

INCREASING THE WILLINGNESS TO COLLABORATE ONLINE: AN ANALYSIS OF SENTIMENT-DRIVEN INTERACTIONS IN PEER CONTENT PRODUCTION

Completed Research Paper

Mihai Grigore

Goethe University
Frankfurt am Main, Germany
grigore@wiwi.uni-frankfurt.de

Christoph Rosenkranz

Goethe University
Frankfurt am Main, Germany
rosenkranz@wiwi.uni-frankfurt.de

Abstract

We investigate mechanisms that trigger collaborative work behavior in online peer communities. We regard the collaboration among Wikipedia editors as a social process influenced by specific communication practices. We analyze and quantify the way Wikipedia editors communicate their feedback and support towards each others' work in form of sentiments and opinions, and explore to what extent this influences online trust among them. We show that peer content production in Wikipedia is influenced by sharing sentiments during discussions among editors. At the global level, sharing sentiments positively influences the level of online trust. We also find a significant difference in the amount of online trust among editors who share mainly positive or mainly negative sentiments. We further suggest that providing and receiving especially supportive feedback expressed in form of positive sentiments and opinions may be beneficial in terms of virtual teamwork.

Keywords: Online Collaboration, Peer Content Production, Affective Communication, Virtual Teamwork, Social Interaction, Online Trust

Introduction

Online communities are a key topic of interest for both researchers and practitioners, especially since Hagel and Armstrong (1997) claimed that commercial success in the online area belongs to those businesses that organize electronic communities. Online users increasingly not only consume the Internet to seek information, but also use it as a virtual communication and collaboration platform. These interactive and collaborative mechanisms have found their way into scientific discourse and discussion mostly under the umbrella term “Web 2.0” (Vossen and Hagemann 2007). Web 2.0 is usually associated with technologies that facilitate interactive information sharing, interoperability, and collaboration on the World Wide Web, leading to the development of social networks and social media (Musser and O'Reilly 2007). The main building blocks of Web 2.0 are principles and technologies that allow an interactive and user-oriented design of Internet applications. Blogs (web logs), wikis, or podcasts allow users to easily create content on their own, to discuss issues of interest, to express opinions, facts, or to share information with others without needing to understand the underlying technologies (Vossen and Hagemann 2007). Web 2.0 further allows various levels of user involvement, as well as unique and unprecedented opportunities for engaging users into mass collaboration.

The emergence of these information and communication technologies fundamentally enables geographically dispersed people to come together with little cost, exchange ideas, and coordinate their activities (Leimeister et al. 2006; Preece 2000; Preece 2001; Rheingold 1993). This phenomenon has led to an increase in the number of online communities that focus on coordinated efforts of volunteers to produce intellectual work – so-called *peer production* (Benkler 2006; Benkler and Nissenbaum 2006). This raises a series of challenges for traditional organization theory, as social peer production systems organize creative mass collaboration through coordination without relying either on markets or on managerial hierarchies to organize decentralized production (Benkler 2006; Stephen and Suzanne 2006). Peer *content* production further offers major advantages over markets and managerial hierarchies in terms of resource allocation and information processing (Andreev et al. 2010; Benkler 2002; Taddeo and Vaccaro 2011). Wikipedia¹, the online encyclopedia “that anyone can edit”, is a prominent example of peer content production, where voluntary contributors coordinate their work and develop mutual understanding on the collaboratively created content of encyclopedic articles (Kittur and Kraut 2010; Tkacz 2010).

Recent years have seen increasing research efforts on investigating how community design affects user behavior in online communities and social networks (Crandall et al. 2008; Ren et al. 2009; Ren et al. 2007; Wang et al. 2009). The effects of social interactions on longitudinal user behavior, however, are barely explored, and neither are the effects of such interactions on user collaboration in online communities and peer production entirely understood. Mass collaboration works surprisingly well, despite its lack of mechanisms of formal authority (Eseryel 2009). However, a major recurring problem for providers of online communities is that a significant part of community members do not contribute with content. That is, even when online platforms provide appropriate tools for collaboration, a large number of their user communities end up being underdeveloped simply because of both non-participative and non-collaborative user behavior (Ling et al. 2005; Maloney-Krichmar and Preece 2005; Susan et al. 2005). Despite the large body of literature on how to build viable online communities (Farzan et al. 2011; Gurzick and Lutters 2009; Lin 2008; Maher et al. 2011; Rosenkranz and Feddersen 2010), we still have no deep understanding of the underlying factors driving online collaboration and peer production. Traditional incentivizing methods are usually unsuitable for such online environments (Bishop 2007). Non-participative online behavior might also be a significant issue for the trend to increasingly “crowdsource” tasks to large masses of workers, for example, by using micro-task markets such as Amazon’s Mechanical Turk² (Downs et al. 2010; Kittur et al. 2008; Ross et al. 2010), where there is a constant need to convince users to provide content for your service in case one either has limited or cannot provide monetary incentives.

The aim of this paper is to provide a novel theoretical perspective on understanding the mechanisms driving productive social interactions online with almost no centralized control. We seek to identify

¹ http://en.wikipedia.org/wiki/Main_Page

² <https://www.mturk.com/mturk>

factors that motivate users to contribute, and that positively influence online collaborative work. For doing this, we analyze and discuss peer content collaboration in the particular case of Wikipedia article editing. Our study regards the collaboration among Wikipedia editors as a *social process* influenced by specific *communication practices*. More concretely, we investigate whether affective communication (Te'eni 2001) among Wikipedia editors acts as an intrinsic motivator of collaborative work behavior. We analyze how users communicate their feedback and support towards each others' work in form of sentiments and opinions expressed during conversations, and explore to what extent peer feedback influences online trust among Wikipedia editors. Our aim is to provide insights towards what makes large-scale collaboration and peer content production in online communities sustainable and productive without relying either on markets or hierarchies (Metiu and Kogut 2001; Stephen and Suzanne 2006).

We discuss the related work and the theoretical background of our research in the next section. Afterwards, we describe and discuss our research method. The dataset used in this study consists of a complete revision history dump of Simple English Wikipedia.³ We employ methods from opinion mining and sentiment analysis to our data. First, we find that sentiments expressed in inter-editor conversations correspond to an increased level of online trust of co-editors on the corresponding article pages. Second, we show a significant difference in the level of online trust between editors who share mainly positive or negative impressions respectively. Finally, we show that reaching mature states of article content is not necessarily linked to the strength of sentiments expressed as feedback in Wikipedia discussions. We discuss our results and conclude with an outlook on further research, as well as on the implications of our study.

Background and Related Work

User Contributions in Online Communities: The “Lurker Problem”

Most traditional research on online communities focuses on the role of information technology in enabling people to interact with each other (Burnett 2000; Preece 2001), on the factors of usability and sociability (Preece 2001; Tarmizi and Vreede 2005), as well as on what principles make online communities successful. For example, Porra and Parks (2006) suggest that the sustainability of online communities requires persistent people, continuous support by an online space, and flexibility for alternative sub-communities to emerge. Ginsburg and Weisband (2002) conclude from their survey that volunteerism is an important aspect for the success of online communities.

Although online communities are becoming increasingly relevant to business (Parameswaran and Whinston 2007), many of them fail and participation drops to zero (Ling et al. 2005). Regardless of the size of online communities, it has been observed that only a small fraction of registered users actively contribute to the process of creating and managing content. In the context of online environments, “lurkers” are usually perceived as users who consume the available information, without further contributing with any content (Gensollen et al. 2007). It is an important and difficult challenge to design technical features of online communities and seed their social practices in a way that generates ongoing contributions from a larger fraction of the participants (Ling et al. 2005). In this context, Ren et al. (2007) developed a theoretical framework to distinguish identity-based from bond-based attachment to online communities. While common identity refers to users who appreciate the community as a whole, common bond attachment refers to users who appreciate other individuals in the group. Fiedler and Sarstedt (2010) performed an evaluation of the influence of social interaction on both common identity and common bond attachment in online communities. They provided valuable insights into the complex relationships underlying user behavior in online communities by showing that both network effects and collectivism influence, among others, membership robustness, and loyalty to the online community.

