representing two different ways the new feature has been integrated.

![Figure 5](image_url): A screenshot of a new feature integrated on the right hand side of the screen (Feedback Widget), here shown in a beta release of the product.

Taking into account the structural information provided by Developer 7 in turn 14, we have a richer understanding of this person’s role in the general development interaction pattern. Developer 7's degree centrality (3.206) indicate that he is a very central participant in the network with a lot of postings and/or responses to his postings. His high betweenness centrality (9.157), well above the average of 0.55, indicates that this person is one of the most powerful (top 5) participants in the network regarding the channeling and/or curtailing of information.

Two developers in this thread are powerful actors that contribute at different intervals (about three years apart) to inform participants about a planned feature in the beginning of the thread (Developer 11) and to conclude its realization at the end of the thread (Developer 7). Their actions, albeit brief, are significant for
realizing “general development.” Excerpt 3 thus illustrates software development as a collaborative effort between less technically skilled end users (customers) and technically skilled developers (employees), brokered by informed end users (champions).

Comparison to related work. When a champion in excerpt 3 requests a new feature and a developer responds after several years by stating that the feature has been developed and integrated in the product (“launched a Product follow feature”), end-user development (Fischer, 2009) appears as part of a bigger eco-system (Pollock et al, 2007). It shows a software artifact that has passed through a “double loop” modification process (outside → inside), incorporating both local adaptation (outside) and integration of the adaptation into the product (inside). This is an extension of an earlier version of the model of mutual development (Andersen & Mørch, 2009; Mørch & Andersen, 2010), now considering it from the perspective of mass collaboration. Progressing from a feature created by an end user to a new version of a commercial (GS) tool is tension laden and can be viewed as a strategy for developing products that embody characteristics relevant to many users, referred to as “generification work” (Pollock et al., 2007). Generification work has technical and managerial elements, and Pollock and colleagues focus mainly on the managerial elements, from a science and technology studies (STS) perspective (i.e. asking different research questions and employing different methods than we do).

The extended model of mutual development proposed here shares traits with the private-collective model of innovation suggested by von Hippel and von Krogh (2003). The two models share a focus on coordinating the activities of multiple stakeholders for the sake of collaboration for a common purpose, in our case further development of a software artifact. Examples of this can be seen in turn 10 where End user 44 elaborates on a solution he has created based on an unresolved feature request in the online community. The two models also share the integration of two levels of activity, in our case adaptation (end-user development) and generalization (professional development) (Andersen & Mørch, 2009; Mørch & Andersen, 2010). However, these two levels differ from those of the private-collective model; transitioning between the levels in our model involves an abstraction in order to integrate the end-user suggested features (redundancy removal, filtering out bad proposals, restructuring old code, etc.), whereas the private-collective model relates to the aggregation of individuals or companies (one to many innovators). The details of the abstraction process are outside the scope of the analysis we have described here and require further methods (e.g., interviews with developers, observation of decision making among developers, involvement of management in final decision making).

5.2.4 Excerpt 4: User-User Collaboration

In excerpt 4, the discussion centers on the collaboration of two end users in the online community to develop a local adaptation or end-user developed solution to a previously requested (and unresolved) feature. This excerpt is extracted from the middle of the same discussion thread as excerpt 3 (general development).

<table>
<thead>
<tr>
<th>Turn</th>
<th>Participant</th>
<th>Text from discussion thread</th>
<th>NrmDegree</th>
<th>NrmBetweenness</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>End user 49</td>
<td>Hi all I've been wanting this feature for a while now, and still do. However there is one important thing to bear in mind that I've learnt with our profile (blue ski). It relies on customers correctly categorizing topics in the first place, and that's if they bother categorizing at all, which many don't. So while it will still be a huge help, we can't rely on it completely and still need people monitoring and</td>
<td>1.371</td>
<td>0.00</td>
</tr>
</tbody>
</table>
This excerpt shows how two end users break away from the traditional pattern of “general development” due to disappointment that an awaited development has not been realized after several years. As a result, these two end users, 49 and 40, start to collaborate on a solution without involving developers, thus the name “user-user collaboration.” Characteristic of this type of interaction, it spans a short, intense time period (about a month), involving collaboration among a few skilled customers. In turn 2, End user 40 states that it is possible to “hard-code” the integration of discussions with a customer-specific product by using the embed code customization functionality. This is supplemented by an example from previous use of GS software and the user’s own product line. The two customers thus act as end-user developers and together collaborate on finding an ad hoc (business-specific) solution to the problem, and one of them, End user 49, implements it. The two customers solve the problem on their own by “hard coding tags” to connect discussions to the products they refer to. Although not an optimal long-term solution (poor software engineering practice), it solves the problem for End user 49.

