

## **WebICs as a new information tool for governments?**

Ulrik Brandes<sup>1</sup>, Patrick Kenis<sup>2,3</sup>, Denise van Raaij<sup>2</sup>, Juergen Lerner<sup>1</sup>, Joerg Raab<sup>2</sup>

Paper to be presented at the Public Management Research Conference, Columbus/Ohio,  
October 1 – October 3, 2009

<sup>1</sup> University of Konstanz, Germany

<sup>2</sup> Tilburg University, The Netherlands

<sup>3</sup> TiasNimbas Business School, The Netherlands

## ABSTRACT

This paper discusses whether and how governments can deal with or use Web-based Internet Communities (WebICs) as a source of information towards their citizens. WebICs are becoming more and more prominent in producing information in areas such as health, software or encyclopedic knowledge and therefore potentially rival or complement official information produced by governments. The topic of this paper is how these new sources of information, which are available to all citizens having access to internet, constitute constraints and opportunities for governments' information policy. We reflect on the status of this type of information from a public policy perspective. Given the outcome that information produced by WebICs is in principal relevant and appropriate from a public policy perspective the next question is how public administration can assess the quality of this type of information. Given the fact that public administration is not the producer of this information but that citizens are nevertheless the consumers of this information, this becomes a pertinent question.

We consequently introduce a method for assessing quality of WebICs by accessing exclusively information on how WebICs are produced that is available on the Internet. The suggested method is also available to public administration and does in no way disturb, control or regulate the work going on in WebICs. The method presented is based on insights from the employee turnover literature and argues that one potential indicator of WebIC quality is the dropout of active members, and subsequently proposes that governance mechanisms contribute in predicting the dropout hazard of WebIC members. We illustrate and test the method on the basis of 200 randomly selected pages in Wikipedia, a WebIC that produces encyclopedic information. The results show that reputation and controversy have different effects for different types of Wikipedians; i.e., an actor's reputation decreases the dropout hazard for *active* Wikipedians, while participation on controversial pages decreases the dropout hazard only for *highly active* Wikipedians.

*Key words: WebICs as a source for public information, online communities, governance, dropouts, network-analysis, reputation, controversies*

## INTRODUCTION

Besides power, law and finances, information is the most recent means of governance to influence the behavior of citizens in order to increase societal welfare (Willke 1995). For most of the 20<sup>th</sup> century public information was used as a tool of government (Weiss 2002) in an entirely top down way, i.e. government agencies producing the information themselves, funding research or hiring experts to produce public information which was then distributed through channels such as radio and TV spots, newspaper advertisements, reports and leaflets to the citizens. This information concerned areas like how to prepare for natural or man made disasters, warn citizens of poisonous or unhealthy substances, raise awareness of fraudulent behavior or distribute information on how to protect themselves against crime, etc. Production and distribution was sometimes coordinated with organized interests as for example victim associations or environmental groups. These top down information campaigns come with considerable cost. Weiss (2002:221) reports, that the US government spent 5-7,7 billion US \$ per year from 1977 to 1999 for information campaigns.

This situation, however, in which governments were very much in control of public information or at least had clear counterparts in organized interest groups that represented a more or less delineated group of citizens has come to an end with the development of information technology, especially the web 2.0, in the last ten years. Governments are now confronted with a diffuse public that creates and disseminates itself information in myriads of blogs, WebICs and other forums on the internet with a speed that is unknown and virtually unmatched by comparatively cumbersome public information campaigns. In addition, top down government information campaigns are increasingly attributed less legitimacy compared to the presumably bottom up process of information creation in WebICs that is perceived as more democratic and less influenced by lobby groups by a large part of the population.

In this paper, we will especially look at web-based information communities (WebICs) and their potential role in the provision of public information WebICs comprises work systems facilitated by the Internet infrastructure and composed of actors voluntarily attempting to create a product or service that is distributed by some type of open source license (Van Raaij et al., 2008). Although scholars have so far been particularly interested in WebICs that produce software, for instance the Linux kernel (e.g., Kuk, 2006; Lee and Cole, 2005; O'Mahony and Ferraro, 2007), WebICs often produce a wide array of other types of products, such as information services (e.g., Psychnet; Wikipedia), portals & management systems (e.g., Sahana) and tools used in plant genetic engineering that can contribute to for instance drugs development (e.g., Cambia). Consequently, WebICs often create societal value (Benkler, 2003; Watson et al., 2005) and are therefore by definition of interest to governments. WebICs can be considered as one, and probably increasingly important, source through which citizens inform themselves. We can readily assume that the source of information has an impact on how citizens behave as a result of receiving that information.

Governments are thus confronted with the question of whether and how they should react to this development and these new type of work system. As formal organizations developed in the 19<sup>th</sup> century and posed a new challenge for governments at that time, they now also are confronted with a new type of work system. At the time, governments started to regulate and control formal organizations or in case of non-profit organizations in the third sector started to subsidize, regulate and control them in order to make use of their capabilities in delivering welfare services to citizens and enhance the overall welfare of societies. The current situation, however, is complicated in as the people behind the WebICs are often not known or difficult to

identify and the work system is virtual and not bound to any geographical location which makes it much harder to stimulate, regulate, monitor, assess or control this type of work system compared to the formal organizations that developed in the 19<sup>th</sup> century.

There are different situations possible, governments might find themselves in with regard to WebICs. First, they might want to produce information in a certain area available to citizens but lack the knowledge or resources to achieve that. A WebIC of dispersed experts or concerned citizens might be able to produce and disseminate this information. Noveck (2009) for example describes a Peer-to-Patent WebIC that connects patent examiners to volunteer scientists and technologists via the web. Second, a situation might occur in which a WebIC has produced information that is creating a counter public to the official information distributed by governments. This could for example be in the area of vaccination, where public health officials would like to propagate a certain vaccination, while there is still some doubt about the effectiveness or side effects that are discussed in this WebIC. Third, a WebIC might produce outright wrong and damaging information, for example that HIV/Aids could be cured with Vitamin C, which governments would like to prevent from being disseminated. Last, government agencies might encounter cases where WebICs produce very useful information, whose quality, however, is unknown and difficult to assess. Depending on their quality, WebICs can therefore be use or harmful in several ways. In this paper we further explore what governments could do in order to assess and secure the quality of information that is potentially beneficial for society, if they decide to use WebICs as a new tool of government (Salamon 2002) in creating and distributing information among citizens. In addition, what then explains the differences in quality? To be clear, we are not advocating that governments must deal with WebICs or even have to use or control them. Our interest is merely, given this new type of work system and potentially high impact on public information, how could governments approach them if they wish to do so.

