AIS Transactions on Human-Computer Interaction

Volume 7 | Issue 4 Article 1

12-2015

The Impact of Sentiment-driven Feedback on Knowledge Reuse in Online Communities

Mihai Grigore
ETH Zurich, mihaigrigore@gmx.de

Christoph Rosenkranz
University of Cologne, Cologne, Germany, rosenkranz@wiso.uni-koeln.de

Juliana Sutanto
Lancaster University, j.sutanto@lancaster.ac.uk

Recommended Citation

Grigore, M., Rosenkranz, C., & Sutanto, J. (2015). The Impact of Sentiment-driven Feedback on Knowledge Reuse in Online Communities. *AIS Transactions on Human-Computer Interaction, 7*(4), 212-232. Retrieved from https://aisel.aisnet.org/thci/vol7/iss4/1
DOI:

Transactions on Human – Computer Interaction

Original Empirical Research

The Impact of Sentiment-driven Feedback on Knowledge Reuse in Online Communities

Mihai Grigore
ETH Zurich
Zurich, Switzerland
mihaigrigore@gmx.de

Christoph Rosenkranz
University of Cologne
Cologne, Germany
rosenkranz@wiso.uni-koeln.de

Juliana Sutanto
Lancaster University
Lancaster, United Kingdom
j.sutanto@lancaster.ac.uk

Abstract:

Knowledge reuse is of increasing importance for organizations. Despite the extant research, we still do not adequately understand the ways peers are motivated to reuse knowledge with the help of wiki technologies. In this paper, we study the motivation for knowledge reuse in a prominent instance of online social production: Wikipedia. Studying knowledge reuse in Wikipedia is important since Wikipedia has been able to leverage the benefits of efficient knowledge reuse to produce knowledge goods of relatively high quality. Specifically, we explore: 1) how Wikipedia editors (any peer who contributes to developing articles in Wikipedia) communicate their feedback toward each other's work in peer conversations and 2) to what extent sentiment-driven feedback impacts the level of knowledge reuse in Wikipedia. The results show that displaying sentiment-driven feedback positively influenced the level of knowledge reuse. Our study further shows a significant difference in the level of knowledge reuse between editors who shared mainly positive or mainly negative sentiments. Specifically, displaying mainly positive feedback corresponded to a superior level of knowledge reuse than displaying mainly negative feedback. We contribute to the extant literature of online social production communities in general and Wikipedia in particular by providing a first building block for research on peer feedback's role in developing and sustaining wiki-based knowledge reuse. We discuss our findings' implications for theory and practice.

Keywords: Knowledge Reuse, Sentiment-driven Feedback, Affective Communication Affect in Information Systems Research, Online Social Production, Online Collaboration.

1 Introduction

Knowledge represents an essential resource for organizations in today's economic environment. Successful organizations build or have built dynamic capabilities to create, acquire, share, use, and reuse knowledge (Alavi & Leidner, 2001; Argote, McEvily, & Reagans, 2003; Argote & Miron-Spektor, 2011; Khodakarami & Chan, 2014). Being localized, embedded, and invested in practice, knowledge represents a critical asset that may prevent organizations from spending time and resources on redeveloping already existing solutions (Carlile, 2002). Researchers have shown that, by leveraging knowledge that already exists, knowledge reuse can enhance efficient and effective problem solving in organizations (Gray, 2001). However, there is infrequent or even lack of knowledge reuse (Liu, Chai, & Nebus, 2013; Rozwell, 2009).

The rise of Internet technologies has facilitated increased access to peers, resources, information, and knowledge even outside the boundaries of traditional organizations. In this sense, new forms of organizing (such as online social production) have emerged and opened substantial opportunities to research knowledge processes at unprecedented scales (Puranam, Alexy, & Reitzig, 2013). Online social production has become an increasingly viable and popular way to create knowledge goods that are often of relatively high quality (Faraj, Jarvenpaa, & Majchrzak, 2011; Giles, 2005; von Hippel & von Krogh, 2003). Although information technologies have been subject to extensive research for their role in facilitating the reuse of knowledge among peers who hold knowledge (Nonaka & Takeuchi, 1995), we still lack empirical research on knowledge reuse in online social production. Indeed, whereas knowledge reuse in traditional organizations has been relatively well researched (e.g., see Majchrzak, Wagner, & Yates, 2013b; Markus, 2001), knowledge reuse in online communities continues to be under-researched (Haefliger, von Krogh, & Spaeth, 2008; von Krogh, Haefliger, Spaeth, & Wallin, 2012).

Wikipedia, a prominent example of an online social production community, has become one of the world's most popular sources of knowledge with more than four million articles. Researchers have repeatedly found Wikipedia entries' quality to be on par with traditionally organized processes carried out by professional editors over several years, such as the Encyclopedia Britannica (Giles, 2005; Tapscott & Williams, 2006). Given the importance of technology-fostered knowledge reuse, there are three specific reasons for considering Wikipedia as a resource for examining a success story of wiki-enabled knowledge reuse. First, the underlying wiki technology records the full editing and interaction activity for each article; thus, Wikipedia enables its users to integrate others' knowledge for efficient knowledge reuse (Grant, 1996). Second, Wikipedia's articles can only be edited using the Wikipedia platform, which allows researchers to have a complete editing and social interaction history of each article. Third, any Internet user can contribute knowledge to the articles, which allows researchers to examine group interactions in an uncontrolled setting.

The extant literature on motivation to contribute to online social production has established that peers follow diverse motivational drives (e.g., the pleasure involved in completing a task) and social signals (e.g., community belonging and social recognition) (Benkler, 2006; Puranam et al., 2013). Moreover, previous research on factors that motivate one's contributing knowledge in communities of practice has focused mostly on factors explaining why peers contribute their personal knowledge (Carlile, 2004; Carlile & Rebentisch, 2003) and little on why peers reuse the knowledge contributed by others in online social production communities (Yates, Wagner, & Majchrzak, 2010). Through a field survey on customer service, researchers have found evidence that intrinsic motivation positively influences knowledge reuse with the help of electronic repositories (Kankanhalli et al., 2011). A recent study in the context of organizational Intranets supported by wikis shows that knowledge shaping promotes knowledge reuse through improved integration of knowledge (Majchrzak et al., 2013b). In this paper, we explain the motivational factors that lead to the success story of knowledge reuse in Wikipedia. Specifically, we address the following research question: "How and why do wiki contributors reuse knowledge?".

To answer this question, we conducted a longitudinal analysis of peers' editing and interaction activity in Wikipedia. We build on the findings of Markus (2001) on knowledge reuse. Specifically, one of the aspects that Markus (2001) stresses is that successful knowledge reuse is in part a matter of how to provide incentives for contributions. In this study, we refer to Wikipedia editors as any individual who contributes to the development of Wikipedia articles, and we regard knowledge reuse among Wikipedia editors as an aspect of *collaboration* that is influenced by specific *communication practices*. We examine the way Wikipedia editors communicate their feedback towards others' work and explore to what extent this communication acts as an incentive to reuse knowledge. In particular, using knowledge reuse as a proxy

for collaboration, we investigate whether affective communication (Te'eni, 2001; Zhang, 2013) in form of sentiment-driven feedback in discussions between Wikipedia editors motivates collaborative work. In doing so, we contribute to the extant literature on knowledge reuse (Majchrzak et al., 2013b; Markus, 2001) by providing the first building block for research on the role of peer feedback on developing and sustaining wiki-based knowledge reuse.

This paper proceeds as follows. In Section 2, we discuss the related work and the conceptual background of our research. In Section 3, we present our research model and methodology. The dataset used in this study comprises a complete revision history dump of Simple English Wikipedia[1]. In Section 4, we employ methods from sentiment analysis, compute Wikipedia-specific metrics, and use regression analysis to produce and analyze our data. Subsequently, in Section 5, we conclude the paper with an outlook for further research.

2 Related Work and Theory Development

2.1 Theoretical Findings on Motivations for Knowledge Reuse and the Role of Peer Feedback

Previous studies have systematically reported that traditional organizational repositories are not suitable to efficiently and effectively leverage the knowledge in organizations (Rafaeli & Ariel, 2008; Yates et al., 2010). In a recent attempt to explain knowledge reuse in communities of practice, Majchrzak et al. (2013b) focused on the unique affordance of wiki technologies to foster online knowledge integration for knowledge reuse. Contributing knowledge to a wiki may involve not only contributing the knowledge of one's domain expertise but also integrating knowledge already contributed to the wiki in order to make it more logically organized. In a wiki-based knowledge-sharing context, one can often visually observe and track knowledge reuse (Chi et al., 2008; Grudin & Poole, 2010).

