

The Mechanism of Spontaneous Order in Online Knowledge Sharing Community: Taking Wikipedia as an Example

Xiaoyu Li

School of Information Management
Wuhan University
Wuhan, China
xiaoyuli@whu.edu.cn

Abstract—In the background of information explosion and big data problem, online knowledge sharing communities confront new challenge of the mechanism of knowledge organization and sharing. Several of these communities, especially Wikipedia, base the knowledge sharing mechanism on spontaneous order by community users. By analyzing the network structure and evolution process of the community policy environment in Wikipedia, this paper finds out that the set of community policies and user guidelines has its rank of priority — some of which are of fundamental place and function. Different policy functions tend to converge in the evolving process, which proves the increasing stabilization of spontaneous order. By illustrating the mechanism of spontaneous order from the angles of user collaboration and CAS theory, this paper points out that the active user participation, the smooth information flow and the prudent control are key points for the Wikipedia knowledge sharing, which is also enlightening for understanding and constructing knowledge sharing mechanism in similar online communities.

Keywords—online community; knowledge sharing; spontaneous order; Wikipedia

I. INTRODUCTION

The users of Web2.0 online community cooperate under the regulations of the community, sharing experiences and knowledge. Along with the explosion of online data amount, traditional information organization and knowledge sharing mechanism can no longer cater to the new phenomena. The spontaneous order generated by community users increasingly plays a significant role in the encouragement and regulation of user behavior, which is also the distinctive representation of the community's structure and functions. The structure and evolution of spontaneous community order is an ideal lens through which the entire mechanism of knowledge sharing can be described and understood.

Wikipedia (www.wikipedia.org), as a free-editing online encyclopedia and community hinged on users' cooperation, is the model of advanced knowledge sharing platform as well as the new research object of information science. Wikipedia has experienced an obvious process from the chaos by massive editing towards an ordered structure. The edits on item pages increase exponentially; the in-degree and out-degree distribution of pages is subject to power law [1]; the topological structure of page networks is similar

with World Wide Web and subject to preferential attachment mechanism. The evolution of Wikipedia is self-similar; its exponential growth is based on large numbers of users [2]. By the comparison of the contribution made by users of different editing quantity [3] and in different time units [4], it can be found that the collaborated editing of vast numbers of users is the fundamental driving force of the self-organization and knowledge sharing.

Wikipedia users pay great attention to self-governance and the construction of rules in the community, among which the community policies and user guidelines are essential. Wikipedia Policy is the most authoritative rule within the community by which users must abide. Guideline is also the legitimate rule of Wikipedia but much more flexible than Policy, and the users wouldn't be forced to follow them. Policy Environment composed by Policy and Guideline is a significant parameter of the spontaneous order and mechanism of knowledge organization and sharing of Wikipedia; it can be regarded as the literal expression of the process of user collaboration and a special function entity promoting the evolution of the entire community [5].

II. METHOD

The statistic date of historical growth of Wikipedia and 242 policy pages of Policy, Guideline and their sub-pages of en.wikipedia.org are gathered and downloaded from stats.wikimedia.org and download.wikipedia.org as of April 15th 2009. By gathering the hyperlinks among the community policies and user guidelines, the author analyzes the citation relationship of the policy pages and points out the roles of different categories of policies playing in the construction of Policy Environment. Furthermore, the author demonstrates the process of self-evolution by means of calculating the numbers of users' edits on different policy pages and related talk pages in the time dimension. On such base, the process of which users generate content and share knowledge cooperatively and maintain the spontaneous order in the community is explored from the perspective of Complex Adaptive System (CAS).

III. STRUCTURE OF WIKIPEDIA POLICY ENVIRONMENT

Wikipedia Policy Environment includes 46 Policies and 29 Guidelines. These Policies and Guidelines constitute the basic structure of Wikipedia Policy Environment.

There exist hyperlinks of mutual citation among Wikipedia policy pages. Some of these links indicate that particular policy is the subpage of another policy; some exist as a result of citing the existed policies to verify themselves; and others come from the functional modules within the policy pages such as Template and List etc. The author drew the oriented but unweighed links among policy pages to establish a network. Based on the classification standard in Beschastnikh's work [6], the Wikipedia policy pages can be divided into 9 categories according to their functions, as showed in Table 1. The author read all the 242 pages, classified them, labeled them with different colors and drew the network with UCINET, as showed in Fig. 1. In this network, there are 242 nodes and 3873 ties; the density is 0.0622, in-degree centralization 0.875, out-degree centralization 0.158, cluster coefficient 0.588, and average path length 2.356. The size of the nodes depends on the in-degree which shows the degree of a particular policy page cited by other pages; hence the analysis of these components leads to the discovery of the basic structure of the Policy Environment.