The first research field we have chosen for identifying and explaining the quality of WebICs is the English version of Wikipedia ([http://en.wikipedia.org/wiki/Main\\_Page](http://en.wikipedia.org/wiki/Main_Page)). In particular, we study the relational structure of the *edit-networks* associated to *individual* Wikipedia entries (i.e., a single page referring to a specific encyclopedic topic, e.g., *gun politics* [[http://en.wikipedia.org/wiki/Gun\\_politics](http://en.wikipedia.org/wiki/Gun_politics)]). The edit-network associated with a Wikipedia page has as actors the authors of the page and encodes how actors edit the page and how they respond to the edits of others. Among other things, this information can be used to determine the role of authors (e.g., to discriminate between those who provide content and those who delete content), to determine authors that erase or defend each other's content most actively, or to determine whether the community decomposes in different groups of opinion. To date, organizational scholars have hardly studied the organizing processes that take place at page level. In contrast, more formal governance that are executed by *Wikipedia administrators* (e.g., formal rules, punishments and the development of a two-tier governance structure distinguishing 'stable' and 'live' versions of entries) have been studied in order to better understand how WebICs differ from other types of governance structures (Garud et al., 2008; Giles, 2005) such as inter organizational networks or hierarchies (Demil and LeCocq, 2006).

### **Assessing the quality of WebICs: Dropout as a WebIC quality indicator**

The quality of Wikipedia is mostly associated to direct assessment of encyclopedic content by scientific experts (e.g., Giles, 2005), or by Wikipedia members themselves (e.g., Stein and Hess, 2007; Brandes et al., 2009a). Although assessment by experts can be a valid approach for assessing the quality of encyclopedic content, it is extremely time-consuming, not the least

because there are 2.6 million entries in English (growing every day). Furthermore, it requires qualitative analysis and does not fit our goal to explore whether online, quantitative data are useful. Others have operationalized Wikipedia’s quality in terms of criteria developed by Wikipedians; e.g., excellent, featured, word reading and controversial pages (Stein and Hess, 2007; Brandes et al., 2009a). Although they operationalize WebIC quality in terms of quantitative data, validity issues challenge them. For instance, controversy not necessarily refers to the quality of encyclopedic information but to conflict, which does not always harm quality (Lee and Cole, 2005).

Here, we propose an alternative approach, which allows us to collect quantitative data about Wikipedia outcomes: the dropout hazard of Wikipedians. This approach is based on the observation that many ‘eyeballs’ are needed for assessing and improving products produced by WebICs (e.g., Garud et al., 2008; Raymond 1999). Moreover, employee turnover is considered an issue affecting every type of organization (Beadles II et al., 2000), having an (negative) impact on organizational performance (Mitchell et al., 2001; Dess and Shaw, 2001; Morrell et al., 2001), particularly if key people with scarce skills and critical knowledge dropout (Mitchell and Lee, 2001; Dess and Shaw, 2001). Yet, whereas the original organizational turnover research started from the assumption that turnover is negative (Beadles II et al., 2000: 331), others have questioned this (e.g., Dalton, Todor and Krackhardt, 1982; Ableson and Baysinger, 1984; Beadles II et al., 2000; Siebert and Zubanov, 2009), and suggested to distinguish between functional and dysfunctional turnover (Dalton, Todor and Krackhardt, 1982). This idea was tested by several others, who conclude that: “The basic premise of turnover functionality is the explicit recognition that different employees are of different value to the organization” (Beadles II et al., 2000: 332). It seems appropriate also in the case of WebICs to distinguish between functional and dysfunctional turnover (Morrell et al, 2001). Consequently, we assume that when highly valued Wikipedians quits an entry (i.e., a single Wikipedia page) the quality of this entry will decrease, while in case hardly valued Wikipedians leave the quality will increase (see table 1).

**Table 1: Dropout Chance as Proxy for Quality of a Wikipedia Entry**

		<b>Degree of chance in dropout of Wikipedians</b>	
		<b>High</b>	<b>Low</b>
<b>Value of the Wikipedians</b>	<b>High</b>	Decrease in Wikipedia entry quality	Increase in Wikipedia entry quality
	<b>Low</b>	Increase in Wikipedia entry quality	Decrease in Wikipedia entry quality

The challenge is to be able to distinguish valuable from low valuable actors, which by definition is a subjective judgment (Dalton, Todor and Krackhardt, 1982). For hierarchical organizations, so-called ‘hard’ measures (e.g., productivity, sales) and ‘soft’ measures (e.g., supervisory appraisals, self-perceptions) have been applied. However, these type of data are unavailable for WebICs, and mostly are not appropriate since WebIC governance differs from the governance of hierarchies (O’Mahony and Bechky, 2008; Ganley and Lampe, 2009). Consequently, we propose to distinguish valuable from less valuable WebIC members by looking at their activity levels; i.e., (highly) active Wikipedians are considered as more valuable compared to less active

Wikipedians, because they are more involved in providing and assessing encyclopedic content. Therefore, we only study (highly) active Wikipedians.

### **Predicting the quality of WebICs: Why WebIC members dropout**

In this section, we will consider whether explanations for employee turnover and research in the field of computer science is helpful for explaining the dropout hazard of WebIC members.

The organizational equilibrium perspective of March and Simon (1958) serves as the major theoretical underpinning for much of today's research on employee turnover (Mitchell and Lee, 2001; Donnelly and Quirin, 2006). "March and Simon identified the turnover decision process as primarily attitude-driven and a function of the perceived ease and desirability of changing jobs" (Donnelly and Quirin, 2006: 59, see also Mitchell et al., 2001, Mitchell and Lee, 2001). Consequently, most research on employee turnover considers perceptual and/or affective factors; i.e., scholars from the psychological school have tried to explain employee turnover by affective constructs (satisfaction, commitment) while others have taken an economic perspective (e.g., perceived job alternatives). Although these constructs show a rather consistent correlation with employee turnover (see e.g., Hom and Griffeth, 1995; Maertz and Campion, 1998), they are generally weak (Donnelly and Quirin, 2006; Mobley et al., 1979). Therefore, several scholars have introduced alternative constructs, some with a different nature, in order to better explain turnover. Mitchell and Lee (2001), for instance, argue that the traditional sequence of "negative attitudes -> search -> quit" is over-simplified and developed a more complex model; i.e., the 'unfolding model' comprising of different decision paths dependent on 1) shocks (e.g., organizational merger, relocation of spouse's job) that make employees reassess their job and 2) decision frames (i.e., "the social and cognitive context that surrounds the experienced shock and which provides a frame of reference for interpreting the event" [2001: 199]). The results of their empirical work suggest that there are multiple ways to leave and that unavoidable and uncontrollable events have a greater impact on decisions to leave than dissatisfaction. Alternatively, by taking a network analytical perspective, Krackhardt and Porter (1986) studied the correlation between employees' perceived informal role similarity and turnover. Krackhardt and Porter argue that whereas traditional research on employee turnover considers each person's behavior to be a stochastic function determined by characteristics attributed to the person, they assume turnover not to be independently distributed. Depending on the nature of the relationship, people affect each other in their behavior (Krackhardt and Porter, 1986). It are *perceptions of similarity* that can affect each other's behavior, not the *actual* network position. Starting from the other side of the coin, i.e., why employees' stay, Mitchell and colleagues introduced the construct job embeddedness (2001). Results suggested that non-attitudinal factors such as how well one *fits* and is *linked* with ones' job and the external community, and which *sacrifices* are perceived when leaving a job, determine staying or leaving. Moreover, it shows that on-the-job external activities (e.g., relationships with colleagues) and off-the-job external activities (e.g., commitment to community, work-life balance) influence people to stay.