Besides the recent interest in wiki technologies for supporting communities of practice in traditional organizations, researchers have intensely investigated wikis as a cornerstone of online social production communities. Online social production communities (also referred to as commons-based peer production) have two defining characteristics: 1) they are based on the online collaboration of volunteers who carry out productive activities primarily for social and psychological purposes rather than for financial remuneration (Benkler, 2006; Shirky, 2010) and 2) the online production apparently happens in the absence of governance mechanisms based on price mechanisms or hierarchical, managerial structures (Aaltonen & Lanzara, 2011; Benkler, 2006). The first characteristic has motivated a bulk of studies in online social production communities to examine the participating individuals' motivational drivers (e.g., Benkler, 2006; Ghosh, 2005; Hahn, Moon, & Zhang, 2008; Lerner & Tirole, 2002; Shah, 2006; Stewart & Gosain, 2006; Weber, 2004). These motivational drivers range from altruism and enjoyment to solving challenging problems, social recognition, and future employment benefit. The second characteristic of online social production communities has motivated research on the governance mechanisms of such communities (e.g., Feller, Finnegan, Hayes, & O'Reilly, 2008; Mehra, 2012; Singh, 2010). Such communities are typically governed by self-organization (Crowston, Li, Wei, Eseryel, & Howison, 2007), a rather slow and difficult process to ensure global coordination out of local interactions between people. Self-organization is fostered by Internet technologies that keep a detailed trace of the community members' interactions while they are interacting in real time (Lanubile, Ebert, Prikladnicki, & Vizcaino, 2010). For example, the extant literature on online social production communities has examined Internet technologies' role in enabling peers to interact with each other (Burnett, 2000; Preece, 2001); usability and sociability are factors that make online communities successful (Preece, 2001; Tarmizi & Vreede, 2005). Based on these findings, Porra and Parks (2006) suggest that the sustainability of online social production 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 online social production communities' success.

Following this line of reasoning, we examine what motivates individuals to collaborate in online social production with an emphasis on knowledge reuse in Wikipedia. In this sense, in an attempt toward establishing a theory of knowledge reuse, Markus (2001) stresses that successful knowledge reuse is in part a matter of how to provide incentives for high-quality contributions. Indeed, researchers have found

[1] http://simple.wikipedia.org

intrinsic motivation to positively influence knowledge reuse through electronic repositories (Kankanhalli et al., 2011). Feng et al. (2004) further emphasize the importance of developing supportive working relationships online.

In online settings, the focus of attention changes from the relationship between peers and the technology to the relationship between peers and the community; peers who never physically meet or know each other get to communicate and work collectively. Researchers have found empathy to be essential in encouraging peers to work together online (Leimeister, Sidiras, & Krcmar, 2006; Maloney-Krichmar & Preece, 2005; Skopik, Truong, & Dustdar, 2009). In the case of online social production communities, the apparent lack of formal authority may be compensated by individuals who *mentor* and *encourage* each other towards contributing knowledge (Eseryel, 2009). That is, providing *sentiment-driven feedback* (or *affective feedback*) may act as a powerful motivational factor towards superior work outcomes (Bateman & Organ, 1983; Zhu, Zhang, He, Kraut, & Kittur, 2013). Further results from offline settings acknowledge that strategically using feedback may increase recipients' motivation to adhere to their goals (Fishbach, Eyal, & Finkelstein, 2010). Concretely, researchers have recognized sentiment-driven feedback to intrinsically motivate goal pursuit, which happens through the affective experience that sentiment-driven feedback is able to produce (Fishbach et al., 2010). We elaborate on the concepts of sentiment-driven feedback, affective communication, and affect in online settings and their usage in information systems in the Section 2.2.

2.2 Sentiment-driven Feedback and its Motivating Role for Knowledge Reuse

Sentiment-driven feedback (or affective feedback) is a form of communication used to express affect such as praise (e.g., "well written") or criticism (e.g., "badly written") (Nelson & Schunn, 2009). Affect—also commonly referred to as sentiment, emotion, or mood—represents "general moods (happiness, sadness) or specific emotions (fear, anger, envy)" as reaction to things one thinks about, to actions one takes, or to various stimuli (Ajzen, 2001, p. 29; Barrett & Russell, 1999; VandenBos, 2006). Social psychologists and IS researchers have paid increasing attention to affect (Scherer, Warnik, Sangsue, Tran, & Scherer, 2004; Van der Heijden, 2004; Zhang, 2013). Recent research studies acknowledge affect's increasing importance on information systems usage and on online work behavior. In this context, researchers have shown affect to be efficiently externalized through computer-mediated communication (Harris & Paradice, 2007). Researchers have further recognized the intensity of affect as a means for coping with communication complexity to achieve communication goals; affect is a suitable means to motivate and to inform (Te'eni, 2001). Stieglitz and Dang-Xuan (2013) investigate the relationship between affect and information diffusion in microblogging websites and find that affective Twitter messages tend to be re-tweeted more often and more quickly compared to neutral ones. Moreover, Aggarwal, Gopal, Sankaranarayanan, and Singh (2012) focus on affect in social media and study the effects of negative posts from employee blogs. They reveal the potentially positive influence of negative posts in the sense that negative posts may act as catalyst that can exponentially increase the awareness of employee blogs. Although several theoretical studies provide a foundation for the concept of affect and indicate that affect indeed impacts on information systems usage and on online work behavior, empirical evidence on the role and influence of affect in online social production communities continues to be underdeveloped (Beaudry & Pinsonneault, 2010; Zhang, 2013).

Transferring these insights to our study, we argue that intrinsic motivations for knowledge reuse may be at least partially explained by the display of affect in inter-editor conversations. Indeed, the articulation of affect both in spoken discourse and in written text (Te'eni, 2001)—that is, affective communication—has the potential to act as awareness catalyst. Affective communication helps coordinating group activity by fostering group bonds (Spoor & Kelly, 2004). Our reasoning is in line with Kankanhalli et al. (2011) who found that intrinsic motivation positively influences knowledge reuse through electronic repositories. To support this reasoning for online social production communities, we first consider prior research on Wikipedia that has shown that an intensification in peer collaboration usually occurs after the initiation of conversations among editors on talk pages (Crandall, Cosley, Huttenlocher, Kleinberg, & SIddharth, 2008). Moreover, researchers have found peer influence exerted across social ties among peers to be a significant predictor of future collaborative behavior in Wikipedia (Crandall et al., 2008). Discussion pages offer Wikipedia editors the means to communicate their achievements and constantly receive feedback on the progress of their work (Reagle, 2010). Peers need to feel that their engagement is beneficial to the organization or community (Haefliger, Monteiro, Foray, & von Krogh, 2011; Stahlbrost & Bergvall-Kareborn, 2011). This perception of appreciation (i.e., affect) can be transmitted through affective communication in form of sentiment-driven feedback on discussion pages in Wikipedia. We further argue

that affective communication in form of sentiment-driven feedback in inter-editor discussions may act as intrinsic motivator for knowledge reuse in Wikipedia.

In this study, we 1) analyze how Wikipedia editors communicate their feedback on others' contribution on discussion pages and 2) explore to what extent affective communication in form of sentiment-driven feedback on discussion pages impacts the level of knowledge reuse in Wikipedia. In doing so, we integrate the analysis of specific peer collaboration in editing Wikipedia articles with the analysis of the intensity of affective communication between the particular Wikipedia editors. Accordingly, we hypothesize that:

H1: The display of sentiment-driven feedback in inter-editors communication corresponds to increased levels of knowledge reuse than in the case of neutral feedback.

The orientation and intensity of affective communication can range on a scale from being very positive to very negative (Barrett & Russell, 1999). The distinction between positive and negative sentiments on the one hand and neutral statements in conversations on the other hand could be useful for explaining collective behavior because sentiments become externalized instances on the collective level (such as "collective sentiment") (Scherer et al., 2004). Depending on the sensitivity to attitudes and changes in disposition voiced in affective communication (Te'eni, 2001), peers should be influenced differently by positive or negative feedback. In this sense, making a habit of dispensing positive feedback rather than negative feedback is more likely to motivate peers to perform with confidence and autonomy (Lickerman, 2012). However, negative feedback is needed when something is being done incorrectly to give peers the opportunity to improve the result. Research has shown that, when possible, positive feedback should be used in public whereas negative feedback is rather effective for correcting problems, behaviors, and attitudes in private (Fishbach et al., 2010). Accordingly, we hypothesize that:

H2: Displaying positive feedback in inter-editors communication corresponds to greater levels of knowledge reuse than displaying negative feedback.

3 Research Method

3.1 Data Set

Wikipedia relies on the open-source model (i.e., providing free products and services for peer review and for the mutual benefit of the peer community) (Bezroukov, 1999). In the case of Wikipedia, the term "open" refers to the fact that any Internet user who has access to the knowledge that other peers produce can freely contribute with knowledge but, at the same time, cannot exert exclusive rights over the collective innovation (Lakhani & Panetta, 2007). The open-source model enables individuals to constantly refine article knowledge through collaboration, which many consider to be one of Wikipedia's main added values.