TABLE 1. FUNCTIONAL CLASSIFICATION OF WIKIPEDIA POLICIES

Code	Type	Description	Example
r	attribution	Citing references	Reliable sources
s	consensus	Consensus-seeking process	consensus
b	bias	Neutrality of content or organization	Neutral point of view
d	disposition	Legitimacy of user actions or intent	Etiquette
o	writing style	Low level style or organization	Manual of Style
g	genre	Defines Wikipedia	NOT
i	inclusion	Non-legal issues of including content	Notability
l	legal	Legality of content or user actions	Copyrights
u	User account	Defines user account	Username policy

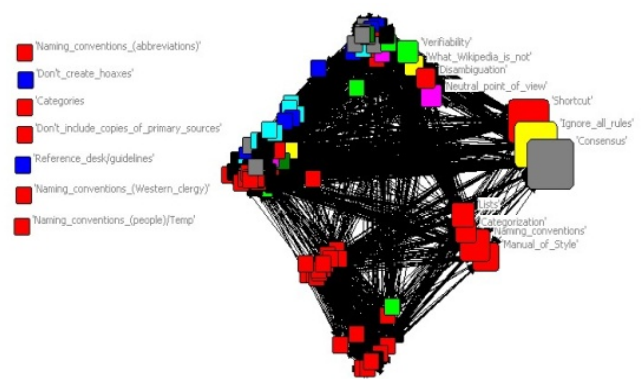


Fig. 1. Principle Component Diagram of Policy Page Network with Type Division

Ranking the in-degree in descending order, Fig. 2 shows the in-degree distribution and the types of top 50 policy pages in histogram. It can be found that the 3 policies with the largest in-degree are: (1) Consensus, (2) Ignore all rules, (3) Shortcut. The policies seeking consensus and defining Wikipedia, with relatively high frequency of citation in the whole policy structure, take an essential position and form the principles of the construction and operation of Wikipedia. The Policy and Guideline demonstrating the manner and style of editing are the largest in number, which shows that Wikipedia is meticulous and comprehensive in editing those policies concerning the editing instructions for users to improve the quality of item pages. The histogram shows apparent exponential characteristic of the in-degree distribution.

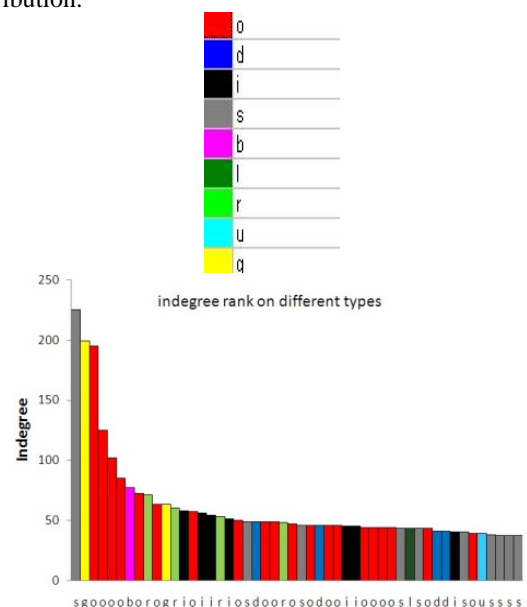


Fig. 2. Functional Classification of the Former 50 Policies Pages

IV. EVOLUTION OF WIKIPEDIA POLICY ENVIRONMENT

The evolution process of Wikipedia Policy Environment refers to the process of the generation, development and maturation of Policy and Guideline pages. The author analyzed the historical evolution of the 242 policy pages. Fig. 3 shows the tendency of accumulated number of the policy pages. The accrument of the total number of policy pages is nearly linear and it began to slow down in the latter half of 2007. During the entire history, the average number of editions on the policy pages per capita is 2.76. However, the average number of editing the talk pages per capita is 6.84, far more than the former. Around February 2006, the exponential increase of the number of users and number of editing on both policy pages and talk pages came to an end and were ushered in a relatively stable state; it is presumed that this is the result of the initial forming of policy environment. It is demonstrated in Fig. 4 that since around the second half of 2005, the average number of editing policy pages has been relatively stable around the mean. And the number of editing talk pages per capita still increases, which shows that the stability of policy environment is relatively dynamic and users are still able to continue discussing and editing the existing Policy and Guideline. The first Wikipedia policy was found on August 17th, 2001, “Wikipedia is not a dictionary”, and its latest edit in the data sample was on April 7th, 2009. The users are still heatedly discussing the applicability of this very first policy and haven’t reached agreement.