Designers of online communities usually face the problem of determining what makes lurkers become active contributors, as previous efforts on understanding why lurkers do not contribute (Maloney-Krichmar and Preece 2005; Susan et al. 2005) do not provide sufficient insights for affecting an established lurking behavior. For example, Rafaeli and Raban (2005) stress the importance of passive reading as a likely precondition for both participating and building a sense of online community. Antin

³ <http://simple.wikipedia.org>

and Cheshire (2010) investigate the way readers learn to get involved in the Wikipedia community, and propose theoretical insights for re-casting lurkers as more valuable participants in online environments. They claim that research on how different categories of users pay attention to the content may reveal insights about particular users' attitudes, behaviors, and intentions. In a study investigating habit formation in online communities, Gan et al. (2009) find that there exist both *intrinsic* and *extrinsic* motivations influencing user participation. Preece et al. (2004) reveal that a significant proportion of lurkers do not belong to the category of self-interested individuals taking advantage of others' work. Lurkers may even be willing to contribute, but are usually hold back by already developed beliefs and values (Bishop 2005). In fact, Preece et al. (2004) identify five important reasons why lurkers do not contribute in any way to the community: there are both no needs and no encouragements to post, lurkers needed to know more about the group before getting involved, they think they are being helpful by being altruistic lurkers, or they are simply not able to use the software functionalities.

To understand the motivation behind technology-mediated social cooperation, Preece and Shneiderman (2009) propose the "Reader-to-Leader" framework. As prior studies (Shah 2006) report that once their needs are met, users are more likely to leave the community than continuing to contribute, Preece and Shneiderman (2009) claim that getting users to revisit online social environments may represent an important premise for users to start contributing content and possibly becoming collaborators. On the one hand, Bishop (2007) recommends that, in order to successfully challenge lurking behavior, community providers should at first attempt to change lurker beliefs by using persuasive discourse, that is, by countering the beliefs of actors and providing them with new information (Chambliss and Garner 1996). On the other hand, Murphy et al. (2003) argue whether using persuasive texts represents an effective way of changing lurker beliefs, and, consequently, lurking behavior. Although challenging these predefined lurker beliefs may generate an increased individual participation in online communities, Bishop (2007) points out that none of the previous theoretical attempts (suggesting that participants are either goal-driven or need-driven) are entirely appropriate for understanding lurking behavior in such settings. Moreover, further empirical studies on the progression from reading to other forms of user participation in online communities are needed (Antin and Cheshire 2010).

Online Collaborative Work: The Role of Affectivity and Trust in Online Settings

What drives people to collaborate in an online setting, becoming active co-workers? Collaboration is a process involving "two or more contributors discussing, cooperating, and working together to create something or share information" (Denning and Yaholkovsky 2008). It takes coordination for people to perform things together and it takes communication to achieve that coordination (Clark 2005). The basis for communication and coordination is shared knowledge, or common ground, between people (Clark 1996, p. 120). Developing mutual understanding and shared beliefs (i.e., a common ground) is an essential aspect of collaboration (Convertino et al. 2008; Convertino et al. 2009). Moreover, participation in any collaborative activity requires a unique motivation to share psychological states with others (Tomasello et al. 2005). Convertino et al. (2009) propose two types of common ground: process and content. Process common ground encompasses "I know that you know that I know how"; content common ground includes "I know that you know that I know what" (Convertino et al. 2009). *Process coordination* implies continuous communication of shared rules, timing, conventions, and manner in which the interaction will be conducted (procedural and strategic knowledge). In contrast, *content coordination* requires exchanging content and mutually checking and signaling understanding to establish a grounded content as result of collaborative work.

Sproull and Kiesler (1991) stretch the importance of developing supportive interpersonal relationships online, especially since the rise of the Internet. In an online setting, the focus of attention changes from the relationship between a person and technology to the relationship between a person and other people; people who never physically meet or know each other get to communicate and work collectively. Especially *trust* and *empathy* are found to be essential in encouraging people to work together online (Leimeister et al. 2006; Maloney-Krichmar and Preece 2005; Skopik et al. 2009).

Trust building increases members' willingness to collaborate, while shared values and the satisfaction with previous user interactions have a positive impact both on trust online and on returning behavior of virtual community members (Wu et al. 2010). Thus, one may count as beneficial the willingness to suspend doubt about others, and start working on group tasks with a positive expectation that the group

activity will be completed (Meyerson et al. 1996). Moreover, research on trust in virtual teams suggests that the initial willingness to show trusting actions leads swiftly to actual trust, and that frequent communication between team members helps to promote trust among them (Wallace 2001). The probability of engaging in trusting behavior is further likely to increase as a result of the “anticipated positive and negative motivational consequences” (Deutsch 1958).

Empathy has been shown to have a significant influence on trust in online settings (Feng et al. 2004). Communication partners who interact in an empathically accurate and supportive way appear to be most trusted by other online participants (Feng et al. 2004). Empathic accuracy itself does not guarantee trust, and, in order to win other’s trust online, it is not nearly enough to correctly infer the other’s feeling, but also to provide *supportive feedback*. Eseryel (2009) further discusses how the lack of formal authority in online communities is efficiently compensated by individuals who *mentor* and *encourage* each other to contribute to the team. Shared affect, that is, subjective experience of feeling - sentiment or emotion - as reaction to things one thinks about, actions one takes, or to various stimuli (VandenBos 2006), helps coordinating group activity through fostering group bonds and group loyalty (Spoor and Kelly 2004).

Moreover, *affective communication* among group members plays a key role in the articulation of sentiments, emotions, or moods both in spoken discourse and in written text. Affective information can be transferred through computer-mediated communication, while message receivers are able to successfully decode affective information (Harris and Paradise 2007). In a recent preliminary study, Lee et al. (2010) investigate whether emotional expressions in threaded discussions influence the quality of knowledge in online communities. More precisely, they argue that emotional expression in inter-user communication is a trigger for sharing experiences in such communities, and they propose that the degree of emotional expression is positively associated with the degree of sharing experiences, the degree of sustaining dialogues, and with the quality of knowledge created.

In the specific case of peer content production in Wikipedia, an analysis of collaboration patterns occurring during article editing revealed that an intensification in collaboration usually occurs after the initiation of conversations among editors (Crandall et al. 2008). Peer influence exerted across social ties among editors can be further used as predictor of future editing behavior (Crandall et al. 2008). Social capital (Shah et al. 2001) represents both a cause and an effect of social selection in Wikipedia, in the sense that articles that reach faster a higher quality level appear to be created by groups of editors who have previously worked together on other articles (Nemoto et al. 2011). This higher pre-existing social capital of co-workers is reported to be especially important in the early phases of Wikipedia article definition and team organization. Once the general direction of the article is set, the team then appears to be able to absorb new contributors more effectively (Nemoto et al. 2011).

To sum up, users who tend to interact and communicate in a supportive way are most likely to be trusted by other online participants. Both pre-existing social capital and supportive feedback among participants appear to have a positive impact both on building trust online, as well as on returning behavior of virtual community members. Trust building has further been shown to increase members’ willingness to collaborate in online settings. Moreover, Convertino et al. (2009) stress the differences between the process of sharing knowledge that supports common ground in conversations (content coordination), on the one hand, and the process of building common ground in the context of complex team activities (process coordination), on the other hand. In the context of teamwork in Wikipedia, we regard the peer content collaboration (i. e., article editing) as a process of content coordination. We further see the communication among Wikipedia contributors as one means by which editors plan editing actions, and develop a shared understanding of the edit policies, procedures, and timing which are meant to guide their interactions (i. e. process coordination). We aim to identify factors that lead to building a sense of trust among Wikipedia editors, and that further drive them towards collaborating on article content.

As a first step, we propose to investigate whether affective communication practices among Wikipedia editors positively affect the level of trust online, and act as intrinsic motivators of peer content collaboration. More precisely, we focus our investigations on how Wikipedia editors communicate their feedback and support towards each others’ work in form of sentiments and opinions expressed in inter-editor discussions, and to further explore to what extent this impacts the collaborative article editing process in Wikipedia. We analyze the level of subjectivity (or, more specifically, the level of sentiments, cf. section “Research Model”) in inter-user communication with regard both to the level of trust between users and to the user collaboration in content creation. For doing so, we propose to integrate the analysis

of specific peer content collaboration - the editing process of Wikipedia articles - with an analysis of informal discourse - the level of sentiments in discussions between Wikipedia editors. This leads us to formulating a general research question:

GRQ: What enables a foundation of trust among peer content contributors? Do peers mentor and encourage others to contribute?