This brief review shows that turnover decisions are determined by a complex set of variables. Both perceptual, affective, relational and external factors influence employee turnover. Moreover, it shows that many factors explaining employee turnover are not relevant for WebICs,

due to their virtual nature and alternative governance structure<sup>1</sup>. Due to our focus on data that are available on-line, only relational explanations (e.g., relationships with colleagues) for employee turnover could be an appropriate starting point for predicting the dropout hazard of WebIC members based on employee turnover literature.

In the organizational literature it is suggested that structural and relational characteristics occurring in a work system's context are indicators of its governance structure influencing organizational actors (see Lewin in: Krackhardt and Porter, 1985). Yet, Krackhardt and Porter argued that this contextual factor is pushed into the background. Therefore, they have re-emphasized the impact of relational factors (e.g., colleagues leaving) on turnover (e.g., Krackhardt and Porter, 1985; 1986), and currently, the effect of networks of people leaving one company for another, is taken into account in popular and scientific literature (e.g., Dess and Shaw, 2001). This illustrates that people affect each other, depending on the nature of their relationship. Also computer scientists report on this relationship. Lento et al., (2006) showed that interacting with active actors is a strong and significant predictor of continued user activity in Weblogs. Also earlier work on the causes of dropout from Wikipedia suggests the relevance of relational explanations; i.e., the process of getting feedback on user talk pages<sup>2</sup> and replying on feedback as a Wikipedian predicts dropout hazards. Getting feedback increased the hazard to dropout, while replying to feedback attenuated this effect (Brandes et al., 2009b).

These relational predictors of turnover are closely related to the more informal, and self-regulating nature of WebIC governance, which is characterized by discretionary feedback, peer review and community norms (Moon and Sproull, 2008), and substitutes economic means to influence employees levels of commitment or satisfaction that are appropriate in hierarchical organizations (Moon and Sproull, 2008; Ganley and Lampe, 2009). Based on these results from the employee turnover literature and empirical studies in computer science, we will explore if the informal governance structure of WebICs predicts dropout hazards.

### **Relational characteristics of a WebICs informal governance structure as predictor for dropout hazards**

Various properties of WebIC governance have been emphasized; some referring to formal features of WebIC governance others to informal ones. Markus defines WebIC governance rather generally as: “the means of achieving the direction, control, and coordination of wholly or partially autonomous individuals and organizations on behalf of an OSS project to which they jointly contribute” (Markus, 2007: 152). Her summary of appropriate governance mechanisms illustrates their informal nature: e.g., peer reviewing as a monitoring mechanism and reputation signaling practices replacing economic incentives. Other mechanisms refer to its more formal features, although not based on contrasts: e.g., structures of roles and responsibilities and formal rules. In this paper we can only analyze if some features of the informal governance structure are significant predictors of Wikipedians dropping out; i.e., actor reputation, actor reputation increase and participation in controversial pages. We have selected these features of WebIC governance

---

<sup>1</sup> For instance, the fact that WebIC participants are volunteers without formal labor contracts and that WebICs are open systems in which open source licenses are used (e.g., GNU General Public License) (Lerner and Tirole, 2001; von Krogh and von Hippel, 2003; Schweick, Evans and Grove, 2005).

<sup>2</sup> User talk pages are discussion pages related to an editor (see [http://en.wikipedia.org/wiki/Talk\\_pages](http://en.wikipedia.org/wiki/Talk_pages)).

since they are typically reflected in network structures emerging in the context of WebICs, and data are available on-line.

### ***Actor reputation and actor reputation growth***

According to Campbell (1960 in: Lee and Cole, 2003) reputation mechanisms contribute to preserving and reproducing selected variations by keeping actors motivated to continue to contribute. This is in line with former research on participation to WebICs. Also in this body of research reputation is referred to as a motivating factor. Thus, reputation mechanisms help to better understand the governance of WebICs by paying attention to individual level motivations for participation (Kuk, 2006, see also Lakhani and von Hippel, 2003; von Hippel, 2005; Hertel, Niedner, and Herrmann, 2003; Shah, 2006; Ganley and Lampe, 2009) and by contributing to the management of intellectual capital. Lee and Cole argue that a node's reputation is based on 1) the appropriateness of the solutions added by a node and 2) the quality of critical evaluations provided by a node. Emphasized here is the quality of contributions. However, it is generally argued in the literature that reputation is a relational construct because reputation and reputation growth depend on the behavior and attitudes of others (e.g., Ganley and Lampe, 2009). Based on earlier results that demonstrate that WebIC members start to participate in order to increase their reputation, we propose that as a node's reputation on the one hand, and the growth of a node's reputation (i.e., developing in a positive direction from low to higher or vice versa in a negative direction) on the other hand decreases the chance the node drops out. Thus:

*H1 If the reputation of a WebIC member is high then the chance this WebIC member drops out is smaller than if the reputation of a WebIC member is low.*

*H2 If the reputation growth of a WebIC member is positive, then the chance this WebIC member drops out is smaller than if the reputation of a WebIC member is negative.*

### ***Controversies***

Controversies within (online) work groups can have a positive or a negative effect on their quality. For one thing, literature on the development of knowledge suggests that creative solutions are built from the recombination of existing ideas (Hargadon and Bechky, 2006), and that variation is a necessary condition for the quality of knowledge creation (Campbell in: Lee and Cole, 2003). This also implies that, at least to some degree, controversies are required.

Moreover, challenging the ideas of others, also referred to as problem solving dialogues between diverse subgroups or experts who bring in exclusive knowledge (Alevi and Leidner, 2001) is suggested to be an important feature of WebICs that are able to produce high quality outcomes. Yet, the question remains whether controversy motivates or demotivates voluntary WebIC participants? Research on why people are motivated to contribute for free demonstrates that users participate because they enjoy working together (Shah, 2005; Roberts, Hann and Slaughter, 2007; Lerner and Tirole, 2005) i.e., the process of working together (Füller, et al., 2006), the fun experienced when working together (Hienerth, 2006; Füller, Jawecki and Muhlbacher, 2006; von Krogh and Seath, 2007; Franke and Shah, 2003, Shah, 2006), and the enjoyment of the work itself (Lakhani and von Hippel, 2003) stimulates people to volunteer. Additionally, Lettl, Herstatt, and Gemuenden (2006) showed that embeddedness in a supportive environment has a positive effect on user participation. An effect of low levels of controversies on user participation can also be deduced from results showing that rivalry between users demotivates WebIC members (Hienerth, 2006; Franke and Shah, 2003; von Hippel, 2002).