End user 49 has a degree centrality of 1.371, which indicates that this person’s contributions are slightly fewer than the average contributor. The person may initiate several requests but be impatient about waiting for a developer-implemented solution and thus wish to create his own. The user’s betweenness centrality of 0 indicates he does not have the ability to exercise control over information flow. End user 40 is better positioned in the network to access information, being among the top 20 most active participants in terms of degree centrality (4.332, see Table 5) and participates in multiple threads (Figure 3). We interpret this to mean that End user 40 can provide useful advice to End user 49 based on past interactions in other discussion threads.

Comparison to related work. User-user collaboration can be compared to collaborative tailoring, (Kahler, 2001). Collaborative end-user tailoring grows out of one or more end users’ need for a feature and the happy coincidence of being in contact with another end user with a similar concern and the required technical skills to solve the problem collaboratively. The data in excerpt 4 shows how two end users collaborate to tailor a common product to fit one of their needs. When end users are empowered and able to collaborate on tailoring a

distributing topics

<table>
<thead>
<tr>
<th>Turn</th>
<th>End user</th>
<th>Text</th>
<th>Degree Centrality</th>
<th>Betweenness Centrality</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>End user 40</td>
<td>This may not help, depending on how you’ve set up your website, but you can hard-code a specific product using the embed code customization. Our “products” correlate to different websites across our properties, so we’re able to use the embed code very effectively to categorize support requests to the appropriate product. However once that is done of course…. we still have ALL our staff getting the notifications, so it’s highly frustrating.</td>
<td>4.332</td>
<td>0.864</td>
</tr>
<tr>
<td>3</td>
<td>End user 49</td>
<td>Our “products” aren’t that clean cut unfortunately as they revolve around the elements of a package holiday. Your suggestion got me thinking about another challenge I have though. Do you know if tags can be hard coded using the embed customization?</td>
<td>1.371</td>
<td>0.00</td>
</tr>
<tr>
<td>4</td>
<td>End user 40</td>
<td>I just looked at the widget customization screen, and yes it does look like you can hard code tags. Maybe that helps</td>
<td>4.332</td>
<td>0.864</td>
</tr>
<tr>
<td>5</td>
<td>End user 49</td>
<td>Brilliant! I always assumed that tag would filter the initial suggestions. Didn’t realize it pre-populated the new topic. Problem solved! Thanks :)</td>
<td>1.371</td>
<td>0.00</td>
</tr>
</tbody>
</table>
certain technical complexity, in this situation by “hard coding tags by using the embed code customization” (turn 2), this can be seen as an example of objectification (Kahler, 2001), meaning that the users or moderators of a software are able to modify or change aspects of it in use. Based on the interactions of these end users, we can infer that this is possible with the GS tools. Furthermore, as excerpt 4 is extracted from the middle of a large discussion thread, it is accessible to the entire support community. This has some similarities to Kahler’s (2001) emphasis on the importance of providing awareness of tailoring activities to others in the community. Despite similarities between user-user collaboration and collaborative tailoring, our work differs in that we have studied tailoring in a mass collaboration context rather than in small-group settings (Kahler, 2007). Moreover, we integrate user-user collaboration (a type of end-user development) with professional development (Andersen & Mørch, 2013).

6 Summary and Conclusions

The research question prompting our research is: “What are the patterns of interaction between end users and professional developers in a mass collaboration community, as seen from a mutual development perspective?” Our data show that in mutual development of the Get Satisfaction social media platform different constellations of, end users, champions, and developers emerge who together create mutually beneficial software in joint processes of mass collaboration. The processes we studied can be seen as examples of how human behavior on a large scale is expressed through a social media technology. Our main finding is the four patterns of mass collaboration in mutual development, representing different relationships or types of interactions between end users, champions, and developers. They were named as follows during the categorization of data:

- **Gatekeeping**: This refers to an actor controlling information by filtering and/or selectively choosing what information to pass on. A gatekeeper demonstrates power by controlling the flow of information in the network in two possible ways: 1) shielding other actors from unnecessary (e.g., obscure, distracting) information and 2) withholding important information (e.g., a suggestion by an end user might go unnoticed, despite having the potential to be novel and interesting).

- **Bridge building**: A bridge builder is a “gatekeeper” who distributes (rather than withholds) information. An advantage of this role is providing other actors with timely access to information in the network; a disadvantage is that information passed on by a bridge builder could generate information overload and worst-case be disrupting to other actors’ primary work.