From this, we draw two subsequent, more focused research questions we address in this study:

RQ1: Is online trust in Wikipedia affected by providing feedback about editors' contributions? If so, does sharing supportive feedback increase promoting a higher level of online trust?

RQ2: To what extent is sharing feedback on others' work beneficial in terms of collaboration effectiveness in Wikipedia?

Research Design and Method

The Data Set

The online encyclopedia Wikipedia uses a wiki technology to support collaborative writing and editing of textual content by its own community of readers. The resulting *article pages* represent the main source of contents used by regular Wikipedia readers. In order for any user to visualize the dynamics of article changes, the *revision history* functionality provides a track of chronologically ordered summaries about the previous versions of the article such as time stamp, editor, actual text resulting from the edit, and comments.

Besides the underlying wiki-based collaboration mechanism, Wikipedia contributors communicate among each others using specific mechanisms such as *talk pages* and *user talk pages*. While performing edits on article entries, Wikipedia editors tend to initiate conversations on article talk pages in order to plan and discuss their work, that is, coordinate their work, share and ask for feedback, report vandalism, or refer to edit guidelines (Schneider et al. 2010). Examples of both supportive and negative editor feedback on Wikipedia's "Talk:NASA" and "Talk:Water_on_Mars" entries are shown below:

*"Many Modules ! Brilliant ! Precise ! Do you know, looking at the changes log, I think the guardians of this page are overworked and under-appreciated. I do hope I can be of assistance wherever possible. Keep up the great work ! Don't give up !Penyulap (talk) 01:08, 18 March 2011 (UTC)"*⁴

*"Not only is it a bit lengthy, but it has very unusual organization for a wikipedia page... perhaps inappropriately so. Also, in some places it isn't written with proper encyclopedic style [...] I hate to merely be a critic, but I'm not nearly qualified enough to attempt rewriting or reorganizing this article. :) The2crowrox (talk) 00:58, 5 November 2010 (UTC)"*⁵

Because of the extremely large size of Wikipedia's revision history and the limited computational power, many previous analyses used only samples of data in order to save computation costs (Arazy and Nov 2010; Javanmardi and Lopes; Muller-Birn et al. 2009; Viegas et al. 2007). However, as there are no general guidelines on how to obtain a good sample from Wikipedia, and since complete revision histories are necessary for computing revision-based metrics, we instead use a complete revision history dump of Simple English Wikipedia, a spin-off of Wikipedia written using basic English vocabulary and uncomplicated grammar constructions. Articles from Simple English Wikipedia are usually not new; their editors use articles from Wikipedia and try to bring them to a simple form. The dump was created in March 2011 and contains over 200.000 pages, out of which around 70.000 are article pages, totaling approximately 16 gigabytes of XML data. Along with anonymous users, there are over 170.000 registered users who contributed to approximately 3 million revisions from the creation of Simple English Wikipedia. The average number of revisions per page is 14.35. Currently, there are over 700 registered users with at least one edit or logged action in the past month.

⁴ retrieved from: <http://en.wikipedia.org/wiki/Talk:NASA>

⁵ retrieved from: http://en.wikipedia.org/wiki/Talk:Water_on_Mars

Research Model

To explore factors that shape socially collaborative behavior, we propose to investigate the role of emotions as perceived feedback and as motivational factor of collaboration in Simple English Wikipedia. We argue that trust-implicating actions of Wikipedia editors help to establish or reinforce an emotional sense of trust and commitment, since positive affect usually circulates among those who express their trust behaviorally, just as negative affect arises among those who ‘betray’ or act distrustfully toward each other (Boyle and Bonacich 1970; Lewis and Weigert 1985). Since emotions are rather transient states of subjectivity and may rapidly decay to the neutral state (Barrett and Russell 1999), we suggest to focus on sentiments and opinions. We propose that those may be more useful for explaining collective behavior, as they, in contrast to emotions, more often become externalized instances on the collective level (such as “public opinion” or “collective sentiment”) (Scherer et al. 2004). We further use the umbrella term *subjectivity* to refer to both sentiments and opinions, which are, in fact, commonly addressed together in the specific literature of the opinion mining and sentiment analysis research field (Pang and Lee 2008; Wiebe and Mihalcea 2006). We observe whether the amount of subjectivity (i. e., the linguistic expression of private states such as opinions, sentiments, or emotions on talk pages, Banea et al. 2008; Furuse et al. 2007; Wilson et al. 2005) is related to building a grounded content (i. e., article page content) as result of collaborative editing. Specifically, we formulate the following hypotheses:

- H1: The total amount of sentiments on talk pages (whether positive or negative) positively influences the trust between editors on corresponding article pages. (RQ1)*
- H2: Positive sentiments during discussions on talk pages are associated with increased trust between editors on corresponding article pages. (RQ1)*
- H3: Semantic convergence of content on article pages corresponds to an increased amount of sentiments on corresponding talk pages. (RQ2)*

Research Method

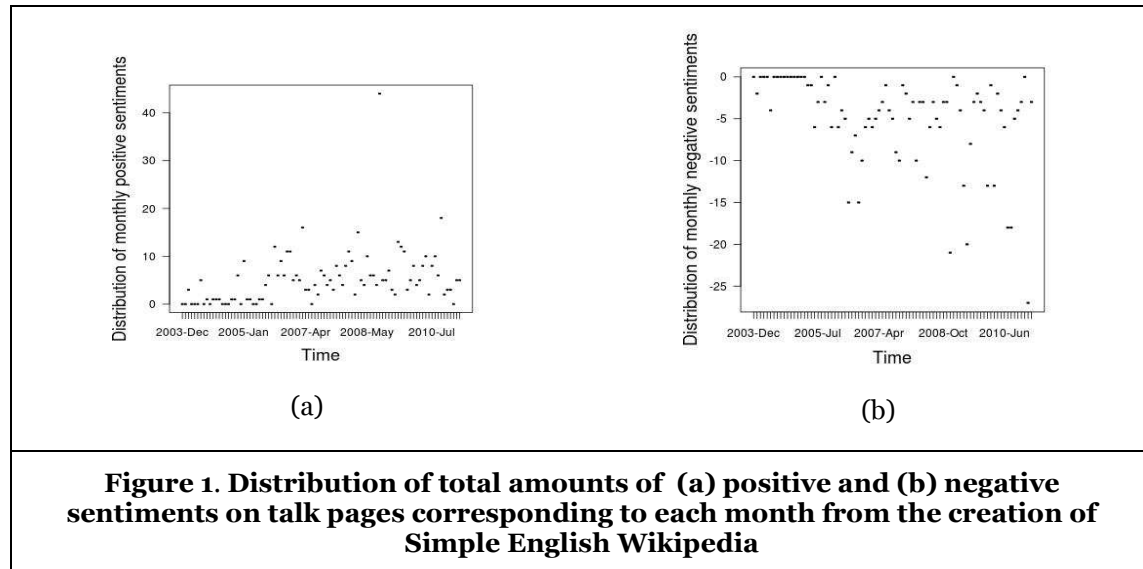
We compute Wikipedia-specific, revision-based measures on the editing process of article pages (trust and content convergence), as well as mining the sentiments expressed on talk pages. To address the above hypotheses, we employ monthly time series analyses of: 1) the user trust that is mediated by the amount of sentiments in inter-user communication, and 2) the level of subjectivity (in terms of sentiments expressed on talk pages) as a measure of coordination effort that is mediated by the convergence of article contents. Below we present each of the measures we use in our analyses.

Sentiment Analysis of Editor Interactions

Subjectivity analysis broadly classifies textual statements into objective statements that express factual information and into subjective statements that reflect holder’s attitudes or perceptions (Banea et al. 2008; Furuse et al. 2007). Opinion mining or sentiment analysis represents a systematic, computer-based analysis of written text or speech excerpts, for extracting the attitude of the author or speaker about a specific topic. Sentiment analysis provides a more fine-grained examination compared to subjectivity analysis, aiming to establish the overall orientation (positive or negative) and intensity (weak or strong) of the opinions or sentiments expressed by statements previously classified as subjective (Pang and Lee 2008). However, subjectivity may often be expressed in a subtle manner, making subjectivity analysis often “more difficult than subsequent polarity classification, so improvements in subjectivity classification promise to positively impact sentiment classification” (Mihalcea 2007). Moreover, Wikipedia editors often use informal vocabulary when writing comments on talk pages, which adds difficulties to the task of identifying subjectivity.