Therefore, we propose that if others disapprove a member's contributions, less community support is experienced and the joy and fun of the working process itself decreases. This is particularly the case for WebICs in which a lot of conflict is going on. Therefore, the following hypothesis is proposed:

*H3 If a WebIC member often participates in WebICs that are characterized by high levels of controversies then the chance this WebIC member drops out is higher than if a WebIC member seldom participates in WebICs that are characterized by high levels of controversies.*

### Methods of analysis

In order to better understand whether a WebIC's informal governance structure predicts Wikipedians to drop out a network-analytic and longitudinal perspective will be applied. We will study how the edit-network *evolves over time* in terms of actor reputation, actor reputation growth, and participation in controversial pages. To explore whether these patterns of edit-network evolution estimate the chance to dropout, lifetime-analysis is applied. Earlier (Brandes et al., 2009b), a statistical model for the dropout hazard (i.e., the conditional probability of users dropping out at time  $t$ , under the precondition that they survived up to  $t$ ) was developed. The dropout rate was modeled in the following functional form:

$$h_u(t) = h_u(W_t; \theta) = \exp \left( \sum_{i=1}^k \theta_i \cdot s_i(u; W_t) \right)$$

Here the dropout hazard of user  $u$  at time  $t$  is assumed to be explained by statistics  $s(u, W)$  that describe specific aspects of the Wikipedia history  $W$  at time  $t$  from the point of view of user  $u$ . (As an example, a specific statistic  $s(u, W)$  might encode the reputation of  $u$  at time  $t$ .) The estimated parameters  $\theta$  give information about whether a hypothetical cause of dropout (e.g., reputation) shows the predicted effect. If the parameter associated with the reputation statistic for example, is significantly negative, then users with higher reputation would have lower chance to drop out – pointing to the fact that users might be motivated by reputation. A significantly positive parameter would indicate that the associated statistic encodes a cause of increased dropout. Details of how the parameters can be estimated from an observed set of dropouts and survivors are given in Brandes et al., (2009b).

### *Sample strategy valuable and very valuable users*

Determining the “lifetimes” of Wikipedians is not straightforward – it is even not obvious how to decide whether a certain Wikipedian “died” or whether he will resume editing at a later time. We restricted the analysis to active users, i.e., users with a specified minimum number of edits, in order to concentrate on Wikipedia's most valuable members, and since it is unclear whether a user with only, say, two edits is a dropout or just rarely involved. To determine active users and divide them into dropouts and survivors we used the revision history of the English Wikipedia dating October 10, 2008, see <http://download.wikimedia.org/>. A user was called active if he did at least 1,000 edits before July 1, 2008, and a user was called highly active if he did at least 10,000 edits before July 1, 2008. Out of several million registered users (excluding Bots; i.e., software programs that perform routine tasks), 17,714 qualified as active and 2,463 as highly active. The time period from July 1, 2008 until October 8, 2008 was used to split (highly) active

users into dropouts and survivors. A user was called a survivor if he did at least one edit after July 1, otherwise he was classified as dropout. This criterion resulted in 13,126 active survivors, 4,588 active user dropouts, 2,130 highly active survivors, and 333 highly active users who dropped out. The requirement of 1,000 edits for active users was motivated by a criterion mentioned on the Missing Wikipedian page, [http://en.wikipedia.org/wiki/Wikipedia:Missing\\_Wikipedians](http://en.wikipedia.org/wiki/Wikipedia:Missing_Wikipedians).

### ***Edit-network structure***

Indicators for an actor's reputation, an actor's reputation increase and participation in controversial pages can be derived from the edit-history of a Wikipedia page; i.e., the edit-network. In Brandes et al., (2009a) it is demonstrated that the edit-network associated with a Wikipedia page  $p$  has as nodes the authors of  $p$  and encodes how authors contributed to  $p$  and how authors interacted with each other while editing  $p$ . This information is computed from the complete history of  $p$ ; i.e., from the sequence of its revisions, by determining which part of the text has been added, has been deleted, or remained unchanged when going from one version of the page to the next.

### ***Actor reputation***

A Wikipedian's reputation refers here to the appropriateness of his solution added perceived by a page's co-authors. This idea is somewhat similar with the work of Adler and Alfaro (2007), who for developing a reputation system for Wikipedia, considered whether text edited by a user persists on the page or gets deleted afterwards. We will measure the reputation of user  $u$  by centrality levels based on the ratio of persistent text over all text added. Since the number of words that are added by  $u$ , as well as the number of persistent words, can change from one edit to the next we get a time-varying measure of reputation for each time point  $t$

$$reputation(u,t)=[a(u,t)-d(u,t)+r(u,t)]/a(u,t)$$

$a(u,t)$  being the number of words added by user  $u$  before or at  $t$ ;

$d(u,t)$  being the number of words added by  $u$  and deleted afterwards (by any user) before or at  $t$ ;

$r(u,t)$  being the number of words added by  $u$ , deleted afterwards and restored later (by any user) before or at  $t$ ;

Consequently  $a(u,t)-d(u,t)+r(u,t)$  is the number of words added that are either not deleted or, if deleted, then restored and  $reputation(u,t)=[a(u,t)-d(u,t)+r(u,t)]/a(u,t)$  is the ratio at time  $t$  of those persistent words over the number of all added words being a centrality measure with a possible score between 0 and 1.

### ***Actor reputation increase***

It might be the case that not only the absolute reputation is a predictor for dropout, but also the increase/decrease of reputation. For instance, if two users  $u$  and  $v$  both have on a particular day<sup>3</sup> a

---

<sup>3</sup> It would not make a big difference for this particular analysis whether we operationalize reputation increase based on the increase during one day, a week (or a month) due to technical reasons: the reputation increase within one week or month is the sum of the daily increase within this week. So the weekly increase is high if and only if it is high for (some of) the days of that week. From another point of view, the weekly increase is the average of the daily increases - up to a constant factor that changes nothing.

reputation score of (say) 0.5 but u's reputation increased recently from 0.3 to 0.5 whereas v's reputation decreased from 0.7 to 0.5, then u might perceive his reputation score as satisfying/motivating while v considers the same value as a setback. To test whether such effects can be observed we define the reputation-increase of user u at time t with respect to a given time lag  $dt$  to be the difference

$$\text{reputation increase}(u,t) = \text{reputation}(u,t) - \text{reputation}(u,t-dt)$$

Reputation increase is thus normalized to the interval [-1.0,1.0]; it is positive if reputation increased and negative if reputation decreased.

### ***Participation in controversial pages***

First, we operationalized controversial pages in terms of the edit-networks levels of bipolarity. Bipolarity characterizes the global collaboration structure and estimates whether the author community of an individual Wikipedia entry decomposes into two groups of opinion that mutually undo the edits of each other. We define bipolarity as follows (see Brandes et al., 2009a, also for visualization on high and low bipolar edit-networks):

$$\text{Bipolarity} = \frac{w - c}{w + c}$$

The author community was partitioned into two groups, A and B, such that the aggregated weight of negative edges between the two groups is maximized. Denote this weight with  $w$  and let  $c$  denote the aggregated weight of negative edges connecting authors within A or within B. The bipolarity score lies between -1 and +1. It is equal to +1 if the graph (e.g., edit-network) is bipartite, i.e., edges connect only members from different groups and, therefore, the division into opposing groups is perfect. The bipolarity score equals 0 if all pairs of actors are connected, and all edges, including loops have the same weight--indicating no opposition at all. Negative weights of the bipolarity are theoretically possible but unlikely and have never been observed in our tests. In fact, it would indicate that all authors delete mostly their own edits but rarely the edits of any other author. Since the bipolarity (controversy) level of a page may change over time, we computed the bipolarity of a given page in separate time intervals of half a year. In the following, let  $\text{bipolarity}(p,t)$  denote the bipolarity of page  $p$  in the half-year interval containing time point  $t$ . Starting from bipolarity as an operationalization of controversy, the statistic  $\text{sum\_controversy}(u,t)$  is defined to be the sum over the bipolarity of pages edited by user  $u$  before or at time  $t$  (where the bipolarity of a page  $p$  is counted as often as  $u$  has edited  $p$ ).