To help its community write and edit textual content collaboratively, Wikipedia uses a wiki technology. The resulting article pages represent the main source of knowledge that regular Wikipedia readers use. Anyone with Internet access can edit almost all Wikipedia articles. For any user to visualize the dynamics of article changes, each article's revision history chronologically tracks previous versions of articles by time stamp, editor, actual text resulting from the edit, and editor comments. Besides article pages, Wikipedia hosts free-form discussion pages called talk pages for each article. Concretely, editors use talk pages to plan and discuss their work; that is, to support and coordinate their work, share and ask for feedback, report vandalism, or refer to edit guidelines (Schneider, Passant, & Breslin, 2010).

In this sense, sentiment-driven feedback emerges from social interactions that occur in the context of inter-personal relationships (Andersen & Guerrero, 1998). In Wikipedia, social interactions often happen on discussion pages in the form of feedback to others' work. We may classify feedback as either being sentiment driven (positive or negative) or neutral. Table 1 presents examples of sentiment-driven feedback (both positive and negative) and neutral feedback expressed by editors on Wikipedia's "Talk:NASA", "Talk:Water_on_Mars", and "Talk:Jupiter" entries.

Table 1. Examples of Positive, Negative, and Neutral Feedback on Article Talk Pages in Wikipedia

Type of feedback		Example
Sentiment-driven feedback	**Positive**	*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)[2]
	Negative	*Not only is it a bit lengthy, but it has very unusual organization for a wikipedia page... perhaps inappropriately so [...] 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) [3]
Neutral feedback		*In the table at the beginning, the amount of hydrogen stated is 85.8 to 89.8%. The source (Williams, Dr. David R. "Jupiter Fact Sheet", NASA) states that the amount of gas is "89.8% (2.0%)", meaning 89.8% plus or minus 2%* Barras (talk) 18:50, 19 July 2009 (UTC) [4]

Articles from Simple English Wikipedia—a spin-off of Wikipedia written using basic English vocabulary and uncomplicated grammar constructions—are usually not new; their editors use articles from Wikipedia and attempt to bring them to a simple form. The dump we used in this study was collected in October 2011 and contained over 200,000 pages (out of which around 70,000 are article pages). It totaled approximately 16 gigabytes of XML data. Along with anonymous users, over 170,000 registered users have contributed to approximately three million revisions from the creation of Simple English Wikipedia in 2003. The average number of revisions per page was 14.35 and Simple English Wikipedia had over 700 active, registered users with at least one edit or logged action in the past month. Because of the extremely large size of Wikipedia's revision history and the limited computational power, many previous analyses have used only samples of data to save computation costs (Arazy & Nov, 2010; Javanmardi & Lopes, 2010; Muller-Birn, Lehmann, & Jeschke, 2009; Viegas, Wattenberg, Kriss, & van Ham, 2007). However, as no general guideline on how to obtain a good sample from Wikipedia exists and since complete revision histories are necessary for computing revision-based metrics (such as knowledge reuse), we instead use a complete revision history dump (as of October 2011) of Simple English Wikipedia.

3.2 Measurement and Construct Operationalizations

We computed Wikipedia-specific, revision-based measures on peers' editing activity of article pages (knowledge reuse) and mined the affect from feedback posted on article talk pages (sentiment-driven feedback) corresponding to each article. To address H1 and H2, we employed monthly time series analyses of the relationship between knowledge reuse and the amount of (positive and negative) sentiments in inter-editors affective communication. Below, we present the operationalizations and measurements we employed for sentiment-driven feedback and knowledge reuse.

3.2.1 Sentiment-driven Feedback in Inter-editor Communication

We mined sentiment-driven feedback from inter-editor communication on Wikipedia article talk pages. To do so, we applied sentiment analysis to distinguish sentiment-driven feedback from neutral feedback. Below we present the approach in detail. Sentiment analysis broadly classifies textual statements into "objective statements" that express factual information and into "subjective statements" that reflect tndividuals' attitudes or perceptions (Banea, Mihalcea, Janyce, & Samer, 2008; Furuse, Nobuaki, Setsuo, & Ryoji, 2007; Pang & Lee, 2008; Wiebe & Mihalcea, 2006). It represents a systematic, computer-based analysis of written text or speech excerpts for detecting the attitude of the author or speaker about a specific topic (Pang & Lee, 2008). Sentiment analysis establishes the overall orientation (positive or

[2] retrieved from: http://en.wikipedia.org/wiki/Talk:NASA
[3] retrieved from: http://en.wikipedia.org/wiki/Talk:Water_on_Mars
[4] retrieved from: http://simple.wikipedia.org/wiki/Talk:Jupiter

negative) and intensity (weak or strong) of sentiments expressed in statements previously classified as subjective. However, sentiments may often be expressed in a subtle manner, which makes subjectivity analysis often more difficult than subsequent polarity classification, so improvements in subjectivity classification promise to positively impact sentiment classification (Mihalcea, 2007). Because Wikipedia editors often use informal vocabulary when writing comments on talk pages, identifying subjectivity is harder to achieve.

Recent algorithms for analyzing sentiment are able to detect positive and negative sentiment strength in short informal texts (Akkaya, Wiebe, & Mihalcea, 2009; Paltoglou & Thelwall, 2010; Shanmugasundaram, Ramachandran, Murugan, & Saranathan, 2009). In our analyses, we used SentiStrength (Thelwall, Buckley, & Paltoglou, 2011; Thelwall, Buckley, Paltoglou, Cai, & Kappas, 2010) to analyze the level of sentiments on article talk pages in Wikipedia. SentiStrength provides a scoring range from -5 (very negative) to +5 (very positive). In case of texts showing an equal amount of positive and negative sentiments, the algorithm can predict which of the two orientations is the prevalent one. Figure 1 shows the distribution of the total monthly amount of positive sentiments expressed by editors on article talk pages; Figure 2 displays the total monthly amount of negative sentiments expressed by editors on article talk pages.

SentiStrength performs best on short texts, such as Twitter postings (Nielsen, 2011) and Wikipedia article talk pages. An evaluation of SentiStrength on short informal texts from Twitter showed that it performed with 96.9 percent accuracy when detecting positive sentiment strength and 95.1% accuracy when detecting negative sentiment strength (Thelwall et al., 2010). To evaluate the performance of SentiStrength on our data set, we constructed a random set comprising 200 sentences from conversations in Simple English Wikipedia talk pages. To produce a gold standard for evaluating SentiStrength, two independent judges (i.e., annotators) classified the data set in terms of the sentiment scores from -5 (very negative) to +5 (very positive). We asked the annotators to individually send us their results. Out of the 200 sentences, the two annotators agreed on 194 cases. A neutral judge adjudicated the remaining six cases. As such, we obtained the gold standard of human ratings. We then evaluated the results of SentiStrength against the gold standard.

To validate the use of SentiStrength for the sentiment analysis, we used precision (P) and accuracy (A) as evaluation metrics (Menditto, Patriarca, & Magnusson, 2007; Powers, 2011). For classification tasks, we used the terms “true positives”, “true negatives”, “false positives” (type I error), and “false negatives” (type II error) to compare the results of the classifier against the gold standard (Goutte & Gaussier, 2005). The terms “positive” and “negative” refer to the result indicated by the classifier, whereas the terms “true” and “false” refer to whether that result corresponds to the gold standard. Precision is the proportion of correctly labeled examples (i.e., the proportion of the true positives against all the positive results (both true positives and false positives)). Accuracy is the proportion of true results (both true positives and true negatives) in the population. While accuracy is the proximity of measurement results to the true value, precision represents the reproducibility of the measurement (Sandhu, Kaur, & Kaur, 2010). Table 2 summarizes the results of the evaluation we performed on the gold standard with respect to three classes: negative, positive, and the overall sentiment score.

Table 2. Results of Evaluation of SentiStrength Against the Gold Standard

SentiStrength evaluation	Negative sentiment	Positive sentiment	Overall sentiment
Precision (P)	0.83	0.93	0.98
Accuracy (A)	0.97	0.98	0.99

The results were consistent with the results of the evaluation of SentiStrength on Twitter. As such, SentiStrength was suitable for analyzing the sentiment of Simple English Wikipedia’s article talk pages.