Recent sentiment analysis algorithms are able to detect positive and negative sentiment strength in short informal texts with a reasonable degree of success (Akkaya et al. 2009; Paltoglou and Thelwall 2010; Shanmugasundaram et al. 2009). In our analyses, we use the SentiStrength tool (Thelwall et al. 2011; Thelwall et al. 2010) to analyze the level of sentiments on talk pages in Wikipedia. SentiStrength provides a scoring range from -5 (very negative) to +5 (very positive). Also, in case of texts showing an equal amount of positive and negative sentiments, the algorithm is able to predict which of the two orientations

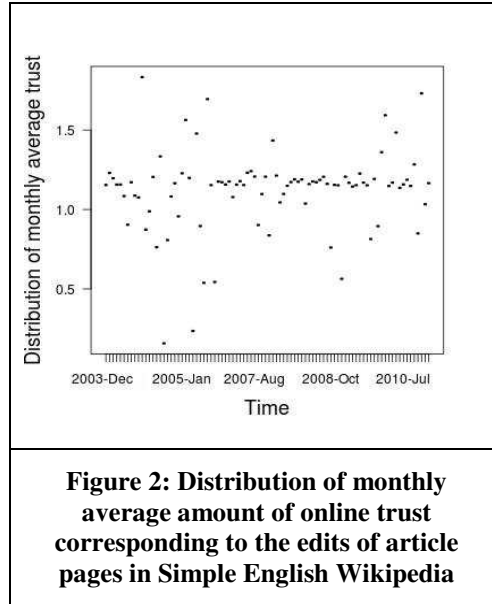
is the prevalent one. Figure 1 shows the distribution of the total monthly amount of sentiments expressed by editors in talk pages.



Online Trust as a Measure of Editor Team Quality

Williamson (1993) considers trust as one's subjective likelihood to engage in collaboration with agents who are perceived as potentially performing "beneficial or at least not detrimental" actions. Trust can in general be conceptualized as "a latent variable resulting from distinct but related (formative) indicators (i. e., propensity to trust and perceived trustworthiness), which lead to (reflective) indicators (i. e., behaviors of cooperation and monitoring between team members)" (Costa and Anderson 2010). In the case of virtual teams, perceived trustworthiness has been shown to accelerate the formation of trust and leads to more solid forms of trust in virtual teams (Hwang et al. 2004). The initial willingness to show trusting actions has been found to lead swiftly to actual trust, and frequent communication between team members helps to promote and gain others' *online trust* (Wallace 2001). Fen et al. (2004) further observe that communication partners who are willing to share feedback and interact in a supportive way are most trusted by other online participants.

Jøsang (2007) elaborates editor's trust towards the correctness of an entry in Wikipedia as perceived reliability of the information provided by other editors. Similar to Turek et al. (2010), we compute the amount of online trust established between two Wikipedia editors of the same article page as the number of words produced by the first author and reused (e. g., copied, moved elsewhere, or restored) by the second one. This indicates a notion of online trust in the sense of "a willingness to show trusting actions" (Wallace 2001). For a given Wikipedia article page, we compute the overall online trust between its editors as a mean of pairwise trust values, weighted by the proportion of editors who contributed to the creation of each specific revision from the total number of article editors. In this way, we reduce the effect of large number of edits on the same page. Figure 2 shows the distribution of the monthly average amount of online trust extracted from the edit logs of article pages.



Maturity of Article Pages and Semantic Convergence of Content

Christopher and Amit (2007) define the semantic convergence of Wikipedia articles as a state in which the content of the article remains stable despite the ongoing edits. More precisely, a document is considered semantically stable if a high semantic similarity between the current version and the previous k revisions can be established.

To measure the semantic similarity between two article revisions, we use the vector-space model for representing the semantics of the content of those revisions (Erk and Pado 2008). For each pair of vector-space representations of edit revisions, we compute their cosine similarity score as the angle between two vectors which represents the inner products of those vectors, the lengths being normalized to the unit (Turney and Pantel 2010). The value of the cosine ranges from -1 (the angle is 180 degrees and the vectors point in opposite directions) to +1 (the angle is 0 and the vectors point in the same direction). We consider a Wikipedia article as mature if its content converges over the series of edits.

In the next section we show the way we employ the three measurements (sentiments, trust, and convergence) in our analysis.

Analysis and Results

To address hypotheses H1 and H2, we examine both the distribution of sentiments on article pages and the trust between editors on the corresponding talk pages to accomplish two related tasks. First, we want to discover whether there is any significant difference between the presence or the absence of sentiments in the talk pages, and the amount of trust between editors. Second, we check if we can observe any discrepancy in terms of trust levels corresponding to positive and negative orientations of sentiments.

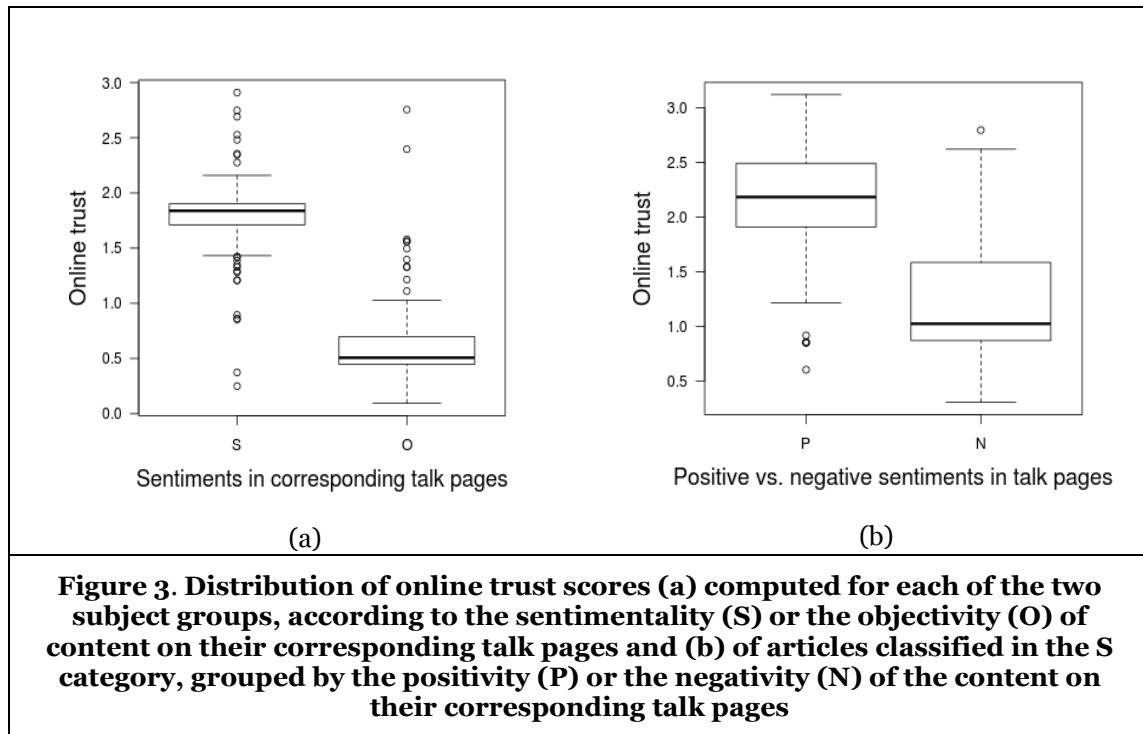
In order to prepare the data for tests, we first perform a time series analysis of all article pages with respect to the trust of their editors. We select all talk pages and store the time stamps of their edits. For those talk pages having an overall number of revisions greater than 30 (approximately double of the average amount of edits per article), we compute the distribution of sentiments over time on their corresponding article pages. We then attempt to discover to what extent sentiments on talk pages trigger collaborative behavior on the corresponding article pages.

To test the third hypothesis, we similarly perform a time series analysis of all talk pages with respect to the sentiments expressed by editors. This time we select all the article pages having an overall number of revisions greater than 30 and divide them into two clusters, whether their content is convergent or not. We aim at discovering whether an increased amount of sentiments on the talk pages corresponds to semantic convergence of content on article pages.