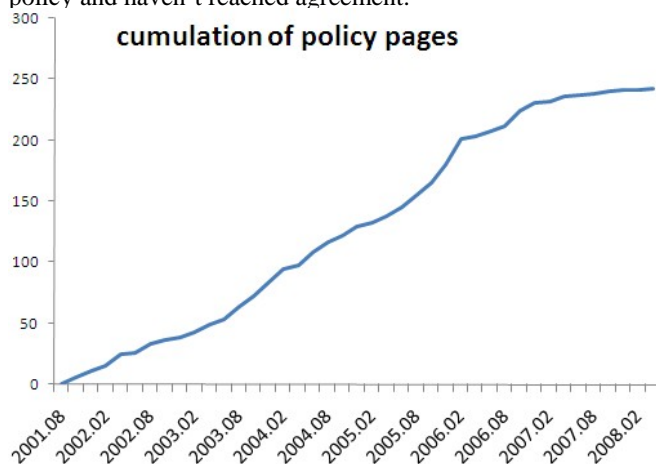


Fig. 3. Growth Tendency of Policy Pages

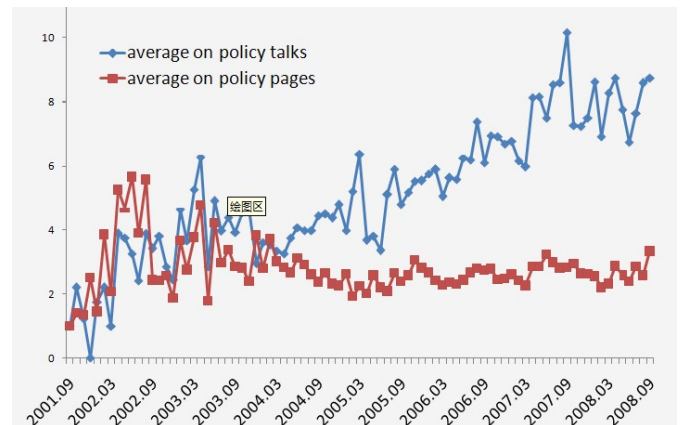


Fig. 4. Average on Policy Pages and Policy Talks

Fig. 5 demonstrates the evolutionary process of founding time and proportion of policy categories. The earliest policy defines Wikipedia (genre): “Wikipedia is not a dictionary.” Later there appeared the policy concerning the non-legal issues of including content: “Patent nonsense”. What content can be included and cited legitimately in Wikipedia is the problem needs solution at the beginning of the founding of Wikipedia. Then appeared the Guideline concerning the manner of editing (writing style): “Manual of Style” and it’s more and more important to regulate the manner of editing items as the number of user increases. These earliest three kinds of policies appeared collectively from August to October, 2001 and they cover the fundamental aspects related to user cooperation in online knowledge sharing community very well. Of all the policies, the consensus-seeking policy, though taking an important position, appeared later than all other policies with the Guideline of “Edit War” appeared on April 26th, 2003. With the increase of the user number, gaining consensus successfully has become the essential community culture to guarantee the editing efficiency.

From August 2001 to April 2003, Wikipedia policy pages generated rapidly, and proportions taken by different categories changed radically. From August 2003 to February 2005, Wikipedia Policy Environment experienced a relatively stable adjustment period when the existed policy pages were still under editing and the generation of new policy pages and subpages tended to be slow. After February, 2005, the proportions of the policies with different functions in the whole policy environment stabilized, which means the functions of policy environment have been mature.