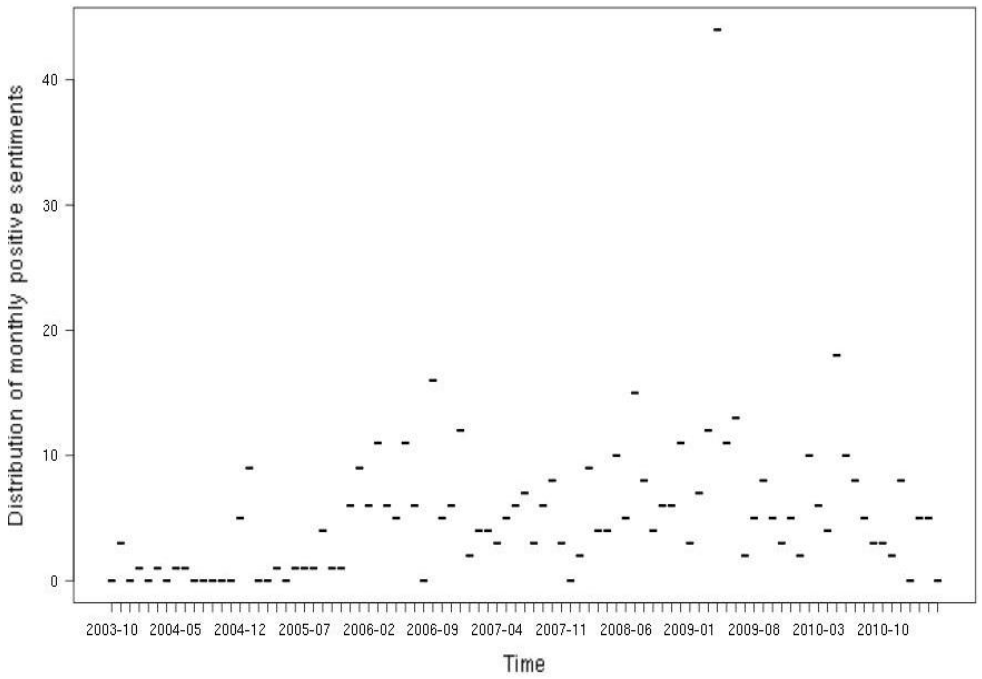

Figure 1. Distribution of Total Amounts of Positive Sentiments on Article Talk Pages Corresponding to Each Month from the Creation of Simple English Wikipedia

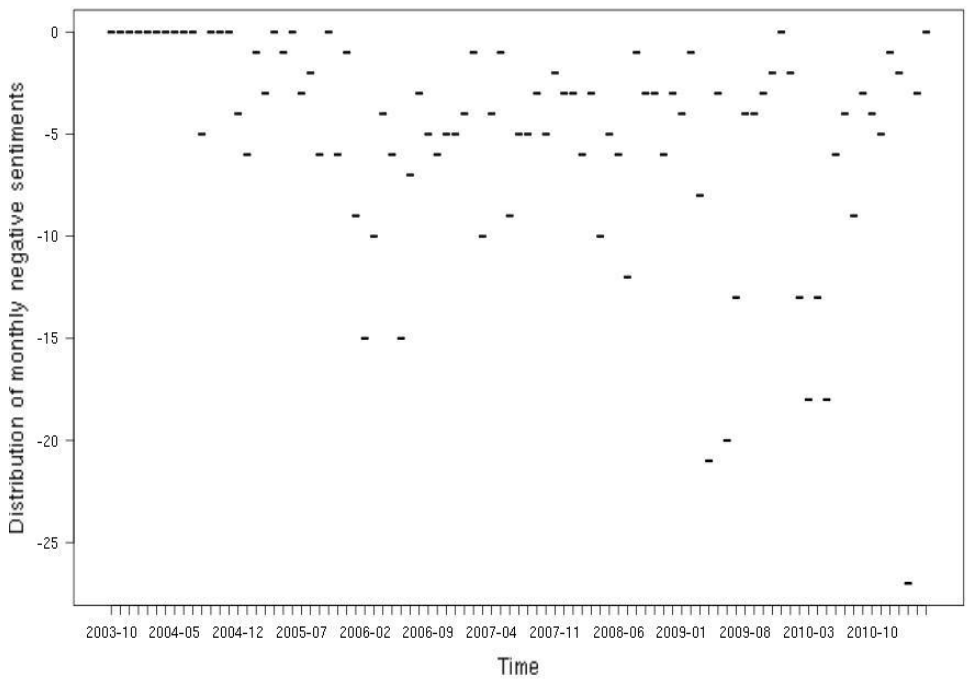

Figure 2. Distribution of Total Amounts of Negative Sentiments on Article Talk Pages Corresponding to Each Month from the Creation of Simple English Wikipedia

3.2.2 Knowledge Reuse

Knowledge reuse enables individuals to repetitively use existing knowledge to create new knowledge. Repetitive use refers to knowledge that is systematically stored in a repository and that can be retrieved

and reused without its reinvention costing anything (Kankanhalli et al., 2011). In an organizational context, knowledge reuse refers to "one individual or group within the firm using knowledge generated by a different individual or group within the same firm in order to be more effective and productive in their work" (Alavi & Leidner, 1999, p. 143; Alavi & Leidner, 2001). In this sense, Majchrzak et al. (2013b) report that companies that use wiki technologies may improve collaboration, work processes, and knowledge reuse. However, quantifying the amount of knowledge being reused is not trivial. One of the settings in which knowledge reuse has been often derived from observable data is software development. Specifically, code reuse in software development, as the name implies, refers to employing previously written code (i.e., objects) as an alternative to writing new (possibly identical) code to perform the same or similar function (Banker, Kauffman, Wright, & Zweig, 1994). Code reuse has been previously measured in terms of the reuse leverage metric. Concretely, the reuse leverage in an application refers to the total number of objects used divided by the number of new objects built (Banker, Kauffman, & Zweig, 1992). For example, if a software application comprises 400 objects (i.e., used objects) of which 100 had to be programmed from scratch (i.e., new objects), the reuse leverage would be 4.0. To indicate how much of a software application can be attributed to reuse (Poulin, 1994), the reuse ratio can be expressed as the ratio of the number of objects that are reused (i.e., 300 = 400-100) to the number of new objects (i.e., 100) (thus, 3:1 = 3.0). Similar to a software application, in the case of Wikipedia, an article represents a dynamic and systematic transformation of existing knowledge. Below, we describe how we quantified knowledge reuse in the context of Wikipedia based on the reuse ratio.

On Wikipedia, the ongoing process of knowledge reuse is facilitated by the wiki technology and is captured by the revision history functionality. The revision history chronologically tracks all the previous versions of an article. To contribute to the development of an article, an editor starts editing the knowledge already contributed in the current revision. In other words, an editor reuses existing knowledge in the previous article revision to create a new article revision. One can quantify the amount of knowledge reuse between two consecutive article revisions as the amount of knowledge from the previous revision that the current revision reuses. Similar to reuse ratio in software development (Banker et al., 1992) and inspired by Turek, Wierzbicki, and Nielek's (2010) work on Wikipedia revision history metrics, we computed the level of knowledge reuse relative to any two consecutive revisions of the same article page as the ratio of the number words reused from the previous revision (e.g., copied, moved elsewhere, or restored) to the number of words newly created in the current revision.

For a given Wikipedia article page, we computed the overall level of knowledge reuse between its revisions as a mean of pairwise levels of knowledge reuse weighted by the proportion of editors who contributed to creating each specific revision from the total number of article editors. In this way, we accounted for the effect of the number of edits and editorial team size, which vary from article to article. A value of 1 indicates equal amount of reused and new words. A value of 2 (i.e., 2:1) indicates that the number of reused words was two times larger than the number of newly created words. Conversely, a value of 0.5 (i.e., 1:2) indicates that the number of reused words was two times larger than the number of newly created words. Figure 3 shows the distribution of the monthly average amount of knowledge reuse extracted from the article pages' edit logs.

We employed two measurement considerations in computing knowledge reuse. The first consideration refers to adapting the measurement of knowledge reuse from software development (Banker et al., 1992). Since reuse in software development employs code (i.e., objects) as units of analysis, reuse in Wikipedia should conversely employ individual key concepts or key phrases expressed in Wikipedia articles as the units of analysis. However, unlike in the case of software development where the code is automatically compiled, the automatic recognition of domain-specific key concepts or phrases from natural language text is not a trivial task in the case of Wikipedia (Boudin & Morin, 2013; Erbs, Santos, Gurevych, & Zesch, 2014). Thus, we decided to measure knowledge reuse in Wikipedia articles at the word level. To avoid considering common words that are not informative, we employed linguistic processing that comprised three steps. First, we divided the textual content into individual strings of characters called tokens. Second, for each token representing a word, we reduced inflected words to their word stem (such as the verbs "represents" and "representing" to "represent"). Finally, we removed non-informative words that tend to occur very often (i.e., employed stop-words filtering), such as "a", "the", "is", "an", "in", "it", "that", and so on. In addition, to measure knowledge reuse, we discarded from the count consecutive word revisions for which the corresponding edit distance was smaller or equal to one (Gonzalo, 2001), which we did because such a small edit distance indicates that the word underwent a minor spelling correction rather than a substantial change in its meaning. Altogether, this linguistic processing reduces the potential pitfall of choosing individual words as the level of analysis for measuring knowledge reuse. Relative to the

second adaptation used in this study, the way we measured knowledge reuse does not penalize repetitive consecutive edits done by the same peers (Turek et al., 2010). We adopted this approach because it is difficult to establish clear criteria on how to group consecutive revisions made by the same editors. Since the time elapsed between consecutive edits done by the same editor may vary up to months or years, we decided to treat every revision as standalone and compute the knowledge reuse with respect to the previous revision.