- **General development**: This is when a local solution (hack, modification, work around, improvement request) proposed by an end user is brought to the attention of developers in the (GS) company and becomes a new feature of the “general product,” in effect making the GS tools available to all of the company’s customers.

- **User-user collaboration**: Customers sometimes need to make adaptations on their own (hacks, work arounds, end-user developed solutions, etc.) to fit the software to a new use situation. In these interactions, they do so by themselves or with the help of other end users (and champions) instead of waiting for professional developers to get involved, which typically takes longer.

The four patterns reveal the following identifiable traits, or shared characteristics, of mass collaboration in mutual development:
- Asymmetrical power relationships (between end users, champions, and professional developers)
- Mass collaboration in the community, initiated mainly by end users but sometimes by champions
- The importance of champions for propelling user driven innovations.
- Different time scales: Interactions can last several years (general development) or be short, intense periods of collaboration over weeks to a few months (user-user collaboration).

We used a mixed-methods approach to analyze our data on two levels: the network level using social network analysis (SNA) and the content level using interaction analysis (IA). This combination gave us both a broad picture and a more detailed look at mass collaboration in an online community. In total, this combination gave us a fuller understanding of mass collaboration that either method could by itself. When performing SNA analysis we draw on relational data, taking the participants contributions in the GS online community as a starting point for performing the analysis. We found SNA to be a useful analytical tool for identifying powerful actors, power relations and the structural attributes of the interaction patterns. However, this was not sufficient. Therefore, we used IA to study the content of the interactions, and together the two methods complemented each other and justify the naming of the interaction patterns.

Furthermore, bringing social network data and interaction data together gave us a richer picture of the flow of information than neither method could provide by itself. In particular, SNA has contributed in the following way to guide our interaction analysis: a) We used SNA to justify our choice of interaction data for content-level analysis (choosing discussion threads based on thread degree centrality) and b) We used SNA to “tag” the different actors’ utterances with network-level (socio-structural) information for informing our interaction analysis (in terms of who has power in the network, calculated by degree and betweenness values for each actor). We conclude that SNA is a valuable and useful tool for analyzing social structure in mass collaboration and especially valuable in combination with IA.

We extend previous research on mass collaboration in two ways: 1) we provide detailed descriptions of the distinct processes of mutual development (a specific form of mass collaboration in product development characterized by power and knowledge asymmetry), and 2) we show how the computation of two centrality measures (degree and betweenness) can identify powerful actors in a network and their different roles. This extends previous work on mass collaboration (Tapscott & Williams, 2007; Cress et al., 2013; Halatchliyski et al., 2014). In previous studies, authors did not distinguish between the multiple roles of participants, except for peers and teachers, nor the distribution of power among actors. Our case involved participants playing one of three different roles (end user, champion, developer). Their centrality measures indicate two types of power in the network in terms of information flow: selection power for high degree actors and control power for high betweenness actors. When we analyzed the emerging patterns of interactions it became apparent that patterns between end users, champions, and developers were different depending on the type of software development activity they engaged in (early in the process; late in the process, etc.). In sum, our social network data and interaction data taken together provides a rich picture of the social interactions and thereby strengthening the belief of the importance of actors’ ability to collaborate and determine information flow. Whereas the vast majority of participants were end users, the interaction patterns also involved champions and developers in different constellations with the end users. The following interaction patterns were identified:
• **Champions**: Champions were active in discussion threads with high degree centrality, which are the threads with many contributors. A reason for this could be that those threads are likely to attract the attention of the company because they have caught the interest of many customers. These threads may indicate discussions about functionality and the need for improvement in order for customers to continue to use GS tools. Because champions are engaged by the company to serve the customer community, they need to seek out threads that generate customer requests.

• **Champions vs. developers**: Champions participated mostly in bridge building activities by negotiating and coordinating information and technical innovations, whereas developers employed a more restrictive or administrative role, controlling information, which resembles gatekeeping.

• **Bridge building and gatekeeping**: Bridge builders were prominent among participants with high betweenness degrees, many of whom were champions (e.g., Champion 1 and Champion 2), whereas gatekeeping was associated with some developers (e.g., Developer 13 and Developer 19). Gatekeeping refers to participants controlling information by filtering and selectively choosing what to pass on. This is partly necessary and exercised by some of the developers in the community in order to prioritize the many proposals for improvements received by a relatively small group of full-time developers. The champions we studied took an opposite role to that of the developers: they spread information to other actors (developers and end users) thus helping to connect developer activity with the activities of end users.