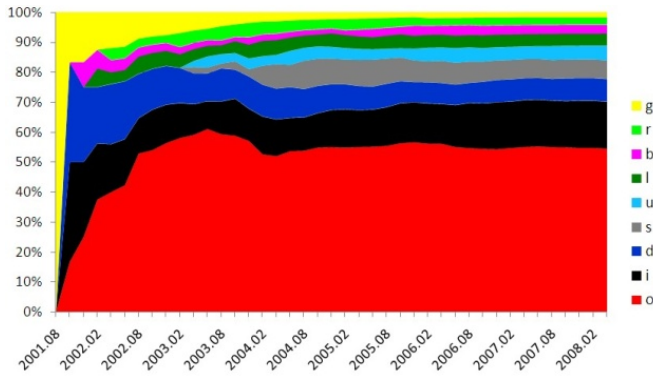


Fig. 5. Evolution of Policy Pages of Different Functions

V. MECHANISM OF SPONTANEOUS ORDER IN WIKIPEDIA BASED ON CAS

Complex Adaptive System (CAS) is an important branch of the systematic self-organization theory and an indispensable dimension to understand the evolutionary process and mechanism of spontaneous order of online community [7].

The basic idea of CAS theory can be generalized as follows. The system is composed of multiple adaptive agents which are neither passive nor static. These agents can constantly interact with other agents in the system to accumulate experience according to which these agents alter their structures and actions to better adapt themselves to the environment. Due to the fact that adaptive agents actively interact with other agents and the environment, the system manifests the feature of emergence. The constant pursuing by the agents to their niches leads to the functional and structural change of the whole system [8]. CAS includes 7 basic concepts, namely 4 properties: aggregation, non-linearity, flows, diversity, and 3 mechanisms: tagging, internal models and building blocks. Other characteristics of complex system can be derived from certain combinations of the 7 points.

A. Adaptive Agents in Wikipedia

In the Wikipedia policy environment, the implementation of the community culture as well as the aim and manner of cooperation is the process of transforming common users to adaptive agents. Besides going through Policy and Guideline, a more substantial way for the users to learn community rules is to review the editing history of pages and to learn from more proficient users [9]. During the growth of Wikipedia, users have been adjusting the rules by consensus-seeking to meet the challenges. Besides, the community makes use of some mechanisms, such as choosing featured items through public appraisal, Wikipedia Barnstar [10], and choosing administrators through public appraisal [11], to encourage users to actively take part in the knowledge construction and sharing in the community.

The Wikipedia technology with various functions facilitates the appearance and improvement of the adaptability of agents. The adaptive behavior of agents can

be the coexistence of competition and cooperation. The competition refers to the defending, maintaining and updating of one's own ideas. And cooperation exists among users [12] [13]; collectively the users participate in the construction of the huge amount of online knowledge and the maintenance of the community order. Wikipedia protects cooperation in order to guarantee the success of community collaboration.

B. Key Points of the Spontaneous Order in Wikipedia

1) *Aggregation*. In CAS, similar agents aggregate through interaction, forming small aggregations which snowball into big ones. Through such hierarchical aggregation, the layered structure within the system manifests complexity in the macroscopic view. The small groups in Wikipedia formed by users with different interests are the demonstration of such aggregation [14]. These small groups have highly efficient edits on their specialized fields and show enthusiasm for some particular maintenance of the community. The folksonomy of Wikipedia creates the macroscopic orderliness in its taxonomy. Such folksonomy is another form of aggregation based upon the content of the pages.

2) *Nonlinearity*. In CAS, the interaction between agents and environment and among agents is complex, dynamic and nonlinear. The whole of the system surpasses the sum of its parts, whether from the angle of structure or function. Such nonlinearity causes the complexity of the communication among the adaptive agents. This communication leads to abundant outcomes. Both the quality and quantity of the items and the number of users and edits have experienced a process of nonlinear exponential increase. Cooperative edits by large number of users guarantee Wiki's fast response to false items [15] [16].

3) *Flows*. Among agents as well as between agents and the environment exists exchanges of flows of materials, energy and information. Smooth channels for these flows play an indispensable role in the evolution of the system. The major flows of Wikipedia system are information flows; and the most important key point of Wikipedia spontaneous order is to make sure the smooth flows of information to which the open editing principles, handy editing manner and the prudence of users' edits make collective contribution. With smooth flows of information, it's easier for user agents to cooperate and more effective to supervise the community.

4) *Diversity*. Diversity refers to the differentiation among agents, which is the reason for agents' interaction and the premise of cooperation. The diversity of the Wikipedia user agents comes from different languages, cultures, beliefs, interest and education backgrounds as well as the division of labor and belonging to various small groups. Such diversity lays the foundation for the complexity of Wikipedia.