Following our analyses of Wikipedia article pages and talk pages, we group article pages by the following criteria:

- *presence or absence of sentiments*: article pages containing sentimental (S) statements on the corresponding talk pages are grouped in an S-cluster (1239 article pages), while the ones containing only objective (O) statements are grouped in an O-cluster (742 article pages);
- *positive versus negative sentiments*: we group each article page belonging to the S-cluster into either a P-cluster, if the main orientation of sentiments expressed in the corresponding talk page is positive (P, 794 articles), or an N-cluster, if the orientation is mainly negative (N, 445 articles); in order to decide upon the main subjectivity orientation, for all the statements on a talk page, we compared the sum of positive sentiment strengths with the corresponding negative one; the orientation was decided by the category corresponding to the higher sum;
- *convergent versus non-convergent content*: article pages whose content has reached a semantically stable state and are thus considered mature are grouped into the C-cluster (1084 article pages), while the others are grouped under the NC-cluster (897 article pages).

To quantify whether the presence or the absence of sentiments in the content of talk pages influences the trust between editors on article pages, we compute monthly averages of trust for each S- and O-clusters according to their article pages' revision flows. Results are shown in Figure 3 (a). We first compare the obtained discrete distributions of trust scores using the Mann-Whitney U test. The median level of trust computed for the O-cluster is significantly lower than the median corresponding to the S-cluster ($Z=-10.26$, $p < 0.001$, $r=0.76$). A linear regression reveals that the presence and the absence of sentiments significantly predict the level of online trust (adj. $R^2 = 0.6388$, $p < 0.001$).



For the articles contained in the S-cluster, we perform a second U test relative to mainly positive (P) and mainly negative (N) sentiment orientation of subjective content on talk pages. We obtain that the median trust for the article pages of the P-cluster is significantly higher than the one corresponding to the N-cluster ($Z=-8.61$, $p < 0.001$, $r=0.64$). The results are displayed in Figure 3 (b). The mainly positive or mainly negative orientation of sentiments also explains a significant proportion of variance in the level of online trust using a linear regression (adj. $R^2 = 0.4284$, $p < 0.001$).

However, we did not find a significant difference in terms of trust among editors between convergent (C-cluster) and non-convergent (NC-cluster) article pages (Mann-Whitney U test, $p=0.17$).

We summarize the results drawn from our data set with regard to our hypotheses:

Sentiments versus level of online trust:

- H1: There is strong statistical evidence that sharing sentimental statements (either positive or negative) on talk pages positively influences the online trust between editors on article pages.
- H2: There is strong statistical evidence that positive sentiment orientation during discussions on talk pages is associated with increased online trust between editors on article pages.

Content convergence versus amount of sentiments:

- H3: There is no statistical evidence that semantic convergence of the content of article pages corresponds to an increased amount of sentiments on talk pages.

On the one hand, our confirmation of H1 and H2 let us conclude that sentiments in discussions on talk pages positively influence the perception of online trustworthiness among Wikipedia editors. On the other hand, a high amount of sentiments does not necessarily imply maturity of content on article pages (rejection of H3).

Discussion

We described and analyzed peer content collaboration in Wikipedia using measures for sentiment strength, online trust between editors, and maturity of content based on the revision history of Simple English Wikipedia articles. Thereby we provided a first building block for research on how to effectively motivate users to collaborate in online peer production communities. We exemplarily studied the way Wikipedia editors give feedback, mentor, encourage, or criticize other editors by means of positive or negative statements on talk pages. We tested three specific hypotheses concerning user collaboration. We found that peer content collaboration in Wikipedia (in terms of higher levels of online trust among editors) is influenced by peer feedback in form of sharing sentiments during inter-editor discussions (H1). We showed that, at the global level, sharing sentiments positively influences the level of online trust. Specifically, Wikipedia talk pages where editors express their affective implication present an increased level of online trust among editors, reflected in the editing of article pages. We also found a significant difference in the amount of trust between editors who share mainly positive or mainly negative sentiments (H2). We conclude that reaching mature states of content on article pages does not particularly imply increased amounts of sentiments (seen as feedback between editors) shared on talk pages (H3). Our results extend the approach of Wierzbicki et al. (2010), who considered online trust among editors as a relevant aspect of peer content collaboration in Wikipedia. We further suggest that receiving (especially positive) feedback in form of sentiments expressed in inter-editor conversations may be beneficial in terms of virtual teamwork.

To our knowledge, we are the first to examine the relationship of the semantic information (provided by the content on Wikipedia article pages via sentiment analysis) with respect to revision-based measures (provided by trust online and maturity of article content). Our findings let us conclude that the approach we considered in this study provides significant insights for researchers looking for factors that influence productive social behavior. We therefore provide a suitable means for analyzing the social interactions in collaborative online communities. Our concepts also create awareness for possible deficiencies and pitfalls. For example, we draw from our findings that it might be very useful to strive for early communication activities on talk pages that include sentimental statements and provide a motivation for other users. This should create trust earlier and may further reduce transaction costs by creating this trust, specifically the costs of cooperation and specialization (Hill 1995).

Wikipedia further provides an unprecedented amount of data suitable for mining practices in computer-based collaborative work. The data set we used in our analyses enabled us to: 1) fully use the information provided by the edit history of highly edited articles (we perform no sampling over the time dimension) to employ revision-based measurements, and 2) exploit the multitude of informal language phenomena (specific to user-generated content) in order to identify subjectivity in the content of talk pages. Additionally, Wikipedia represents an ideal environment for studying the cumulative effect of social and affective interactions among editors on collaborative work. In fact, recent investigations found social interaction to be both a cause and an effect of social selection in Wikipedia: articles that reach faster a

higher quality level appear to be created by groups of editors who have previously worked together on other articles (Nemoto et al. 2011). In a more general setting, our findings may illustrate that the need for explicit coordination does not necessarily come along with the maturity stage of collaboration. We propose that we need to further compare individual profiles (e. g., activity patterns, coordination efforts, or communication activities) of users whose group collaboration reaches mature stages in order to distinguish whether affective implication is a matter of individual work habits or if it indeed comes along as a consequence of social interaction among users. In this sense, Aral et al. (2009) further stretch the increasing importance of clearly separating social influence from homophily in order to be able both to understand how behavior spreads (i. e., what enables social contagion, and what makes behaviors propagate viral), as well as to effectively design and support policies that encourage or combat the spread of specific behaviors in online communities.

Our research offers several advantages to researchers and practitioners interested in the social dynamics of online communities, peer content production, and virtual collaboration. First, we offer a form for carrying out explanation and prediction. The application of our approach illustrates its potential for analyzing users of online communities, their behavior, their relations, and their activities in form of online collaboration and peer content production. Second, our research is capable of guiding both research and practice of online collaboration. If a “sound” understanding of online collaboration as a socially productive process is among the desired goals, researchers and practitioners may benefit from our insights on how to conceptualize the relationship between collaborators, affective communication, and trust. In particular, our findings on online trust in the context of subjectivity show a great potential for gaining insights on the way peer production develops with almost no centralized control. We acknowledge other areas dealing with the dynamics of social interaction in online collaboration, such as the analysis of gestures, problems of culture, or issues of social power as important (Baym 2006; Jiang et al. 2011). Pragmatically, there may be several perspectives for examining social interaction in online communities. Conditions other than sentiments or trust – users’ capabilities, characteristics and goals, their interpretations of technology, their work practices, and institutional contexts, power, or culture – may play key roles in causal explanations. Our results open a link to further experiments observing the affective implication of individuals who share a high value of trustworthiness.

Our work is not without limitations regarding both the semantic analysis of content and the generalizability of our findings. Each processing step is limited to the accuracy of the measurements we compute. For instance, the online trust measurement does not penalize repetitive edits in any way (Turek et al. 2010). We are also aware of the potential deficiencies associated with the existent behavioral approaches to measuring trust (McEvily 2011). Moreover, the accuracy of the identification of sentiments in talk pages is limited by the performance of the SentiStrength tool. Thelwall et al. (2010) performed an evaluation of SentiStrength on short informal texts, and reported that it performs with 96.9% accuracy on positive sentiment strength detection, while the detection of negative sentiment strength yielded 95.1% accuracy. Among the current sentiment analysis approaches, SentiStrength also performs best on Twitter postings (Nielsen 2011). We plan to use and compare alternative approaches from sentiment analysis in order to perform a more detailed comparison and benchmark of the results. Furthermore, the semantic convergence of articles takes into account only revision-based heuristics. An improvement to this measurement would be to compute the convergence only for those articles proven to meet a minimum standard of quality.