### ***Number of Edits***

Since  $\text{sum\_bipolarity}(u,t)$  is, by design, monotonically increasing with the number of edits done by  $u$ , a positive (respectively negative) parameter associated with participation to controversial pages could just be due to the possibility that users wear out over time (respectively get more robust against dropout over time). Thus, we included the number of edits as a control variable when analyzing the effect of participation to controversial pages. To define the statistics encoding how much a particular user  $u$  contributed to the edit-network up to a time point  $t$ , let  $E_{u,t}$  denote the set of revisions that  $u$  performed on pages of the main namespace on or before time  $t$ . The respected statistic is defined by:

$$\text{Edit}(u; W, t) = \text{Log}(1 + |E_{u,t}|)$$

The logarithmic scaling of the number of revisions has been chosen due to the extremely skewed distribution (there are users who perform more than 100,000 revisions, while most of the selected users have a count of only slightly more than 1,000). A significant positive (negative) parameter associated with edits implies that users with higher numbers of revisions to the main namespace have a higher (or lower) hazard to drop out.

#### *Approximation for computational reasons*

Since it is computationally infeasible to compute the edit-networks of all (more than two million) Wikipedia pages, we approximate the reputation scores in the following way. First we randomly select 200 pages in an importance-driven sampling procedure<sup>4</sup>: a page is selected with probability proportional to the number of its edits that are done by active users. Therefore, pages on which active users are highly involved have a higher probability to be selected. The reputation scores that are computed on the selected pages serve as an approximation of the true reputation scores. Due to the large N (more than 17,000 users) we strongly believe that this approximation does not severely change our parameter estimates – although the approximate reputation of some individual users might be far of the true value.

## RESULTS

In the result section we report on three models that explored whether informal governance mechanisms predict the dropout hazard of (highly) active Wikipedians. The first hypothesis was partly confirmed, the second hypothesis was rejected, and for the third hypothesis we found counter-expected results.

#### ***Reputation***

The parameter associated with *reputation* is significantly negative for active users, and is non-significant for highly active users (see table 1).

---

<sup>4</sup> Contrary to a uniformly random sample which would give the same probability to everyone of 2 million pages. Because most pages are very small, containing very little revisions, a uniformly random sample would decrease our chance to learn something about, for instance, the reputation of (*highly*) active users.

**Table 1: Estimated reputation parameter, standard errors and *t*-ratios for dropout chance of (highly) active users<sup>5</sup>**

Statistic	Active users (at least 1,000 edits)		Highly active users (at least 10,000 edits)	
	Parameter (s.e)	<i>t</i> -ratio	Parameter (s.e.)	<i>t</i> -ratio
<b>Reputation</b>	<b>-0.1930****</b> (0.015)	12.87	-0.140 (0.147)	0.95

\*\*\*\*= parameter significantly different from zero at 0.1%.

This indicates that the dropout hazard of active users decreases with an actor's higher reputation scores, but that reputation is not a good predictor for the hazard of highly active users to dropout. Apparently, for the highly active users, reputation in terms of the persistence of words is not that important. This could be explained by the different roles that are fulfilled by active versus highly active Wikipedians. If a user does 10,000 edits, and sometimes up to more than 100,000 edits, than he cannot add much text as 10,000 paragraphs or even sentences is an incredible amount of work. Hence, *highly active* Wikipedians mostly do revision work: spell-checking, reverting vandalism, or checking formal requirements on text written by others, while in general *active* Wikipedians are content providers. For the latter, reputation in terms of persisted text is more motivating.

#### **Reputation increase**

If we add the parameter *reputation increase* to the model, the hazard of (highly) active users to drop out is not predicted (see table 2).

**Table 2: Estimated reputation parameter, reputation increase parameter, standard errors and *t*-ratios for dropout chance of (highly) active users**

Statistic	Active users (at least 1,000 edits)		Highly active users (at least 10,000 edits)	
	Parameter (s.e)	<i>t</i> -ratio	Parameter (s.e.)	<i>t</i> -ratio
<b>Reputation</b>	<b>- 0.195 (0.015) ****</b>	13	-0.154 (0.148)	1.04
<b>Reputation increase</b>	0.093 (0.166)	0.56	1.620 (0.976)	1.65

\*\*\*\*= parameter significantly different from zero at 0.1%.

Hence, reputation increase does not predict either the dropout hazard for active or for highly active Wikipedians. Moreover, if we add reputation increase to the reputation model, the predictive value of reputation only slightly improves for active users. Hence, our second hypothesis is rejected. It could be that for active Wikipedians a certain level of reputation is a

<sup>5</sup> The constant just normalizes the model to the empirical time scale in which one unit corresponds to the expected time-to-dropout of a (hypothetical) user for which the effects of all other statistics add up to zero. Since the value of the constant does not provide much information, we do not report on its values.

sufficient motivator; i.e., that increasing of already high reputation or vice versa does not motivate or demotivate. It could also be the case the reputation increase remains unnoticed.

***Participation in controversial pages***

The data demonstrate that if highly active users participate in controversial pages, the dropout hazard of highly active users decreases. Hence, participation in controversial pages seems to motivate highly active Wikipedians, but by contrast, has no effect on the hazard that active users will leave (see table 3).

**Table 3: Estimated participation on controversial pages parameter, edit parameter, standard errors and *t*-ratios for dropout chance of (highly) active**

Statistic	Active users (at least 1,000 edits)		Highly active users (at least 10,000 edits)	
	Parameter (s.e.)	<i>t</i> -ratio	Parameter (s.e.)	<i>t</i> -ratio
Participation on controversial page	-0.001 (0.001)	1.00	<b>-0.0058 ****</b> (0.0015)	3.87
Number of edits	0.00048 (0.001)	0.478	<b>0.0048 ****</b> (0.0012)	4.00

\*\*\*\*= parameter significantly different from zero at 0.1%.

Hence, hypothesis 3 is rejected. Instead, participation in controversial pages has a positive effect. We checked whether our results could be influenced by our operationalization, by developing three alternative statistics; 1) counting number of edits to pages with bipolarity bigger than 0.7, 2) taking the average bipolarity of pages edited by the user, and 3) taking the average bipolarity of pages edited today by the user. None of these led to significant results.

Moreover, we included the number of edits in the model to check the robustness of the effect of participation on controversial pages on the dropout chance of (highly) active users. Earlier (Brandes et al., 2009b), it was reported that the number of edits has a negative significant effect on the dropout hazard of active Wikipedians. Without taking the number of edits into account, participation in controversial pages had a significantly negative effect (i.e., more bipolarity - less dropout). The sum\_controversy statistic (i.e., the statistic referring to the sum over the bipolarity of pages edited by user *u*) has the property that it is monotonically increasing with the number of edits (more edits, higher value of this statistic). So it could be the case that users become just less likely to drop out when they accumulate edits. Indeed, in a model with just the number of edits and a constant, the number of edits is significantly negative as well. The results presented here demonstrate that the effect of participation in controversial pages is robust for highly active users when both statistics are included; but not for active users. This could be explained by the nature of the product produced; encyclopedic information. Our hypotheses were based on results about WebICs producing software and sport equipments. It could be that these developers are indeed demotivated by rejections of their contributions, while highly active Wikipedians reverting what they see as ‘vandalism’ are motivated by rejections; that is to say, are motivated to reject the contribution of the other again.