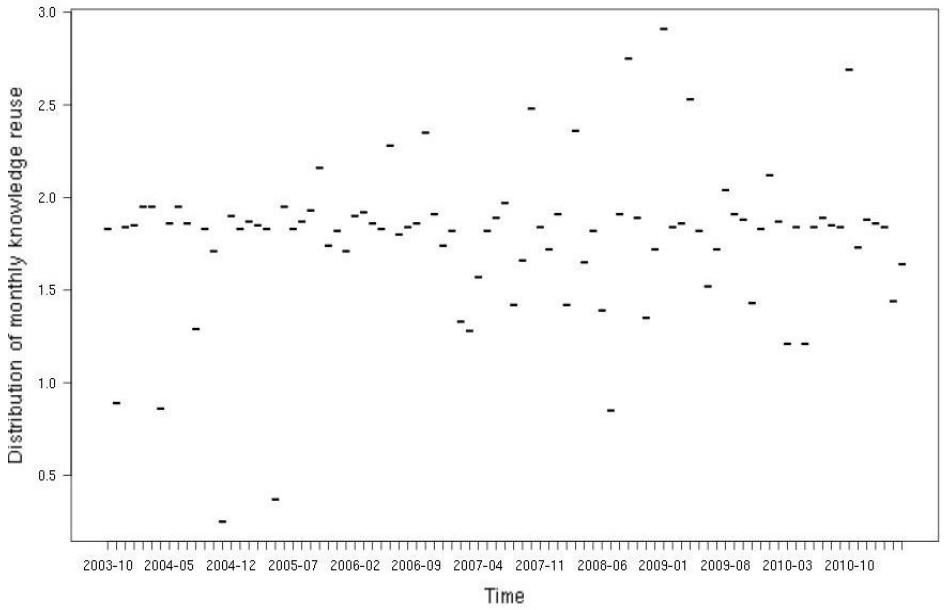

Figure 3. Distribution of Monthly Average Amount of Knowledge Reuse Corresponding to the Edits of Article Pages in Simple English Wikipedia

4 Analysis and Results

To test hypotheses H1 and H2, we examined both the distribution of sentiments on article pages and level of knowledge reuse between editors on the corresponding article pages. Similar to Turek et al. (2010), we used the number of edits on an article as a proxy to avoid considering articles that were in their very initial stages of development. After analyzing the distribution of edits, similar to Turek et al. (2010), we decided to use a threshold of 30 edits for analyzing those articles whose knowledge is a result of collaboration among their editors. For those article pages having an overall number of revisions greater than this threshold, we computed the distribution of sentiments over time on their corresponding article talk pages. To prepare our dataset for hypothesis testing, we first performed a time series analysis of all article pages with respect to the amount of both knowledge reuse and sentiments expressed during conversations among their editors. Following our analyses of Wikipedia article pages and article talk pages, we grouped article pages by the following criteria:

- **Presence or absence of sentiments**: we grouped article pages containing subjective (S) statements on the corresponding article talk pages in an S cluster (1239 article pages) and the ones containing only objective (O) statements in an O cluster (742 article pages).
- **Positive versus negative sentiments**: we grouped each article page belonging to the S cluster into either a P cluster if the main orientation of sentiments displayed in the corresponding talk page was positive (P, 794 articles) or an N cluster if the orientation was mainly negative (N, 445 articles). To decide on the main orientation of the sentiments, for all the statements on a talk page, we compared the sum of positive sentiment strengths with the corresponding negative one; we decided the orientation based on the category corresponding to the higher sum.

To quantify whether the presence or absence of subjectivity in the content of article talk pages influenced the level of knowledge reuse, we compute monthly averages of the level of knowledge reuse for each S and O cluster according to their article pages' revision flows. Figure 4 shows the results. We first compare

the obtained discrete distributions of the levels of knowledge reuse. On the one hand, the median level of knowledge reuse in the case of sentiment-driven feedback (S) appeared to be close to 2, which indicates that the number of reused words was almost twice as much as the number of newly created words. On the other hand, in the case of neutral feedback (O), the sub-unitary level of knowledge reuse shows that the number of newly created words appeared to overcome the number of reused words. As expected, the median level of knowledge reuse computed for the O cluster was significantly lower than the median of the S cluster (Mann-Whitney U test; $Z = -10.26$, $p<0.001$, $r = 0.76$). A linear regression showed that the presence and absence of sentiments significantly explained the level of knowledge reuse (*adj.* R^2 $=0.6322$, $p < 0.001$), which confirms H1. In other words, as expected, providing feedback appears to be beneficial for knowledge reuse in Wikipedia.

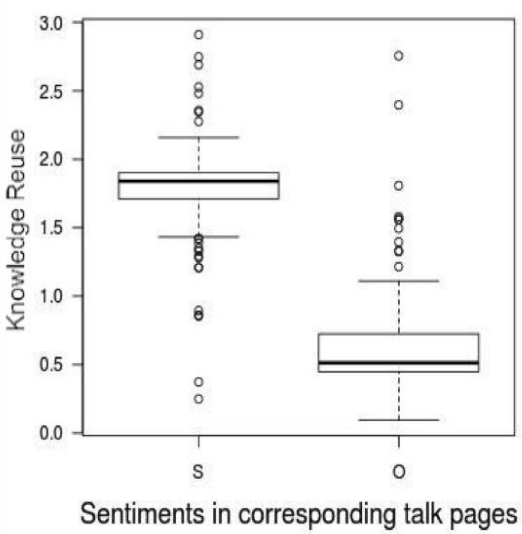

Figure 4. Distribution of Knowledge Reuse Scores Computed for Each of the Two Subject Groups According to the Sentimentality (S) or the Objectivity (O) of Content on Their Corresponding Article Talk Pages

For the articles contained in the S cluster, we compare mainly positive (P) and mainly negative (N) orientation of subjective content on article talk pages. Figure 5 shows the results. The median level of reuse for the article pages of the P-cluster was significantly higher than the one corresponding to the N cluster (Mann-Whitney U test; $Z=-8.61$, $p<0.001$, $r=0.64$). The mainly positive or mainly negative orientation of subjectivity also explained a significant proportion of variance in the level of knowledge reuse using a linear regression (*adj.* $R^2 = 0.4284$, $p < 0.001$). This means that, indeed, providing positive feedback appears to be more effective in Wikipedia with respect to knowledge reuse than providing negative feedback. The median level of knowledge reuse in the case of positive feedback (P) appears to be more than two times higher than in the case of negative feedback (N). This result further confirms H2.

In a nutshell, our results confirm that receiving (especially positive, rather than negative) feedback in form of sentiments that are expressed in inter-editor conversations is beneficial in terms of sustaining knowledge reuse in Wikipedia; moreover, giving either positive or negative feedback appears to be more effective than providing no feedback at all.

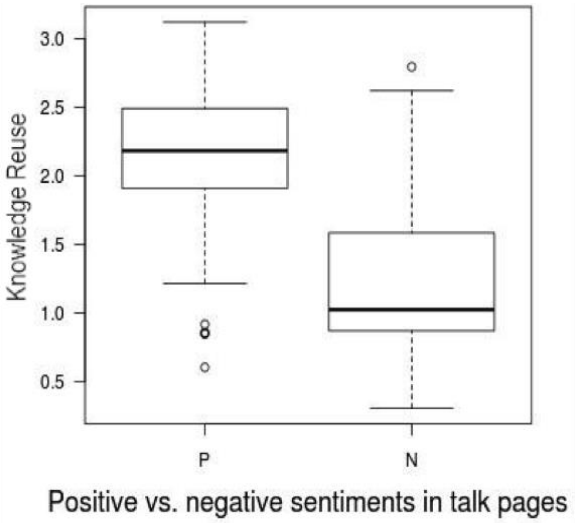

Figure 5. Distribution of Knowledge Reuse Scores of Articles Classified in the S Category Grouped by the Positivity (P) or the Negativity (N) of the Content on Their Corresponding Article Talk Pages

5 Discussion and Implications

In this paper, by conducting a longitudinal analysis of peers' editing and interaction activity in Simple English Wikipedia, we investigate how peers are motivated to reuse knowledge in online social production communities. In line with Markus's (2001) findings on knowledge reuse, we consider knowledge reuse among Wikipedia editors as an essential aspect of collaboration that is influenced by specific communication practices. Specifically, we examine the way peers in Wikipedia communicate their feedback and support towards the work of other peers; we quantify to what extent providing sentiment-driven feedback acts as incentive towards the reuse of knowledge among peers. We found that peer content collaboration in Wikipedia in terms of higher levels of knowledge reuse appears to be strongly influenced by either positive or negative sentiment-driven feedback in inter-editor discussions. We also found a significant difference in the level of knowledge reuse between editors who share mainly positive or mainly negative feedback. Indeed, displaying mainly positive sentiments in form of peer feedback corresponded to a superior level of knowledge reuse than displaying mainly negative sentiments, which suggests that making a habit of dispensing positive feedback in public conversations (such as in the case of Wikipedia) is more likely to motivate peers to perform with confidence and autonomy, than giving negative feedback. At the same time, non-public negative peer feedback could increase one's likelihood to engage in online social production by correcting inherent problems, behaviors, and attitudes in private peer conversations, which also strongly suggests that mechanisms for providing non-public negative feedback should be designed, incorporated, and tested in collaborative platforms such as wikis. Thus, we contribute to the extant literature of online social production communities in general and Wikipedia in particular by providing the first building block for research on the role of sentiment-driven feedback for developing and sustaining wiki-based knowledge reuse.