As regards generalizability, there are several characteristics that make Wikipedia significantly different from other online communities and virtual environments: established edit policies and norms, many active volunteers, the presence of administrators, as well as established social reputation among users. Moreover, if we consider the main usages of Wikipedia talk pages (request edit coordination, discuss controversial edits, ask for feedback, report vandalism, or refer to edit guidelines (Schneider et al. 2010)), researchers may even transfer and test our findings from Wikipedia to more general scenarios involving peer collaboration. An immediate point of interest would be to investigate team collaboration and peer content production in the context of another similar and rapidly growing resource, Wikia.⁶ With more than 34,000 of established wikis, Wikia is fundamentally different from the Wikipedia community in terms of having more permissive guidelines and policies, as well as a high number of small, topic-centered communities. A further direction worth investigating would be to analyze and test patterns of social

⁶ <http://www.wikia.com/Wikia>

collaboration in open source software development communities (e.g., Linux, Apache, GitHub, and SourceForge).

At a higher level, a large number of processes are currently increasingly conducted electronically (Overby 2008), as the trend of virtual collaborations that are organized without markets or managerial hierarchies is emerging in the context of information production systems. Understanding the factors that drive socially productive behavior is therefore increasingly important, as the virtual, self-organizing workplace constantly evolves towards more spontaneous and decentralized forms of mass collaboration (Tapscott and Williams 2006). Despite that the effect of virtualization of organizational structures on collaborative work has been in focus for some time (Assmann et al. 2010; Korsgaard et al. 2010; Markus 2004; Picot et al. 2008), the investigation of organizational behavior change has only shown that user online involvement in peer production does not necessarily follow the normal signals generated by either market-based or hierarchical models (Andreev et al. 2010; Benkler and Nissenbaum 2006). If both issues of motivating user contribution and organizing peer collaboration are overcome, then peer production of information potentially presents a systematic advantage over markets and firms in matching the best available human capital to the best available information inputs (Benkler 2006).

Conclusion

Online collaboration is becoming more and more important. Although the concept has found broad acceptance in practice, research or studies both on the factors influencing and on the effects of online collaboration are still underdeveloped. In this paper, we attempt to transfer concepts from sentiment analysis to the study of peer content collaboration in Wikipedia as a social process. To our knowledge, we are the first to examine the trust behavior modeled by the subjective user implication. Our main finding is that sharing feedback in form of sentiments and opinions positively affects online trust in inter-user interactions. We hope that our transfer of concepts will be useful for other researchers and practitioners who investigate factors that shape online collaborative behavior. As Wikipedia has a series of specific qualities that make it difficult to generalize, we intend to apply our approach to other wiki communities (such as Wikia).

As further work, we propose to analyze the amount of sentiments in peer production communities supporting social networking. It would be worthwhile to investigate the evolution of affective implication of users in connection to changes in the group collaborative behavior. Moreover, at the community level, one may investigate whether the amount of sentiments contained in discussions can be correlated to the frequency of interactions between members. Furthermore, we intend to replicate and test our findings by further performing controlled experiments with regard to the level of trust and the amount of sentiments.

References

- Akkaya, C., Wiebe, J., and Mihalcea, R. 2009. "Subjectivity word sense disambiguation," in: *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1 - Volume 1*. Singapore: Association for Computational Linguistics.
- Andreev, P., Feller, J., Finnegan, P., and Moretz, J. 2010. "Conceptualizing the Commons-Based Peer Production of Software: An Activity Theoretic Analysis," *ICIS 2010 Proceedings*.
- Antin, J., and Cheshire, C. 2010. "Readers are not free-riders: reading as a form of participation on wikipedia," in: *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. Savannah, Georgia, USA: ACM, pp. 127-130.
- Aral, S., Muchnik, L., and Sundararajan, A. 2009. "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *Proceedings of the National Academy of Sciences* (106:51), pp. 21544-21549.
- Arazy, O., and Nov, O. 2010. "Determinants of wikipedia quality: the roles of global and local contribution inequality," in: *CSCW '10: Proceedings of the 2010 ACM conference on Computer supported cooperative work*. New York, NY, USA: ACM.
- Assmann, J.J., Drescher, M.A., Gallenkamp, J., Picot, A., and Welp, I.M. 2010. "Here Be (No More) Dragons: Pushing the Frontier of Research on Virtual Organizations and Teams," in: *Proceedings of the 2010 43rd Hawaii International Conference on System Sciences*. IEEE Computer Society, pp. 1-10.

- Banea, C., Mihalcea, R., Janyce, W., and Samer, H. 2008. "Multilingual subjectivity analysis using machine translation," in: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Honolulu, Hawaii: Association for Computational Linguistics.
- Barrett, L., and Russell, J. 1999. "The Structure of Current Affect: Controversies and Emerging Consensus," *Current Directions in Psychological Science* (8:1).
- Baym, N.K. 2006. "Interpersonal life online," *Handbook of new media: social shaping and social consequences of ICTs*, pp. 35-54.
- Benkler, Y. 2002. "Coase's penguin, or Linux and the nature of the firm», " *Yale Law Journal* (4).
- Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*.
- Benkler, Y., and Nissenbaum, H. 2006. "Commons based Peer Production and Virtue*," *Journal of Political Philosophy* (14:4), pp. 394-419.
- Bishop, J. 2005. "A model for understanding and influencing behaviour in virtual communities," *Proceedings of the Post-Cognitivist Psychology Conference 2005*.
- Bishop, J. 2007. "Increasing participation in online communities: A framework for human-computer interaction," *Comput. Hum. Behav.* (23:4), pp. 1881-1893.
- Boyle, R., and Bonacich, P. 1970. "The Development of Trust and Mistrust in Mixed-Motive Games," *Sociometry* (33:2), 06/01, pp. 123-139.
- Burnett, G. 2000. "Information exchange in virtual communities: a typology," *Information Research* (5:4).
- Chambliss, M.J., and Garner, R. 1996. "Do Adults Change their Minds after Reading Persuasive Text?," *Written Communication* (13:3), July 1, 1996, pp. 291-313.
- Christopher, T., and Amit, P.S. 2007. "Semantic Convergence of Wikipedia Articles," in: *Proceedings of the IEEE/WIC/ACM International Conference on Web Intelligence*. IEEE Computer Society.
- Clark, H.H. 1996. *Using Language*. New York, NY, USA: Cambridge University Press.
- Clark, H.H. 2005. "Coordinating with each other in a material world," *Discourse Studies* (7:4-5), pp. 507-525.
- Convertino, G., Mentis, H., Rosson, M.B., Carroll, J., Slavkovic, A., and Ganoe, C. 2008. "Articulating Common Ground in Cooperative Work: Content and Process," *Proceeding of the Annual SIGCHI Conference: Human Factors in Computing Systems* Florence, Italy: ACM, pp. 1637-1646.
- Convertino, G., Mentis, H.M., Rosson, M.B., Slavkovic, A., and Carroll, J.M. 2009. "Supporting content and process common ground in computer-supported teamwork," *27th International Conference on Human Factors in Computing Systems*, Boston, MA, USA: ACM, pp. 2339-2348.
- Costa, A.C., and Anderson, N. 2010. "Measuring trust in teams: Development and validation of a multifaceted measure of formative and reflective indicators of team trust," *European Journal of Work and Organizational Psychology* (20:1), pp. 119 - 154.
- Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., and Siddharth, S. 2008. "Feedback effects between similarity and social influence in online communities," *14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Las Vegas, Nevada, USA: ACM.
- Denning, P.J., and Yaholkovsky, P. 2008. "Getting to "we": Solidarity, not software, generates collaboration," *Communications of the ACM* (51:4), pp. 19-24.
- Deutsch, M. 1958. "Trust and suspicion," *Journal of Conflict Resolution* (2:4), December 1, 1958, pp. 265-279.
- Downs, J.S., Holbrook, M.B., Sheng, S., and Cranor, L.F. 2010. "Are your participants gaming the system?: screening mechanical turk workers," *28th International Conference on Human Factors in Computing Systems*, Atlanta, GA, USA: ACM.
- Erk, K., and Pado, S. 2008. "A structured vector space model for word meaning in context," in: *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Honolulu, Hawaii: Association for Computational Linguistics.
- Eseryel, U.Y. 2009. "Leadership in a Non-Traditional Setting: Self-Managing Virtual IS Development Teams," in: *ICIS 2009 Proceedings*.
- Farzan, R., Dabbish, L.A., Kraut, R.E., and Postmes, T. 2011. "Increasing commitment to online communities by designing for social presence," in: *Proceedings of the ACM 2011 conference on Computer supported cooperative work*. Hangzhou, China: ACM, pp. 321-330.
- Feng, P., Lazar, J., and Preece, J. 2004. "Empathy and online interpersonal trust: A fragile relationship," *Behaviour & Information Technology* (23:2).
- Fiedler, M., and Sarstedt, M. 2010. "Influence of Community Design on User Behaviors in Online Communities," *ICIS 2010 Proceedings*, St. Louis.