**DISCUSSION**

The results suggest that different features of informal WebIC governance have varying impact on different types of Wikipedians, which potentially differ in their value for Wikipedia entries. This

suggests that these online data sources can be helpful for explaining and identifying WebIC quality, although better operationalizations are needed for WebIC quality in the future. Yet, the preliminary results suggest that different governance mechanisms can be effective under various conditions. While currently organization scholars have mostly described the general features of WebIC governance (Bechky and O'Mahony, 2009; Demil and Lecocq, 2006), the results suggest that not each governance mechanism is equally effective, and that the effectiveness of WebIC governance depends on the type of WebIC. While studies on software WebICs report on the effectiveness of the absence of controversies for sustaining volunteers, this does not seem to work in encyclopedic WebICs. By contrast, controversy motivates highly active Wikipedians. Also, reputation is generally understood as an effective feature of WebIC governance by motivating users to participate (e.g., Markus, 2007; Lakhani and von Hippel, 2003; von Hippel, 2002; Hertel Niedner and Herrmann, 2003; Shah, 2000, 2006; Hars and Ou, 2002;). Only a couple of studies argue that reputation is not a significant predictor of participation (Jeppersen and Frederiksen, 2006; Franke and Shah, 2003). Our research provides nuances to these conclusions by suggesting that reputation is not always effective as a mean of achieving direction, control, and coordination of volunteering individuals.

These early results could also be relevant for studies on hierarchical organizations and employment turnover. A tendency is noted that hierarchies move towards less formal, hierarchical governance structures (e.g., Sinh and Van de Ven, 2005), for instance when they open up R&D departments, engage in open innovation processes and crowd surfing (Chesbrough, 2006), and start co-producing with WebICs. Under these circumstances, in which contracts and economic incentives are less appropriate (O'Mahony and Bechky, 2008), organizations need alternatives to governance mechanisms appropriate for hierarchies. Better understanding the effects of WebIC governance, among others, reputation and how to deal with controversies, could be helpful.

Moreover, it has been noticed that employee turnover research is challenged by the difficulty of explaining turnover based on affective and perceptual constructs and it has been shown that taking a relational perspective adds to a better understanding of this phenomenon (e.g., Krackhardt and Porter, 1985; 1986). Here, we have also started from such relational perspective, although by concentrating on relational characteristics that identify WebIC governance features. In line with results showing the importance relationships with colleagues (Mitchell and Lee, 2001), it could be beneficial to study the predictive power of other relational features for employee turnover, such as reputation.

## **IMPLICATIONS FOR PUBLIC POLICY**

Although the object of analysis in this study, the production WebICs can be considered to be a completely private endeavor (citizens producing information for citizens without in the majority of cases, any use of tax money) they nevertheless turn out to have considerable public policy implications.

First, WebICs are prevalent and we expect them to be more and more prevalent in the future. Secondly, WebICs touch on issues, governments traditionally have a role to inform citizens about (e.g. social and health risks). Third, WebICs have turned out to inform citizens where governments have failed to do so (e.g. catastrophe alerting systems) or are competing and providing different perspectives on an issues than the one to be found in WebICs (e.g. the information provided about the risks and advantages about an electronic patient registry in the Netherlands). Fourth, we easily can assume that the content of WebICs will vary in quality. This

does, however, not mean that they are not valuable by definition. WebICs often contain valuable ‘bottom up knowledge’ (e.g. tips on how to increase adherence to medication or other therapies) which can certainly be worthwhile to be promoted by public administration.

In order to do so, public administration might want to know two things: first, is the type and content of information provided in line with formal policy making principles on the issue at hand? Secondly, is the information provided of good quality? Crossing the element of value (i.e. fitting the normative choices made by government) with the element of fact (whether the information provided is of good quality or in other words based on evidence) produces the following four-fold table.

**Table 4: Possible responses by public administration to different categories of WebICs**

		Element of value (assessment in terms of normative alignment with policy)	
		Acceptable	Non-acceptable
Element of facts (assessment in terms of quality or evidence base)	High quality	<b>Promote</b>	<b>Tolerate (freedom of information provision)</b>
	Low quality	<b>Facilitate for better information</b>	<b>Activate counter evidence</b>

As can be seen from the table above there are a number of potential actions necessary for government in order to deal with WebICs in case the overall objective is to provide good information to citizens about the causes government has identified as important. Whereas it belongs the logic of governing itself to know which causes should be advanced (the element of value), the question of how to assess WebICs on their quality in terms of evidence base needs to be assessed in an objective way. We assume that in most cases public administration has not itself the content knowledge to assess the quality of the information in a WebIC. Therefore, we have presented here an alternative route to externally assess the quality of WebICs.

In a third step, what then are the main differences between a traditional top down information campaign and the inclusion of WebICs with regard to the process? Weiss (2002) identified five ideal typical steps in delivering information to the public (in the old paradigm): First, policymaker have to define the target audience. Who is it exactly they want to influence. Second, they have to decide which information will be useful, i.e. which information is appropriate and convincing to a particular audience, in order to influence their behavior in the intended way. Third, the information needs to be obtained. If the information does not yet exist, the policy must include provisions for the information to be gathered or produced. Fourth, the information has to reach the target audience, i.e. it has to be delivered in a way that insures maximum attention and transfer either directly or via intermediaries. Fifth, ideally the information delivered does not only reach the intended individual but also mobilizes his or her

respective social and institutional context. Since behavior is embedded in this context, a change in behavior very often requires a change in this context as well.

If we look at the information produced by WebICs, where governments intend to use existing information if they could assess the quality of the information (once the content has been valued as “acceptable”), we can see an important difference in the sequence of these steps if compare the traditional top down approach with the inclusion of WebICs. While in the traditional information campaign the audience needs to be defined and the information very often first be produced, steps 1 and 3 have already been taken beforehand. Governments now have to assess the quality of the information and decide whether they would like to broaden the target population on top of the people who already correspond the WebIC in question for example by including a text on the government website that is set up in a way so it can easily be found by the search engines and include a link to that WebIC, i.e. promote the information source or facilitate it to improve the information.

A traditional information campaign or information production and dissemination by a WebIC, information, however, is just one of many tools of government (Salamon 2002) that might be applied and effective only under specific circumstances (Weiss 2002:233): First, information policies might be useful if information asymmetry exists for example between producers and consumers or information costs for citizens are very high. Second, they might be successfully applied if the target population is very broadly dispersed and not organized. i.e. existing societal organizations cannot be used for the dissemination of the information. Third, information policies are likely to be successful, if the interests of policymakers and citizens are similar or when there is broad agreement of desired outcomes like reducing infant mortality and voluntary compliance by citizens is therefore likely to occur. Fourth, information policies might be successfully applied in instances, in which no alternative means like law or finances are available or acceptable. This might be the case in private areas such as sexual behavior where intervention with “hard” government tools would violate fundamental rights of citizens but a change of behavior nonetheless is in the interest of society. This was and is for example the background for the HIV/Aids campaigns.