C. Mechanism of Spontaneous Order for Knowledge Sharing in the Community

1) *The first and foremost key point of Wikipedia is to ensure sufficient diverse and active Adaptive Agents.* Wikipedia makes sure the agents are adaptive via a series of incentive and control mechanisms which are the foundation stones for the community order. The incentive mechanism stimulates users to join in the discussion of community rules [17] while the control mechanism asks for rather cautious edits on policy pages. Some policy pages are of rigid priority, which do not encourage much revision [18]. In Wikipedia, the incentive mechanism is ensured by the official authority of the community to a great extent while the control mechanism is enforced and participated by common users generally. This is far beyond our traditional thinking of management. First, the target of collaboration can be established with the help of the community culture; and the identity with the collaborating target results in the policies and guidelines made by the users. Second, with the help of wiki-tech, Wikipedia has made great efforts to reduce the restrains for editing and users' behaviors, hence the users will find the item pages they are interested in and other users who share the same interest quickly so that the efficiency of aggregation will be improved. Third, to ensure the users to be positive, Wikipedia decentralizes the management into general editing process of common users, and mediation and arbitration from higher authority are asked to be operated cautiously only when the seeking for consensus among users fails; the decentralizing management system shows high efficiency for the supply of public order.

2) *Wikipedia realizes self-organization through ensuring smooth flows of information.* Competition and cooperation are the major behaviors of user agents in the system. The interaction of competition and cooperation forms the driving force for the self-organization of the system; competition keeps the system away from the equilibrium state while cooperation magnifies the motions of the non-equilibrium sub-systems. In such process, smooth flows of information help users to choose the strategy of competition or cooperation, as well as the degree of participation. During the process of editing users constantly learn and reinforce their behavior strategy which, based upon users' cognition of the system, is the internal model of the evolution of the system's self-organization. The model, existing in the brains of user agents and reflected through behavior, is the mechanism of building block in the theory of Complex Adaptive System. Through users' selection and learning, the permutation and combination of effective internal models produce new internal models. The Wikipedia users incessantly learn and accumulate experience upon which they form new strategies in order to adapt to the new environment. In this regard, the

spontaneous order emerged in the system comes from the optimal combination of each user's behavior model, and the macroscopic order of Wikipedia comes from the order of the cognitive structure and knowledge structure of its users.

3) *Prudential authority is an external force effecting on the system.* In Wikipedia, the spontaneous order does not only come from the users, but also from its relatively simple but effective hierarchical structure. The management functional entity in Wikipedia such as Arbitration Committee plays a significant role in preserving the ordering process in the entire community. Even the power originating outside the community, say, the decision of Jimmy Wales, is also very important to answer the controversial but fundamental questions and keep the general daily work of Wikipedia from endless debate. Such decisive power is a kind of protection mechanism for the Wikipedia community to pursue the efficiency of users' collaboration. However, the principle of exerting such power should obey the object of Wikipedia community. The coercive decision must come after all the possibilities of seeking consensus by the users.

VI. CONCLUSION

This paper sums up the key points of the mechanism of spontaneous order in online knowledge sharing community by means of exploring the mechanism of spontaneous order of Wikipedia. First, sufficient adaptive agents are the fundamental driving force of the whole evolutionary mechanism of spontaneous order. Second, smooth flows, including expression behavior and management, secure the operation of the mechanism. Third, cautious control is the external force to facilitate the evolution of the mechanism.

Online knowledge sharing community with huge amount of data resource, especially in the background of big data, confronts the challenge of exploring new knowledge organization mechanism. Spontaneous order generated and maintained by community users has expanded our view of new types of knowledge organizing and sharing and deepened our thinking of this topic. It offers a practical and effective way for organizing vast data utilizing decentralized system of user self-organization. Although Wikipedia is a special example, the mechanism based on user collaboration solving technological problem is worthy noticing by both the Internet industry and the academic research.

ACKNOWLEDGMENT

This research has been made possible through financial support by the National Science Foundation of China under grant number 70833005.

REFERENCES

- [1] L. Buriol, C. Castillo, D. Donato, S. Leonardi, and Millozzi, S. "Temporal Analysis of the Wikigraph," In Proc. of the 2006 IEEE/WIC/ACM International Conference on Web Intelligence, Hong Kong, China, December 18-22, 2006, IEEE Computer Society Washington, DC, pp.45-51.