Our findings have implications for user driven innovation by identifying champions (not only lead users) as originators of ideas. Lead users are customers (end users) who are well qualified to participate in proposing new ideas along the lines suggested by lead users in user driven innovation research (Von Hippel, 2005). We found that also champions were important originators of new ideas, and their point of reference is anchored twice: a) Champions are partly paid by the company (GS), which implies a commitment to this company (e.g., brand advocates), which may cause some of their contributions to be biased, and b) champions started out as ordinary end users, representing a customer organization like a lead user (Von Hippel, 2005) or a super user (Mørch, Hansen & Ludvigsen, 2007). Therefore, we cannot say that the innovations proposed by the participants in our study were entirely motivated by customer needs; it is better thought of as a compromise between two worlds: the needs of the developers and the needs of the customers.

Our findings also have implications for CSCL by proposing an extended model of collaborative knowledge creation in situations where:

• Peer collaboration is necessary but not sufficient for sustaining and directing mass collaboration toward productive interactions ➔ the emergence of champions as intermediaries between learners and educators

• A tangible and extensible (software) artifact is used as a shared object in collaborative knowledge creation ➔ integrating software artifacts (in our study represented as user interfaces) with discussions about the software.

Studying mass collaboration in the mutual development of a software product has implications for traditional product development processes and innovation studies. When participants who are geographically spread around the world are enabled with tools to collaborate online, join forces, share knowledge with the purpose of co-creating a product, this may lead to new forms of online collaboration. Along this line, our
findings indicate that all participants, including end users and champions, act as contributors, designers, and co-creators of the company’s software product. This demonstrates how some end users have changed from consumers of finished products to active and committed contributors to processes of end-user development (Fischer, 2009). These are processes of mutual development that are shaped by the needs, expectations, and skills of end users, champions, and professional developers using tools of various degree of complexity, allowing changes to be made at different levels of abstraction (from identifying problems to proposing solutions). The different social structures between the participants allowed us to identify four patterns of interaction in mass collaboration. However, further studies may be able to identify additional patterns and refine the patterns we have described. There is plenty of scope for more research in this area, as long as social media technologies continue to facilitate collaboration across boundaries of expertise, roles, and geography in online communities.

6.1 Limitations and Directions for Further Research

The limitations of our study include the following:

- Our data was collection from mass collaboration in the development of a single software product. Further work ought to study a different product development community or a different product mediated by GS tools to see if the same interaction patterns can be identified in order to extend, verify, or adapt this work.

- We have studied software product development as it co-evolves among end users, champions, and professional developers. This excludes managers. Managers have important decision-making functions in commercial companies and can impact the final product in many ways, for example, regarding the direction a product will develop in with the aim of increasing revenue. Our study was limited in this regard due to practical reasons. We have excluded those stakeholders who do not have an online profile in the community and who did not contribute to any of the forums. A manager might, for example, be able to exercise power and influence by tagging an idea as “not planned” and thereby shutting down an opportunity for this feature to be developed. However, those managers who also have developer status (e.g., technology managers) were included in the developer category.

- We selectively chose empirical data from two out of four types of discussion threads, “Share an idea” and “Give praise,” for practical reasons, specifically to limit the length of the write up. It could be interesting to supplement this study with data from the categories “Report a problem” and “Ask a question.”

- SNA analysis, conducted manually by the first author, was very time-consuming as relationship data had to be hand coded one-by-one into an Excel sheet for further processing by UCINET. This was partially automated through the use of the DNA software. It is technically possible to generate sociograms and SNA analyses more directly from a social networking site, but this needs to be balanced against the privacy rights of individuals.

Directions for future studies based on the work we have presented here include but are not limited to the following questions:

- What motivates different participants to contribute to and spend much of their (leisure, without pay) time improving products belonging to a company that may profit from their work?
- How can a company ensure that large crowds of users will generate quality input and not just noise and distraction?
• How can an evolving product be prevented from becoming over-specialized or feature excessive? Is it realistic that customers can continue to request improvements of a product over the duration of its lifetime and expect to be satisfied?

• How might managers use SNA+IA type information to improve the company or organization? Understanding how customers see the product and how champions and developers solve issues could help improve fundamental business practices.

Acknowledgements
We wish to thank the members of the Mediate Research Group at Department of Education, University of Oslo for detailed and constructive feedback throughout the process of writing this article. Special thanks go to Sten Ludvigsen (University of Oslo) and Yvonne Dittrich (IT University of Copenhagen) for useful comments on a previous version of the article. We thank the anonymous reviewers for their valuable comments on our manuscript. The first author received financial support by a PhD fellowship from the Faculty of Educational Sciences, University of Oslo.

References