Gross modo we therefore do not see a fundamental difference with regard to these circumstances in which information policies might successfully be applied in traditional campaigns and the dissemination of information via WebICs. However, as stated before this strategy stays and falls with the ability of governments to assess the quality of the current and possibly future output of a WebIC. We therefore developed a method that could be used in assessing a WebIC by using information that is relatively easily accessible on the Internet about the structure and process of information production. It is based that this information is indicative for the quality of what they produce. The method presented is a first attempt to assess these new forms of information production. There clearly is room for improvement. All further attempts, however, should stick to the principal that the assessment itself should in no way disturb, control or regulate the work going on in WebICs.

## REFERENCES

- Abelson, M. and Baysinger, B (1984) Optimal and dysfunctional turnover: toward an organizational level model. *Academy of Management Review*, 9, 331-342.
- Adler, B. T. and de Alfaro, L (2007) A content-driven reputation system for the Wikipedia. In *Proceedings 16<sup>th</sup> Intl. Conf. WWW2007*: 261-270.
- Beadles II, C. M., Lowery, N. A., Petty, M. M., and Ezel, H. (2000) An examination of the relationship between turnover, functionality, turnover frequency and organizational performance. *Journal of Business and Psychology*, 15, 331-337.
- Benkler, Y. (2003) Coase's Penguin, or, Linux and the nature of the firm. *The Yale Law Journal*, 112, 369-446.
- Brandes, U. and Erlebach, T. (2005) *Network Analysis: Methodological Foundations*. Heidelberg, Springer.
- Brandes, U., Fleischer, D. and Lerner, J. (2006) Summarizing dynamic bipolar conflict structures. *IEEE Transactions on Visualization and Computer Graphics*, 12, 1486-1499.
- Brandes, U., Kenis, P., Lerner, J. and van Raaij, D. (2009a) Network analysis of collaboration structure in Wikipedia. In *Proceedings 18th Intl. World Wide Web Conference (WWW2009)*.
- Brandes, U., Kenis, P., Lerner, J. and van Raaij, D. (2009b) Is editing more rewarding than discussion? A statistical framework to estimate the causes of dropout from Wikipedia. In *Proceedings 1<sup>st</sup> Intl. Workshop Motivation and Incentives on the web (Webcentives '09) at the 18<sup>th</sup> Intl. World Wide Web Conference (WWW2009)*.
- Brown, J.S. and Duguid, P. (2001) Knowledge and organizations. A social practice perspective. *Organization Science* 12, 198-213.
- Chesbrough, H. (2006) *Open Innovation: The New Perspective for Creating and Profiting from Technology*, Harvard Business School Press, Cambridge.
- Coleman, J. S. (1982) *The Asymmetric Society*. Syracuse University Press, Syracuse.
- Crowston, K. and Annabi, H. (2005) Effective work practices for FLOSS development: A model and propositions. International Conference on System Sciences, Hawaii.
- Crowston, K. and Howison., J. (2005) The social structure of free and open source software development. *First Monday*, 10  
<http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/1207/1127>
- Cummings, J. and Cross, R. (2003) Structural properties of work groups and their consequences for performance. *Social Networks*. 25, 197-210.

- Dalton, D. R., Todor, W. D. and Krackhardt, D. M. (1982) Turnover overstated: the functional taxonomy. *Academy of Management Review*, 7, 117-123.
- Demaziere, D., Horn, F., and Zune, M. (2007) The functioning of the free software community: entanglement of three regulation modes - control, autonomous and distributed. *Science Studies*, 20, 34-54.
- Demil, B. and Lecocq, X. (2006) Neither market nor hierarchy nor Network: The Emergence of Bazaar Governance. *Organization Studies*, 27, 1447-1466.
- Dess, G. G. and Shaw, J. D. (2001) Voluntary turnover, social capital, and organizational performance. *Academy of Management Review*, 26, 496-524.
- Donnelly, D. P. and Quirin, J. J. (2006) An extension of Lee and Mitchell's unfolding model of voluntary turnover. *Journal of Organizational Behavior*, 27, 59-77.
- Franke, N. and Shah, S. (2003) How communities support innovative activities: an exploration of assistance and sharing among end-users. *Research Policy*, 32, 157-178.
- Füller, J., Bartl, M., Ernst, H. and Mühlbacher, H. (2006a) Community based innovation: how to integrate members of virtual communities into new product development? *Electronic Commerce Research*, 6, 57-73.
- Füller, J., Jawecki, G. and Mühlbacher, H. (2006b) Innovation creation by online basketball communities. *Journal of Business Research*, 60, 60-71.
- Ganley, D. and Lampe, C. (2009) The ties that bind: social network principles in online communities. *Decision Support Systems*, 47, 266-274.
- Garud, R., Jain, S., and Tuertscher, P. (2008) "Incomplete by design and designing for incompleteness. *Organization Studies*, 29, 351-371.
- Giles, J. (2005) Internet encyclopedias go head to head. *Nature*, 438, 900-901.
- Hars, A. and Ou, S. (2002) Working for free. Motivations for participating in OSS projects. *International Journal of electronic Commerce*, 6, 25-39.
- Hargadon, A. B. and Bechky, B. A. (2006) When collections of creatives become creative collectives. A field study of problem solving at work, *Organization Science*, 17, 484-500.
- Hertel, G., Niedner, S. and Herrmann, S. (2003) Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel." *Research Policy*, 32, 1159-1177.
- Hienert, C (2006) The commercialization of user innovations: the development of the Rodeo kayak industry. *R&D Management*, 36, 273-293.

- Hom, P. W. and Griffeth, R. W. (1995) *Employee turnover*. South-Western College Publishing, Cincinnati.
- Ibarra, H. (1992) Structural alignments, individual strategies, and managerial action: elements towards a network theory of getting things done, in. Nohria, N. and Eccles, R. G. (eds), *Networks and Organizations*. Boston, Harvard Business School Press, 165-188.
- Jeppesen, L. B. and Molin, M. J. (2006) Why do users contribute to firm-hosted user communities? The case of the computer-controlled Music instruments. *Organization Science*, 17, 45-63.
- Kidane, Y. H. and Gloor, P. A. (2007) Correlating temporal communication patterns of the eclipse open source community with performance and creativity. *Computational and Mathematical Organization Theory*, 13, 17-27.
- Krackhardt, D. and Porter, L.W. (1985) When friends leave: a structural analysis of the relationship between turnover and stayer's attitudes. *Administrative Science Quarterly*, 30, 242:261.
- Krackhardt, D. and Porter, L.W. (1986) The snowball-effect: turnover embedded in communication networks. *Journal of Applied Psychology*, 71, 50-55.
- Kuk, G. (2006) "Strategic interaction and knowledge sharing in the KDE developer mailing list" *Management Science*, 52, 1031-1042.
- Lakhani, K. R. and von Hippel, E. (2003) How OSS works: "free" user-to-user assistance. *Research Policy*, 32, 923-943.
- Lee, G.K. and Cole, R. (2003) From a firm-based to a community-based model of knowledge creation: the case of the Linux kernel development. *Organization Science*, 14, 633 – 649.
- Leenders, R. Th. A. J., van Engelen, J. M. L. and Kratzer, J. (2003) Virtuality, communication and new product team creativity: a social network perspective. *Journal of Engineering and Technology Management*, 20, 69-92.
- Lento, T., Welsch, H. T., Gu, L. and Smith, M. (2006) The ties that blog: examining the relationship between social ties and continued participation in the Wallop weblogging system. In *Proc. 3rd Annual Workshop on the Weblogging Ecosystem*.
- Lerner, J., and Tirole, J. (2001) The open source movement: key research questions. *European Economic Review*, 45, 819-826.
- Lerner, J. and Tirole, J. (2005) The economics of technology sharing: Open source and beyond. *The Journal of Economic Perspectives*, 19, 99-120.