One should view our findings in light of several limitations. With regard to generalizability and endogeneity, we acknowledge that we did not examine several areas dealing with the dynamics of social interaction in online collaboration, such as the issues of social power or culture (Baym, 2006; Jiang, Bazarova, & Hancock, 2011). Pragmatically, one can take several other perspectives for examining the role of sentiment-driven feedback in online social production communities. Conditions other than peer feedback—such as group interactions, group composition characteristics, peers' capabilities and goals, their interpretations of technology, and institutional contexts, power, or culture—may play key roles in causal explanations. Due to the nature of our observational data and the possibility of simultaneous causality bias and errors-in-variable bias, future research should examine our identified relationships using more controlled settings or methods such as instrumental variables regression or controlled experiments. Moreover, we discuss the limitations of measuring knowledge reuse at word level are discussed in detail in Section 3.2. Researchers could further address these limitations by employing individual key concepts or key phrases as unit of analysis when computing the level of knowledge reuse. Furthermore, no clear guidelines on how to select the articles for analysis exist; similar to Turek et al.

(2010), we decided to use a threshold of 30 edits for analyzing those articles that result from peer collaboration. We acknowledge that this choice may introduce a bias in selecting only those articles where editors appeared motivated. To filter out articles that are not a result of peer collaboration, one could use alternative methods such as filtering out either relatively new articles based on their age or those articles that have not received enough attention from the community based on the number of times they have been read. Finally, SentiStrength's performance limits how accurately we identified sentiments in the article talk pages; we manually evaluate sentiment analysis algorithms in Section 3.2. Future research may want to use and compare alternative approaches from sentiment analysis to perform a more detailed comparison and benchmark of the results.

Nonetheless, by examining the talk pages of *all* articles in Simple English Wikipedia (we performed no sampling), our study provides important insights for the literature on online social production. Thereby, we provide the first building block for research on how to understand peer collaboration in online social production communities in terms of knowledge reuse. Wikipedia provides an unprecedented amount of data that enabled us to 1) fully use the information provided by the edit history to quantify the amount of knowledge reuse and 2) exploit the multitude of informal language to identify subjectivity in the content of article talk pages. Wikipedia is an ideal environment for studying the cumulative effect of social and sentiment-driven interactions among editors on collaborative work. By extending previous work on knowledge reuse, our study contributes to the existing research on online social production along several dimensions of interests to researchers and practitioners.

From a theoretical perspective, the collective ethics of online social production appears to be in conflict with traditional policies, perceptions, and theories of organizational work (Arvidsson, 2008; Banks & Deuze, 2009; Puranam et al., 2013; Sanger, 2009). Indeed, social production systems raise a series of challenges for traditional organization in so far as researchers have shown that peers do not necessarily follow the normal signals generated by firms or markets either as employees in the firms following managerial directions or as individuals in the markets following price signals (Benkler & Nissenbaum, 2006; Tapscott & Williams, 2006). In this sense, a micro-foundation of peer production is important to develop up-to-date theoretical concepts for management and organizational sciences. To design efficient policies that boost an innovative, networked economy, we need a systematic empirical analysis and an empirically grounded theoretical understanding of knowledge processes in peer production. Relative to the focus of this paper, knowledge reuse in online communities continues to be under-researched (Haefliger et al., 2008; von Krogh et al., 2012). Although Majchrzak, Faraj, Kane, and Azad (2013a) provide comprehensive theorizing of how peers engage in knowledge sharing via online knowledge conversations, we still lack a theoretical understanding of technology-enabled knowledge reuse in online communities. This exploratory study helps to discover strategies to encourage collaboration and foster knowledge reuse in online communities and make the crowd sustainable without relying either on markets or hierarchies (Metiu & Kogut, 2001; Stephen & Suzanne, 2006).

Our results open a link to further controlled studies such as experiments observing the affective implication of individuals who reuse content. Researchers may 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 online social production in the context of another similar and rapidly growing resource, Wikia. With more than 370,000 established communities, Wikia is fundamentally different from the Wikipedia community in terms of having more permissive guidelines and policies and a high number of small, topic-centered communities. Another direction worth investigating would be to analyze and test patterns of social collaboration in communities of open source software development (OSS) (e. g., Linux, Apache, GitHub, or SourceForge).

From a managerial perspective, organizations increasingly consider outsourcing knowledge tasks to large masses of workers via distributed labor networks using limited or no monetary incentives, which is possible, in part, due to the fact that the virtual, self-organizing workplace constantly evolves towards more spontaneous and decentralized forms of collaboration. While open collaborative innovation can potentially displace producer innovation at many parts of the economy (Baldwin & von Hippel, 2011; Maiolini & Naggi, 2011), the fluid generativity of distributed innovation suggests that knowledge resources will be increasingly heterogeneous and often only temporarily integrated (Yoo, Boland, Lyytinen, & Majchrzak, 2012). Reflecting from the Wikipedia case, having insights about practical mechanisms to motivate the refinement of collectively produced knowledge resources is important for organizations that would like to outsource knowledge tasks to large masses of online distributed workers. This study

provides insights that sentiment-driven feedback appears to be an effective way to motivate collaborative work in online social production.

Acknowledgements

This research received financial support from Swiss National Science Foundation Grant No. 100018_146444.

References

Aaltonen, A., & Lanzara, G. (2011). *Governing complex social production in the Internet: The emergence of a collective capability in Wikipedia.* Paper presented at the Decade in Internet Time Symposium.

Aggarwal, R., Gopal, R., Sankaranarayanan, R., & Singh, P. V. (2012). Blog, blogger, and the firm: Can negative employee posts lead to positive outcomes? *Information Systems Research*, *23*(2), 306-322.

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology*, *52*(1), 27-58.

Akkaya, C., Wiebe, J., & Mihalcea, R. (2009). Subjectivity word sense disambiguation. In *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing (ACL)*.

Alavi, M., & Leidner, D. (1999). Knowledge management systems: Issues, challenges, and benefits. *Communications of the AIS*, *1*, 1-37.

Alavi, M., & Leidner, D. E. (2001). Review: Knowledge management and knowledge management systems: Conceptual foundations and research issues. *MIS Quarterly*, *25*(1), 107-136.

Andersen, P., & Guerrero, L. (1998). Principles of communication and emotion in social interaction. In P. Andersen & L. Guerrero (Ed.), *Handbook of communication and emotion*. New York Academic Press.

Arazy, O., & Nov, O. (2010). Determinants of wikipedia quality: The roles of global and local contribution inequality. In *Proceedings of the 2010 ACM Conference on Computer Supported Cooperative Work*. New York, NY, USA: ACM.

Argote, L., McEvily, B., & Reagans, R. (2003). Managing knowledge in organizations: An integrative framework and review of emerging themes. *Management science*, *49*(4), 571-582.

Argote, L., & Miron-Spektor, E. (2011). Organizational learning: From experience to knowledge. *Organization Science*, *22*(5), 1123-1137.

Arvidsson, A. (2008). The ethical economy of customer coproduction. *Journal of Macromarketing*, *28*(4), 326-338.

Baldwin, C., & von Hippel, E. (2011). Modeling a paradigm shift: From producer innovation to user and open collaborative innovation. *Organization Science*, 22(6), 1399-1417.

Banea, C., Mihalcea, R., Janyce, W., & Samer, H. (2008). Multilingual subjectivity analysis using machine translation. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*.

Banker, R., Kauffman, R., & Zweig, D. (1992). Automating output size and reuse metrics in a repository-based computer aided software engineering (CASE) environment. *Information Systems Working Papers Series*.

Banker, R. D., Kauffman, R. J., Wright, C., & Zweig, D. (1994). Automating Output Size and Reuse Metrics in a Repository-Based Computer-Aided Software Engineering (CASE) Environment. *IEEE Transactions on Software Engineering*, *20*(3), 169-187.