- Furuse, O., Nobuaki, H., Setsuo, Y., and Ryoji, K. 2007. "Opinion sentence search engine on open-domain blog," in: *Proceedings of the 20th international joint conference on Artificial intelligence*. Hyderabad, India: Morgan Kaufmann Publishers Inc.
- Gan, L., Gu, B., Jarvenpaa, S., and Yang, G. 2009. "Habit Formation in Online Communities," *Thirtieth International Conference on Information Systems (ICIS 2009)*, Phoenix, Arizona.
- Gensollen, M., Brousseau, E., and Curien, N. 2007. *Information goods and online communities*. Cambridge University Press.
- Ginsburg, M., and Weisband, S. 2002. "Social capital and volunteerism in virtual communities: the case of the Internet Chess Club," *35th Annual Hawaii International Conference on System Sciences*, Hawaii, USA, pp. 2225-2234.
- Gurzik, D., and Lutters, W.G. 2009. "Towards a design theory for online communities," in: *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology*. Philadelphia, Pennsylvania: ACM, pp. 1-20.
- Hagel, J., and Armstrong, A.G. 1997. *Net Gain: Expanding Markets through Virtual Communities*. Boston, MA, USA: Harvard Business School Press.
- Harris, R.B., and Paradice, D. 2007. "An Investigation of the Computer-mediated Communication of Emotions," *J. of Applied Sciences Research* (3), pp. 2081-2090.
- Hill, C.W.L. 1995. "National Institutional Structures, Transaction Cost Economizing and Competitive Advantage: The Case of Japan," *Organization Science* (6:1), 1995/01, pp. 119-131.
- Hwang, H.-G., Ku, C.-Y., Yen, D.C., and Cheng, C.-C. 2004. "Critical factors influencing the adoption of data warehouse technology: a study of the banking industry in Taiwan," *Decision Support Systems* (37:1), pp. 1-21.
- Javanmardi, S., and Lopes, C. "Statistical measure of quality in Wikipedia," in: *Proceedings of the First Workshop on Social Media Analytics*. Washington D.C., District of Columbia: ACM.
- Jiang, L., Bazarova, N.N., and Hancock, J.T. 2011. "From Perception to Behavior: Disclosure Reciprocity and the Intensification of Intimacy in Computer-Mediated Communication," *Communication Research*.
- Jøsang, A. 2007. "Prospectives for online trust management," *IEEE Transactions on Knowledge and Data Engineering*.
- Kittur, A., Chi, E.H., and Suh, B. 2008. "Crowdsourcing user studies with Mechanical Turk," *26th Annual SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy: ACM.
- Kittur, A., and Kraut, R.E. 2010. "Beyond Wikipedia: coordination and conflict in online production groups," in: *Proceedings of the 2010 ACM conference on Computer supported cooperative work*. Savannah, Georgia, USA: ACM.
- Korsgaard, M.A., Picot, A., Wigand, R.T., Welp, I.M., and Assmann, J.J. 2010. "Cooperation, Coordination, and Trust in Virtual Teams: Insights from Virtual Games," in: *Online Worlds: Convergence of the Real and the Virtual*, W.S. Bainbridge (ed.). Springer London, pp. 253-264.
- Lee, K.Y., Lee, M., Bassellier, G., and Faraj, S. 2010. "THE IMPACT OF EMOTIONAL EXPRESSIONS ON KNOWLEDGE CREATION IN ONLINE COMMUNITIES," *ICIS 2010 Proceedings*.
- Leimeister, J.M., Sidiras, P., and Krcmar, H. 2006. "Exploring Success Factors of Virtual Communities: The Perspectives of Members and Operators," *Journal of Organizational Computing and Electronic Commerce* (16:3 & 4), pp. 277-298.
- Lewis, D., and Weigert, A. 1985. "Trust as a Social Reality," *Social Forces* (63:4), pp. 967-985.
- Lin, H.-F. 2008. "Determinants of successful virtual communities: Contributions from system characteristics and social factors," *Information & Management* (45:8), pp. 522-527.
- Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Li, X., Cosley, D., Frankowski, D., Terveen, L., Rashid, A.M., Resnick, P., and Kraut, R. 2005. "Using social psychology to motivate contributions to online communities," *Journal of Computer-Mediated Communication* (10:4), p. article 10.
- Maher, M.L., Paulini, M., and Murty, P. 2011. "Scaling Up: From Individual Design to Collaborative Design to Collective Design," in: *Design Computing and Cognition '10*, J.S. Gero (ed.). Springer Netherlands, pp. 581-599.
- Maloney-Krichmar, D., and Preece, J. 2005. "A multilevel analysis of sociability, usability, and community dynamics in an online health community," *ACM Trans. Comput.-Hum. Interact.* (12:2), pp. 201-232.
- Markus, M.L. 2004. "Technochange management: using IT to drive organizational change," *Journal of Information Technology (Palgrave Macmillan)* (19:1), pp. 4-20.