Lettl, C., Herstatt, C. and Gemuenden, H. G. (2006) User's contributions to radical innovation: evidence from four cases in the field of medical equipment technology. *R&D Management*, 36, 251-272.

Maertz, C. P. and Campion M. A. (1998) 25 years of voluntary turnover research: a review and critique. *International Review of Industrial and Organizational Psychology*, 13, 49-81.

Markus, M. L. (2007) The Governance of free/open source software projects: monolithic, multidimensional, or configurational? *Journal of Management and Governance*, 11, 151-163.

Marsden, P. V. (2005) The sociology of James S. Coleman. *Annual Review of Sociology*, 31, 1-24.

Mitchell, T. R., Holtom, B. C., Lee, T. W. and Erez, M. (2001) Why people stay: Using job embeddedness to predict voluntary turnover. *Academy of Management Journal*, 44, 6, 1102-1121.

Mitchell, T. R. and Lee, T. W. (2001) The unfolding model of voluntary turnover and job embeddedness: foundations for a comprehensive theory of attachment. *Research in organizational behavior*, 23, 189-246.

Mobley, W. H., Griffeth, R. W., Hand, H. H. and Meglino, B. M. (1979) Review and conceptual analysis of the employee turnover process. *Psychological Bulletin*, 86, 493-522.

Moon, J. Y. and Sproull, L. S. (2000) The essence of distributed work: The case of the Linux kernel. *First Monday*, 5, [Http://ftp.firstmonday.dk/www/issues/issue5\\_11/moon/](http://ftp.firstmonday.dk/www/issues/issue5_11/moon/).

Moon, J. Y. and Sproull, L. S. (2008) The role of feedback in managing the Internet-based volunteer work force. *Information Systems Research*, 19, 494-515.

Molm, L. D. and Cook, K. S. (1995) Social Exchange and Exchange Networks, In Cook, K. S., Fine, G. A. and House, J. S. (eds), *Sociological Perspectives on Social Psychology*, Boston, Allyn and Bacon, 209-235

Monge, P. R. and Contractor, N. S. (2003) *Theories of Communication Networks*. Oxford University Press, Oxford.

Morrell, K., Loan-Clarke, J. and Wilkinson, A. (2001) Unweaving leaving: the use of models in the management of employee turnover. *International Journal of Management Review*, 3, 219-244.

Noveck, B. S. (2009). *Wiki Government. How Technology Can Make Government Better, Democracy Stronger, and Citizens More Powerful*. Brookings Institution Press, Washington.

O'Mahony, S. and Bechky, B. (2008) Boundary Organizations: Enabling Collaboration among Unexpected Allies. *Administrative Science Quarterly*, 53, 422 – 459.

- O'Mahony, S. and Ferraro, F. (2007) The emergence of governance in an open source community. *Academy of Management Journal*, 50, 1079-1106.
- Powell, W. W. (1990) Neither market nor hierarchy. Network forms of organization. *Research in Organizational Behavior*, 12, 295-336.
- Raymond, E. S. (1999) *The Cathedral and the Bazaar: Musings on Linux and Open Source by Accidental Revolutionary Revised Edition*. O'Reilly, Cambridge.
- Roberts, J. A., Hann, I. H. and Slaughter, S. A. (2006) Understanding the motivations, participation, and performance of OSS developers: a longitudinal study of the Apache Projects. *Management Science*, 52, 984-999.
- Salamon, L.M. (Ed.) (2002). *The Tools of Government. A Guide to the New Governance*. Oxford University Press, Oxford.
- Schweick, C., Evans, T. and Grove, J. M. (2005) Open source and open content: a framework for global collaboration in social-ecological research. *Ecology and Society*, 10, 33 (<http://www.ecologyandsociety.org/vol10/iss1/art33/>).
- Siebert, W.S. and Zubanov, N. (2009) Searching for the optimal level of employee turnover: a study of a large U.K. retail organization. *Academy of Management Journal*, 52, 294-313.
- Sinha, K. K. and Van de Ven, A. H. (2005) Designing work within and between organizations. *Organization Science*, 16, 389 - 408.
- Shah, S. (2005) Open Beyond Software, In Dibona, C., Cooper, D and Stone, M. (eds), *Open Sources 2*. Sebastopol, O'Reilly, 339-360.
- Shah, S. (2006) Motivation, governance, and the viability of hybrid forms in open source software development. *Management Science*, 52, 1000-1014.
- Stein, K. and Hess, C. (2007) Does it matter who contributes? A study on featured articles in the German Wikipedia. In *Proc. 18th ACM Conf. Hypertext and Hypermedia (Hypertext 2007)*, 171-174.
- Takahashi, N. (2000) The emergence of generalized exchange. *American Journal of Sociology*, 105, 1105-1134.
- Tichy, N. and Fombrun, C. J. (1979) Network analysis in organizational settings." *Human Relations*, 32, 923-965.
- Van Raaij, D., Brandes, U., Kenis, P. and Lerner, J. (2008) *Determining the Quality of WebICs by Structural Parameters*. Paper prepared for EGOS conference, Amsterdam.

Van Wendel de Joode, R., Lin, Y. and David, S. (2006) Rethinking free, libre and open source software. *Knowledge, Technology and Policy*, 18, 5-16.

Viegas, F., Watternberg, M. and Dave, K. (2004) Studying cooperation and conflict between authors with history flow visualizations. In *Proceedings of CHI'04*, 575-582.

Von Hippel, E. (2002) Open source software projects as user innovation networks. *Open Source Software : Economics, Law and Policy*, Toulouse, 20-21.

Von Hippel, E. (2005) *Democratizing Innovation*. MIT Press, Boston,.

Von Krogh, G. and Speath, S. (2007) The OSS phenomenon: characteristics that promote research. *Journal of strategic information systems*, 16, 236-253.

Von Krogh, G. and von Hippel, E. (2003) Special issue on open source software development. *Research Policy*, 32, 1149-1157.

Watson, R. T., Boudreau, M.-C., Greiner, M., Wynn, D., York, P. and Gus., R (2005) Governance and global communities. *International Journal of Management*, 11, 125 - 142.

Weiss, J.A. (2002). Public Information, in Salamon, L.M. (Ed.) (2002). *The Tools of Government. A Guide to the New Governance*. Oxford University Press, Oxford. pp. 217-254.

Willke, H. (1994). *Systemtheorie III: Steuerungstheorie*. Gustav Fischer, Stuttgart.

Wellman, B. S. and Gulia, M. (1999) Net-surfers don't ride alone: virtual community as community, In Wellman, B. S. (ed), *Networks in the Global Village: Life in Contemporary Communities*, Westview, Boulder, 331-367.