Banks, J., & Deuze, M. (2009). Co-creative labour. *International Journal of Cultural Studies*, *12*(5), 419-431.

Barrett, L., & Russell, J. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science*, *8*(1), 10-14.

Bateman, T. S., & Organ, D. W. (1983). Job satisfaction and the good soldier: The relationship between affect and employee “citizenship”. *Academy of Management Journal*, *26*(4), 587-595.

Baym, N. K. (2006). Interpersonal life online. *Handbook of new media: social shaping and social consequences of ICTs*: 35-54.

Beaudry, A., & Pinsonneault, A. (2010). The other side of acceptance: Studying the direct and indirect effects of emotions on information technology use. *MIS Quarterly*, *34*(4), 689-710.

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. London: Yale University Press.

Benkler, Y., & Nissenbaum, H. (2006). Commons based peer production and virtue. *Journal of Political Philosophy*, *14*(4), 394-419.

Bezroukov, N. (1999). Open source software development as a special type of academic research. *First Monday*, *4*(10). Retrieved from http://firstmonday.org/article/view/696/606

Boudin, F., & Morin, E. (2013). *Keyphrase extraction for n-best reranking in multi-sentence compression*. Paper presented at the Proccedings of the NAACL HLT 2013 Conference.

Burnett, G. (2000). Information exchange in virtual communities: A typology. *Information Research*, *5*(4). Retrieved from http://www.informationr.net/ir/5-4/paper82.html

Carlile, P. R. (2002). A pragmatic view of knowledge and boundaries: boundary objects in new product development. *Organization Science*, *13*(4), 442-455.

Carlile, P. R. (2004). Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries. *Organization Science*, *15*(5), 555-568.

Carlile, P. R., & Rebentisch, E. S. (2003). Into the black box: The knowledge transformation cycle. *Management Science*, *49*(9), 1180-1195.

Chi, E. H., Pirolli, P., Suh, B., Kittur, A., Pendleton, B. A., & Mytkowicz, T. (2008). *Augmented social cognition.* Paper presented at the AAAI Spring Symposium: Social Information Processing.

Sandhu, P. S., Kaur, M., & Kaur, A. (2010). A density based clustering approach for early detection of fault prone modules. In *Proceedings of the International Conference On Electronics and Information Engineering.*

Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., & Siddharth, S. (2008). *Feedback effects between similarity and social influence in online communities*. Paper presented at the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Las Vegas, Nevada, USA.

Crowston, K., Li, Q., Wei, K., Eseryel, U. Y., & Howison, J. (2007). Self-organization of teams for free/libre open source software development. *Information and software technology*, *49*(6), 564-575.

Erbs, N., Santos, P. B., Gurevych, I., & Zesch, T. (2014). DKPro keyphrases: Flexible and reusable keyphrase extraction experiments. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics* (pp. 31-36).

Eseryel, U. Y. (2009). Leadership in a non-traditional setting: Self-managing virtual IS development teams. In *Proceedings of the International Conference on Information Systems*.

Faraj, S., Jarvenpaa, S. L., & Majchrzak, A. (2011). Knowledge collaboration in online communities. *Organization Science*, *22*(5), 1224-1239.

Feller, J., Finnegan, P., Hayes, J., & O'Reilly, P. (2008). Governance structures for open innovation: A preliminary framework. In G. Leon, A. Bernardos, J. Casar, K. Kautz, & J. DeGross (Eds.), *Open IT-based innovation: Moving towards cooperative IT transfer and knowledge Diffusion* (pp. 511-525). New York: Springer.

Feng, J., Lazar, J., & Preece, J. (2004). Empathy and online interpersonal trust: A fragile relationship, *Behaviour & Information Technology*, *23*(2), 97-106.

Fishbach, A., Eyal, T., & Finkelstein, S. R. (2010). How positive and negative feedback motivate goal pursuit. *Social and Personality Psychology Compass*, *4*(8), 517-530.

Furuse, O., Nobuaki, H., Setsuo, Y., & Ryoji, K. (2007). Opinion sentence search engine on open-domain blog. In *Proceedings of the 20th International Joint Conference on Artifical intelligence*.

Ghosh, R. A. (2005). *CODE: Collaborative ownership and the digital economy.* Cambridge, MA: The MIT Press.

Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*, *438*(7070), 900-901.

Ginsburg, M., & Weisband, S. (2002). *Social capital and volunteerism in virtual communities: The case of the Internet chess club.* Paper presented at the 35th Annual Hawaii International Conference on System Sciences, Hawaii, USA.

Gonzalo, N. (2001). A guided tour to approximate string matching. *ACM Computing Survey*, *33*(1), 31-88.

Goutte, C., & Gaussier, E. (2005). A probabilistic interpretation of precision, recall and F-score, with implication for evaluation. In D. E. Losada & J. M. Fernandez-Luna (Eds.), *Advances in information retrieval* (pp. 345-359). New York: Springer.

Grant, R. M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*, *17*, 109-122.

Gray, P. H. (2001). A problem-solving perspective on knowledge management practices. *Decision Support Systems*, *31*(1), 87-102.

Grudin, J., & Poole, E. S. (2010). *Wikis at work: Success factors and challenges for sustainability of enterprise wikis*. Paper presented at the Proceedings of the 6th International Symposium on Wikis and Open Collaboration.

Haefliger, S., Monteiro, E., Foray, D., & von Krogh, G. (2011). Social software and strategy. *Long Range Planning*, *44*(5), 297-316.

Haefliger, S., von Krogh, G., & Spaeth, S. (2008). Code reuse in open source software. *Management Science*, *54*(1), 180-193.

Hahn, J., Moon, J. Y., & Zhang, C. (2008). Emergence of new project teams from open source software developer networks: Impact of prior collaboration ties. *Information Systems Research*, *19*(3), 369-391.

Harris, R. B., & Paradice, D. (2007). An investigation of the computer-mediated communication of emotions. *Journal of Applied Sciences Research*, *3*, 2081-2090.

Javanmardi, S., & Lopes, C. (2010). Statistical measure of quality in Wikipedia. In *Proceedings of the First Workshop on Social Media Analytics*.

Jiang, L., Bazarova, N. N., & Hancock, J. T. (2011). From perception to behavior: Disclosure reciprocity and the intensification of intimacy in computer-mediated communication. *Communication Research*, 40(1), 125-143.

Kankanhalli, A., Lee, O.-K. D., & Lim, K. H. (2011). Knowledge reuse through electronic repositories: A study in the context of customer service support. *Information & Management*, *48*(2), 106-113.

Khodakarami, F., & Chan, Y. E. (2014). Exploring the role of customer relationship management (CRM) systems in customer knowledge creation. *Information & Management*, *51*(1), 27-42.

Lakhani, K. R., & Panetta, J. A. (2007). The principles of distributed innovation. *Innovations: Technology, Governance, Globalization*, *2*(3), 97-112.

Lanubile, F., Ebert, C., Prikladnicki, R., & Vizcaíno, A. (2010). Collaboration tools for global software engineering. *IEEE Software, 27*(2), 52-55.

Leimeister, J. M., Sidiras, P., & Krcmar, H. (2006). Exploring success factors of virtual communities: The perpesctives of members and operators. *Journal of Organizational Computing and Electronic Commerce*, *16*(3-4), 277-298.

Lerner, J., & Tirole, J. (2002). Some simple economics of open source. *The Journal of Industrial Economics*, *50*(2), 197-234.

Lickerman, A. (2012). *The undefeated mind: On the science of constructing an indestructible self.* Deerfield Beach, FL: HCI Books.

Liu, H., Chai, K.-H., & Nebus, J. F. (2013). Balancing codification and personalization for knowledge reuse: A Markov decision process approach. *Journal of Knowledge Management*, *17*(5), 755-772.

Maiolini, R., & Naggi, R. (2011). Crowdsourcing and SMEs: Opportunities and challenges. In A. D'Atri, M. Ferrara, J. F. George, & P. Spagnoletti (Eds.), *Information technology and innovation trends in organizations.* New York: Pyhsica-Verlag.

Majchrzak, A., Faraj, S., Kane, G., & Azad, B. (2013a). The contradictory influence of social media affordances on online knowledge sharing. *Journal of Computer Mediated Communication*, 19, 38-55.

Majchrzak, A., Wagner, C., & Yates, D. (2013b). The impact of shaping on knowledge reuse for organizational improvement with wikis. *MIS Quarterly*, *37*(2), 455-469.

Maloney-Krichmar, D., & Preece, J. (2005). A multilevel analysis of sociability, usability, and community dynamics in an online health community. *ACM Transactions on Computer-Human Interaction*, *12*(2), 201-232.

Markus, M. L. (2001). Toward a theory of knowledge reuse: Types of knowledge reuse situations and factors in reuse success. *Journal of Management Information Systems*, *18*(1), 57-94.