- McEvily, B. 2011. "Reorganizing the Boundaries of Trust: From Discrete Alternatives to Hybrid Forms," *Organization Science*, p. orsc. 1110.0649 v1111.
- Metiu, A., and Kogut, B. 2001. "Distributed knowledge and the global organization of software development," *Wharton School, University of Pennsylvania*.
- Meyerson, D., Weick, K.E., and Kramer, R.M. 1996. "Swift trust and temporary groups," *Trust in organizations: Frontiers of theory and research* (166), p. 195.
- Mihalcea, R. 2007. "The Multidisciplinary Facets of Research on Humour," in: *Proceedings of the 7th international workshop on Fuzzy Logic and Applications: Applications of Fuzzy Sets Theory*. Camogli, Italy: Springer-Verlag.
- Muller-Birn, C., Lehmann, J., and Jeschke, S. 2009. "A Composite Calculation for Author Activity in Wikis: Accuracy Needed," in: *Proceedings of the 2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology - Volume 01*. IEEE Computer Society.
- Murphy, P.K., Long, J.F., Holleran, T.A., and Esterly, E. 2003. "Persuasion online or on paper: a new take on an old issue," *Learning and instruction* (13:5), pp. 511-532.
- Musser, J., and O'Reilly, T. 2007. *Web 2.0 - Principles and Best Practices*. Sebastopol, CA, USA: O'Reilly Media.
- Nemoto, K., Gloor, P., and Laubacher, R. 2011. "Social capital increases efficiency of collaboration among Wikipedia editors," in: *Proceedings of the 22nd ACM conference on Hypertext and hypermedia*. Eindhoven, The Netherlands: ACM, pp. 231-240.
- Nielsen, F.Å. 2011. "A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," *CoRR* (abs/1103.2903).
- Overby, E. 2008. "Process Virtualization Theory and the Impact of Information Technology," *Organization Science* (19:2), March 1, 2008, pp. 277-291.
- Paltoglou, G., and Thelwall, M. 2010. "A study of information retrieval weighting schemes for sentiment analysis," in: *Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics*. Uppsala, Sweden: Association for Computational Linguistics.
- Pang, B., and Lee, L. 2008. "Opinion Mining and Sentiment Analysis," *Found. Trends Inf. Retr.* (2:1-2), pp. 1-135.
- Parameswaran, M., and Whinston, A.B. 2007. "Research Issues in Social Computing," *Journal of the Association for Information Systems* (8:6).
- Picot, A., Reichwald, R., and Wigand, R. 2008. *Information, Organization and Management*. Berlin, Germany et al.: Springer.
- Porra, J., and Parks, M.S. 2006. "Sustainable virtual communities: suggestions from the colonial model," *Information Systems and E-Business Management* (4:4), pp. 309-341.
- Preece, J. 2000. *Online Communities. Designing Usability, Supporting Sociability*. Chichester et al., UK: John Wiley & Sons.
- Preece, J. 2001. "Sociability and usability in online communities: determining and measuring success," *Behaviour & Information Technology* (20:5), pp. 347-356.
- Preece, J., Nonnecke, B., and Andrews, D. 2004. "The top five reasons for lurking: improving community experiences for everyone," *Computers in Human Behavior* (20:2), pp. 201-223.
- Preece, J., and Shneiderman, B. 2009. "The Reader-to-Leader Framework: Motivating technology-mediated social participation," *AIS Transactions on Human-Computer Interaction* (1:1), pp. 13-32.
- Rafaeli, S., and Raban, D. 2005. "Information sharing online: a research challenge," *International Journal of Knowledge and Learning* (1), pp. 62-79.
- Ren, Y., Harper, M., Drenner, S., Kiesler, S., Terveen, L., Riedl, J., and Kraut, R. 2009. "Building Attachment to Online Communities: Design Principles for Identity and Bonds," *Working Paper*.
- Ren, Y., Kraut, R., and Kiesler, S. 2007. "Applying Common Identity and Bond Theory to Design of Online Communities," *Organization Studies* (28:3), pp. 377-408.
- Rheingold, H. 1993. *The Virtual Community: Finding Connection in a Computerized World*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA.
- Ross, J., Irani, L., Silberman, M.S., Zaldivar, A., and Tomlinson, B. 2010. "Who are the crowdworkers?: shifting demographics in mechanical turk," *28th of the International Conference on Human Factors in Computing Systems*, Atlanta, GA, USA: ACM.
- Scherer, Klaus, R., Wranik, Tanja, Sangsue, Janique, Tran, ronique, Scherer, Ursula, Averill, James, R., Ben, Z.E., Aaron, Revhon, Netta, Goldie, Peter, Izard, Carroll, E., King, Kristen, A., Trentacosta,

- Christopher, J., Katz, Jack, Le, B., David, Parkinson, Brian, Wilhelm, Peter, Schoebi, Dominik, Perrez, and Meinrad. 2004. *Emotions in everyday life: probability of occurrence, risk factors, appraisal and reaction patterns. Commentary. Authors' reply*. London, ROYAUME-UNI: Sage.
- Schneider, J., Passant, A., and Breslin, J. 2010. "A Qualitative and Quantitative Analysis of How Wikipedia Talk Pages Are Used," *Web Science Overlay Journal*.
- Shah, D.V., Kwak, N., and Holbert, R.L. 2001. "'Connecting" and "disconnecting" with civic life: Patterns of Internet use and the production of social capital," *Political Communication* (18:2), pp. 141-162.
- Shah, S.K. 2006. "Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development," *Manage. Sci.* (52:7), pp. 1000-1014.
- Shanmugasundaram, H., Ramachandran, S., Murugan, S., and Saranathan, P. 2009. "Opinion mining and summarization of reviews in web forums," in: *Proceedings of the Third Annual ACM Bangalore Conference*. Bangalore, India: ACM.
- Skopik, F., Truong, H.-L., and Dustdar, S. 2009. "Trust and Reputation Mining in Professional Virtual Communities," in: *Web Engineering*, M. Gaedke, M. Grossniklaus and O. Diaz (eds.). Springer Berlin / Heidelberg, pp. 76-90.
- Spoor, J.R., and Kelly, J.R. 2004. "The Evolutionary Significance of Affect in Groups: Communication and Group Bonding," *Group Processes & Intergroup Relations* (7:4), October 1, 2004, pp. 398-412.
- Sproull, L., and Kiesler, S. 1991. *Connections: New ways of working in the networked organization*. Cambridge, MA, USA: The MIT Press.
- Stephen, M.M., and Suzanne, S. 2006. "Open Source Software: The New Intellectual Property Paradigm," National Bureau of Economic Research, Inc.
- Susan, L.B., Andrea, F., and Amy, B. 2005. "Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia," in: *Proceedings of the 2005 international ACM SIGGROUP conference on Supporting group work*. Sanibel Island, Florida, USA: ACM.
- Taddeo, M., and Vaccaro, A. 2011. "Analyzing Peer-to-Peer Technology Using Information Ethics," *The Information Society* (27:2), pp. 105-112.
- Tapscott, D., and Williams, A.D. 2006. *Wikinomics: How Mass Collaboration Changes Everything*. New York: Portfolio Hardcover.
- Tarmizi, H., and Vreede, G.-J.d. 2005. "A Facilitation Task Taxonomy for Communities of Practice," *11th Americas Conference in Information Systems (AMCIS 2005)*, Omaha, NE, USA, pp. 3545-3554.
- Te'eni, D. 2001. "Review: A Cognitive-Affective Model of Organizational Communication for Designing IT," *MIS Quarterly* (25:2), pp. 251-312.
- Thelwall, M., Buckley, K., and Paltoglou, G. 2011. "Sentiment in Twitter events," *J. Am. Soc. Inf. Sci. Technol.* (62:2), pp. 406-418.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., and Kappas, A. 2010. "Sentiment in short strength detection informal text," *J. Am. Soc. Inf. Sci. Technol.* (61:12), pp. 2544-2558.
- Tkacz, N. 2010. "Wikipedia and the Politics of Mass Collaboration," *Journal of Media and Communication*.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. 2005. "Understanding and sharing intentions: The origins of cultural cognition," *Behavioral and Brain Sciences* (28:675-691).
- Turek, P., Wierzbicki, A., and Nielek, R. 2010. "WikiTeams: Evaluating Teamwork in Wikipedia," in: *WebScio: Extending the Frontiers of Society On-Line*. Raleigh, NC: US.
- Turney, P.D., and Pantel, P. 2010. "From frequency to meaning: vector space models of semantics," *J. Artif. Int. Res.* (37:1), pp. 141-188.
- VandenBos, G.R. 2006. *APA dictionary of psychology*. Washington, DC :: American Psychological Association.
- Viegas, F.B., Wattenberg, M., Kriss, J., and van Ham, F. 2007. "Talk Before You Type: Coordination in Wikipedia," in: *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*. IEEE Computer Society.
- Vossen, G., and Hagemann, S. 2007. *Unleashing Web 2.0: From Concepts to Creativity*. Burlington, MA, USA: Morgan Kaufmann
- Wallace, P.M. 2001. *The psychology of the Internet*. Cambridge Univ Pr.
- Wang, X., Butler, B., Burke, M., Kraut, R., and Joyce, E. 2009. "Beyond Information: Developing the Relationship Between the Individual and the Group in Online Communities," *Working Paper*.
- Wiebe, J., and Mihalcea, R. 2006. "Word sense and subjectivity," in: *Proceedings of the 21st International Conference on Computational Linguistics and the 44th annual meeting of the Association for Computational Linguistics*. Sydney, Australia: Association for Computational Linguistics.

- Wierzbicki, A., Turek, P., and Nielek, R. 2010. "Learning about team collaboration from Wikipedia edit history," in: *Proceedings of the 6th International Symposium on Wikis and Open Collaboration*. Gdansk, Poland: ACM.
- Williamson, O.E. 1993. "Calculativeness, Trust, and Economic Organization," *Journal of Law & Economics* (36:1), pp. 453-486.
- Wilson, T., Paul, H., Swapna, S., Jason, K., Janyce, W., Yejin, C., Claire, C., Ellen, R., and Siddharth, P. 2005. "OpinionFinder: a system for subjectivity analysis," in: *Proceedings of HLT/EMNLP on Interactive Demonstrations*. Vancouver, British Columbia, Canada: Association for Computational Linguistics.
- Wu, J.-J., Chen, Y.-H., and Chung, Y.-S. 2010. "Trust factors influencing virtual community members: A study of transaction communities," *Journal of Business Research* (63:9-10), pp. 1025-1032.