- [2] R. Almeida, B. Mozafari, and J. Cho, "On the Evolution of Wikipedia," In Proc. of Int. Conf. on Weblogs and Social Media, Boulder, Colorado, U.S.A., March 26-28, 2007.
- [3] A. Kittur, B. Suh, B. Pendleton, and E. Chi, "He Says, She Says: Conflict and Coordination in Wikipedia," In Proc. of the SIGCHI conference on Human factors in computing systems, San Jose, California, U.S.A., April 28- May 3, 2007, CHI'07. ACM, NY, pp.453-462.
- [4] F. Ortega, and J. Barahona, "Quantitative analysis of the wikipedia community of users," In Proceedings of the 2007 international symposium on Wikis, Montréal, Québec, Canada, Oct. 21-23, 2007. WikiSym'07. ACM, New York, NY, pp.75-86.
- [5] B. Butler, E. Joyce, and J. Pike, 2008. "Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in wikipedia," In Proceeding of the twenty-sixth annual SIGCHI conference on Human factors in computing systems, Florence, Italy, April 5-10, 2008, CHI'08. ACM, New York, NY, 1101-1110.
- [6] I. Beschastnikh, T. Kriplean, and D. McDonald, "Wikipedian Self-Governance in Action: Motivating the Policy Lens," In Proc. of Int. Conf. on Weblogs and Social Media, Seattle, Washington, U.S.A., March 30- April 2, 2008, Association for the Advancement of Artificial Intelligence Press, pp.27-35.
- [7] F. Ma, and Y. Xia, "Mechanism of Ordering in Wikipedia Based on the CAS Theory," Library Tribune, 2008, vol. 28(6), pp.85-92.
- [8] J. Holland, *Emergence: from chaos to order*. Shanghai Scientific & Technical Publishers, Shanghai, 2001, p. 246.
- [9] S. Bryant, A. Forte, and A. Bruckman, "Becoming Wikipedian: Transformation of Participation in a Collaborative Online Encyclopedia," In Proceedings of the 2005 international ACM SIGGROUP Conference on Supporting Group Work, Sanibel Island, Florida, U.S.A., November 06-09, 2005, CSGW'05. ACM, New York, NY, pp.1-10.
- [10] T. Kriplean, I. Beschastnikh, and D. McDonald, "Articulations of WikiWork: Uncovering Valued Work in Wikipedia through Barnstars," In Proceedings of the ACM 2008 conference on Computer supported cooperative work, San Diego, California, U.S.A., November 8-12, 2008, CSCW'08. ACM, New York, NY, pp.47-56.
- [11] M. Burke, and R. Kraut, "Mopping up: modeling wikipedia promotion decisions," In Proc. of the 2008 ACM conference on Computer supported cooperative work, San Diego, California, U.S.A., November 8-12, 2008, CSCW'08. ACM, New York, NY, pp.27-36.
- [12] F. Viégas, M. Wattenberg, and K. Dave, "Studying Cooperation and Conflict between Authors with history flow Visualizations," In Proc. of the SIGCHI conference on Human factors in computing systems, Vienna, Austria, April 24-29, 2004, CHI'04. ACM, New York, NY, pp.575-582.
- [13] F. Viégas, M. Wattenberg, J. Kriss, and F. Ham, "Talk Before You Type: Coordination in Wikipedia," In Proc. of the 40th Hawaii International Conference on System Sciences, Waikoloa, Big Island, Hawaii, Jan. 03-06, 2007, HICSS 2007. IEEE Computer Society, Washington, DC, pp.78-87.
- [14] P. Konieczny, "Governance, Organization, and Democracy on the Internet: The Iron Law and the Evolution of Wikipedia," *Sociological Forum*, 2009, vol. 24(1), pp. 162-192.
- [15] B. Stvilia, M. Twidale, L. Smith, and L. Gasser, "Information Quality Work Organization in Wikipedia," *JASIST*, 2008, vol. 59(6), pp. 983 – 1001.
- [16] D. Wilkinson, and B. Huberman, "Assessing the Value of Cooperation in Wikipedia," <http://arxiv.org/abs/cs/0702140>, 2008-02-01.
- [17] A. Forte, and A. Bruckman, "Scaling Consensus: Increasing Decentralization in Wikipedia Governance," In Proc. of the 41st Hawaii International Conference on System Sciences, Waikoloa, Big Island, Hawaii, January 07-10, 2008, HICSS 2008, Hawaii, pp.157-167.
- [18] T. Kriplean, I. Beschastnikh, D. McDonald, and A. Golder, "Community, consensus, coercion, control: cs*w or how policy mediates mass participation," In Proceedings of the 2007 international ACM conference on Supporting group work, Sanibel Island, Florida, U.S.A., Nov. 04-07, 2007, GROUP'07. ACM, New York, NY, pp.167-176.