Mehra, S. (2012). Platforms, teamwork and creativity: Mediating hierarchs in the new economy. *Journal of Law, Economics, & Policy*, *9*, 15-29.

Menditto, A., Patriarca, M., & Magnusson, B. (2007). Understanding the meaning of accuracy, trueness and precision. *Accreditation and Quality Assurance*, *12*(1), 45-47.

Metiu, A., & Kogut, B. (2001). *Distributed knowledge and the global organization of software development* (working paper). Philadelphia, PA: Wharton School, University of Pennsylvania.

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*.

Muller-Birn, C., Lehmann, J., & 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* (vol. 1). New York: IEEE Computer Society.

Nelson, M. M., & Schunn, C. D. (2009). The nature of feedback: How different types of peer feedback affect writing performance. *Instructional Science*, *37*(4), 375-401.

Nielsen, F. Å. (2011). A new ANEW: Evaluation of a word list for sentiment analysis in microblogs. In *Proceedings of the ESWC2011 Workshop on "Making Sense of Microposts"* (pp. 93-98).

Nonaka, I., & Takeuchi, H. (1995). T*he knowledge-creating company: How Japanese companies create the dynamics of innovation*. New York: Oxford University Press.

Paltoglou, G., & 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., & Lee, L. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval, 2*(1-2), 1-135.

Porra, J., & Parks, M. S. (2006). Sustainable virtual communities: Suggestions from the colonial model. *Information Systems and E-Business Management*, *4*(4), 309-341.

Poulin, J. S. (1994). Measuring software reusability. In *Proceedings of Software Reuse: Advances in Software Reusability* (pp. 126-138). IEEE.

Powers, D. M. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, *2*(1), 37-63.

Preece, J. (2001). Sociability and usability in online communities: Determining and measuring success. *Behaviour & Information Technology*, *20*(5), 347-356.

Puranam, P., Alexy, O., & Reitzig, M. (2013). What's "new" about new forms of organizing? *Academy of Management Review*, *39,* 162-180.

Rafaeli, S., & Ariel, Y. (2008). Online motivational factors: Incentives for participation and contribution in Wikipedia. In A. Barak (Ed.), *Psychological aspects of cyberspace: Theory, research, applications* (pp. 243-267). Cambridge, UK: Cambridge University Press.

Reagle, J. M. (2010). "Be nice": Wikipedia norms for supportive communication. *New Review of Hypermedia and Multimedia*, *16*(1-2), 161-180.

Rozwell, C. (2009). Socialization of knowledge management drives greater reuse. *Gartner Research Note.*

Sanger, L. (2009). The fate of expertise after Wikipedia. *Episteme*, *6*(1), 52-73.

Scherer, K., Wranik, T., Sangsue, J., Tran, V., & Scherer, U. (2004). Emotions in everyday life: Probability of occurrence, risk factors, appraisal and reaction patterns. *Social Science Information*, *43*(4), 499-570.

Schneider, J., Passant, A., & Breslin, J. (2010). A qualitative and quantitative analysis of how Wikipedia talk pages are used. In *Proceedings of the WebSci10: Extending the Frontiers of Society*.

Shah, S. K. (2006). Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, *52*(7), 1000-1014.

Shanmugasundaram, H., Ramachandran, S., Murugan, S., & Saranathan, P. (2009). Opinion mining and summarization of reviews in Web forums. In *Proceedings of the 3rd Annual ACM Bangalore Conference*.

Shirky, C. (2010). *Cognitive surplus: Creativity and generosity in a connected age*. New York: Penguin.

Singh, P. V. (2010). The small-world effect: The influence of macro-level properties of developer collaboration networks on open-source project success. *ACM Transactions on Software Engineering and Methodology, 20*(2), 1-27.

Skopik, F., Truong, H.-L., & Dustdar, S. (2009). Trust and reputation mining in professional virtual communities. In M. Gaedke, M. Grossniklaus, & O. Díaz (Eds.), *Web engineering* (vol. 5648, pp. 76-90). New York: Springer.

Spoor, J. R., & Kelly, J. R. (2004). The evolutionary significance of affect in groups: Communication and group bonding. *Group Processes & Intergroup Relations*, *7*(4), 398-412.

Stahlbrost, A., & Bergvall-Kareborn, B. (2011). Exploring users motivation in innovation communities. *International Journal of Entrepreneurship and Innovation Management*, *14*(4), 298-314.

Stephen, M. M., & Suzanne, S. (2006). *Open source software: The new intellectual property paradigm*. National Bureau of Economic Research (NBER Working Paper No. 12148). Natuonal Bureau of Economic Research.

Stewart, K. J., & Gosain, S. (2006). The impact of ideology on effectiveness in open source software development teams. *MIS Quarterly*, *30*(2), 291-314.

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and information diffusion in social media—sentiment of microblogs and sharing behavior. *Journal of Management Information Systems*, *29*(4), 217-248.

Tapscott, D., & Williams, A. D. (2006). *Wikinomics: How mass collaboration changes everything.* New York: Portfolio Hardcover.

Tarmizi, H., & Vreede, G.-J. d. (2005). *A facilitation task taxonomy for communities of practice*. Paper presented at the 11th Americas Conference on Information Systems, Omaha, NE, USA.

Te'eni, D. (2001). Review: A cognitive-affective model of organizational communication for designing IT. *MIS Quarterly*, 25(2), 251-312.

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406-418.

Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., & Kappas, A. (2010). Sentiment in short strength detection informal text. *Journal of the American Society for Information Science and Technology*, *61*(12), 2544-2558.

Turek, P., Wierzbicki, A., & Nielek, R. (2010). WikiTeams: Evaluating teamwork in Wikipedia. In *Proceedings of WebSci10.*

Van der Heijden, H. (2004). User acceptance of hedonic information systems. *MIS Quarterly*, *28*(4), 695-704.

VandenBos, G. R. (Ed.). (2006). *APA dictionary of psychology*. Washington, DC: American Psychological Association.

Viegas, F. B., Wattenberg, M., Kriss, J., & van Ham, F. (2007). Talk before you type: Coordination in Wikipedia. In *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*.

von Hippel, E., & von Krogh, G. (2003). Open source software and the "private-collective" innovation model: Issues for organization science. *Organization Science*, *14*(2), 209-223.

von Krogh, G., Haefliger, S., Spaeth, S., & Wallin, M. W. (2012). Carrots and rainbows: Motivation and social practice in open source software development. *MIS Quarterly*, *36*(2), 649-676.

Weber, S. (2004). *The success of open source.* Cambridge, UK: Cambridge University Press.

Wiebe, J., & Mihalcea, R. (2006). Word sense and subjectivity. In *Proceedings of the 21st International Conference on Computational Linguistics.*

Yates, D., Wagner, C., & Majchrzak, A. (2010). Factors affecting shapers of organizational wikis. *Journal of the American Society for Information Science and Technology*, *61*(3), 543-554.

Yoo, Y., Boland, R. J., Lyytinen, K., & Majchrzak, A. (2012). Organizing for innovation in the digitized world. *Organization Science*, *23*(5), 1398-1408.

Zhang, P. (2013). The affective response model: A theoretical framework of affective concepts and their relationships in the ICT context. *MIS Quarterly*, 37(1), 247-274.

Zhu, H., Zhang, A., He, J., Kraut, R. E., & Kittur, A. (2013). *Effects of peer feedback on contribution: A field experiment in Wikipedia.* Paper presented at the Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.

About the Authors

Mihai Grigore obtained his PhD degree in Management of Information Systems from the Swiss Federal Institute of Technology Zurich (ETH Zurich, Switzerland), working on the strategic implications of digital content in online communities and companies. His research focuses on digital strategy, big data analytics, strategic information systems, information aggregation and markets, and digital business models. He currently works as big data and smart analytics expert in the financial services and information technology industries.

Christoph Rosenkranz is Full Professor of Information Systems at University of Cologne. He holds Diploma and doctoral degrees from the University of Münster and Goethe University, respectively. His research focuses on integrated information systems as socio-technical systems, business process management, information systems development and IT project management, and online communities. He has published papers in such outlets as *ACM Transactions on Management Information Systems*, *Business & Information Systems Engineering*, *Information Systems Journal*, *Journal of Database Management*, *Journal of Information Technology*, and *Journal of the Association for Information Systems*.

Juliana Sutanto is a Professor for Information Systems at Lancaster University Management School in Lancaster University. Her research focuses on artifact design and behavioral analysis in digital communications and interactions; as such, she subscribes herself to the design and behavioral sciences paradigms. She won INFORMS ISS Design Science Award 2013 for an empirical work on privacy-safe personalized offering. Her research work has been published in leading information systems journals, such as *Management Science*, *MIS Quarterly*, *Information Systems Research*, *Journal of Management Information Systems*, *IEEE Transactions on Engineering Management,* and others.

Transactions on Human – Computer Interaction

Editors-in-Chief

http://thci.aisnet.org/