Research Policy 44 (2015) 1226–1239

Contents lists available at ScienceDirect

Research Policy

journal homepage: www.elsevier.com/locate/respol

Big from the beginning: Assessing online contributors' behavior by their first contribution

Sylvain Dejean [a], Nicolas Jullien [b,*]

[a] Université de La Rochelle, CEREGE, France
[b] Institut Mines Télécom Bretagne, iSchool, ICI-M@rsouin, France

ARTICLE INFO

Article history:
Received 28 October 2013
Received in revised form 3 March 2015
Accepted 6 March 2015

JEL classification:
L17
L86
H41

Keywords:
Voluntary participation
Open online communities
Survey
Econometric studies
Heckman
Wikipedia

ABSTRACT

The objective of this paper is to investigate the process of involvement in open online communities producing knowledge, via the link between the first contribution and the level of contribution reached. While most studies consider the career of contribution following the first contribution, we focus on what happened before and during the first contribution. We challenge the fact that becoming a core member starts with peripheral contributive activities and results from a continuous learning process, as explained by the theory of community of practice. On the contrary, and coherent with epistemic community theory, our results, based on 13,000 answers to a survey on the use of, and contributions to Wikipédia, show that the future level of users' involvement depends on the period of time between the discovery of Wikipedia and the first contribution (negative effect), and of the effort made in the first contribution (positive effect). Implications for management are also discussed.

1. Introduction

The mobilization of hundreds (Linux) or thousands of contributors (Wikipedia), volunteer online open projects aimed at creating new knowledge, online "communities of creation", as named by Rullani and Haefliger (2013), is viewed as central in the generation of new, innovative knowledge by and for firms. But the path to successful community building is still risky and uncertain, and as for business building, most of the attempts fail, no matter how many hundreds of thousands of dollars were spent (Worthen, 2008).

One of the key elements in the development of a successful and sustainable community, as explained a quarter of century ago by Eric von Hippel (1986), is to attract enough highly competent and "committed/committing" contributors, being they named "lead users", "core", or "big" contributors (Mahr and Lievens, 2012; Fang and Neufeld, 2009), i.e. the most productive people, who are also those with more responsibility in the management of the project (Rullani and Haefliger, 2013). This focus is explained by the fact that, since both big and small contributors are needed (O'Mahony and Bechky, 2008; Hemetsberger and Reinhardt, 2009), as in most collective actions and projects (Oliver et al., 1985; Ostrom, 1990), the former are much rarer than the latter, in addition to being more productive (Lakhani and von Hippel, 2003; Maillart et al., 2008; Voss, 2005).

New big contributors are constantly needed as they are subject to disengagement after some years (Ortega and Izquierdo-Cortazar, 2009 for open online communities, Borzillo et al., 2011 for intra-organization communities), and they are difficult to recruit and retain (Von Krogh et al., 2003 in the case of open source software communities, Halfaker et al., 2013 for Wikipedia). This echoes more general findings about the efficiency of groups. As shown by Uzzi and Spiro (2005) in the case of musical comedies, and Uzzi (2008) in the case of a social network, for a creative group to be successful, it needs to fine tune the level of newcomers to an already constituted group, for fresh ideas (Guimera et al., 2005; Defélix et al., 2005). Wikipedia, for instance, is known for its gender bias amongst editors, which would explain why certain topics are less well covered

* Corresponding author. Tel.: +33 229 001 245.
E-mail addresses: Sylvain.Dejean@univ-lr.fr (S. Dejean), Nicolas.Jullien@telecom-bretagne.eu (N. Jullien).

http://dx.doi.org/10.1016/j.respol.2015.03.001
0048-7333/

by the encyclopedia.[1] Specific programs, targeting new contributors, or "newcomers", have been designed to facilitate the supposed learning curve leading to regular contribution, thus "mentoring" them (Wikipedia program terminology). This means, using Kram (1983)'s typology, assigning to each newcomer (newly registered person or beginner contributor) a volunteer coach/counselor, who is a regular contributor, to guide him/her through the different contributing steps and rules (see Mateos-Garcia and Steinmueller, 2008 for open-source project Debian, and Musicant et al., 2011, for Wikipedia).

In this article we ask whether it is possible to identify the different contributors early enough in the process to adapt the mentoring to their profiles, and thus make it possible to decrease the high discouragement rate of both newcomers and mentors (Musicant et al., 2011). This question is based both on studies showing that those future big contributors may be identified from their very first contributions (Fang and Neufeld, 2009 for open source, Panciera et al., 2009 for Wikipedia), and on the theory of epistemic community (Cohendet et al., 2001; Edwards, 2001), which stresses that those communities are project-oriented communities of experts, whose expertise is acquired outside the community. According to this theory, entering a community is equivalent to starting to contribute: there is no peripheral participation.

In addition to providing the knowledge-community managers with results on how to better recruit future core contributors, this article aims at contributing to the studies of social practices in context and to the characterization of such online communities, leveraging on and discussing previous works such as Amin and Roberts (2008)'s on the different models of "knowing in action" (craft/task-based, professional, epistemic/creative, virtual), and Rullani and Haefliger (2013)'s analysis of the dynamics of intra-organizational communities of creation. This contribution is based on econometric analyses of a survey of more than 13,000 Wikipedia users and sometime contributors, their first contribution and their current level of contribution.

The article is organized as follows: in Section 2 a review of the literature used to construct our framework of investigation, in Section 3 the formulation of our hypotheses, in Section 4 the data collection strategy (choice of the community and definition of the questions), in Section 5 the results. We discuss the consequences of this work, its limits and future research in Section 6 before concluding.

2. A career in communities of creation: from peripheral to big contributor?

There is a consensus that big and small contributors do not have the same aims when contributing: in the case of open source software, Shah (2006) showed that regular contributors enjoy programming and interacting with the rest of the community (i.e., labeled as "hobbyists"), whereas new or sporadic contributors are typically driven by an immediate need for software (i.e., use value). For the most involved Wikipedians, the recognition from their peers ('credit') is an important motivation (Forte and Bruckman, 2005; Bryant et al., 2005), as is the sense of mission (Liang et al., 2008; Prasarnphanich and Wagner, 2009); for most of the (small) contributors, the will to fix mistakes is the principal motivation, meaning that they are not strongly committed to the project (Kamata et al., 2010). According to Shah (2006), this echoes the more general sociological notion of "career" (Becker, 1960, 1963), which stresses that people's motivations and actions are curved by the social interactions they meet in their practice. In other words, individual involvement means two things in these communities: knowledge production and acquisition activities, and the social process inherent to any group, since interacting with people leads to (is equivalent to?) socializing. Those two facets are well illustrated by Butler et al. (2007), who stressed that in those groups, people come for the knowledge and stay for the people.

Considering this, and because it stresses these two facets, the concept of legitimate peripheral participation, or LPP (Wenger, 2006, 1998), has been used to explain how contributors learn how to propose new knowledge, in the online communities: future contributors start to get involved by observing, "dipping their toes in to passively participate while learning more about a complex system" (Antin et al., 2012) before editing (for Wikipedia) or coding (for open-source), then interacting with the experienced member at the margin, and so on. There would appear to be a slow process of "decantation" in the group of readers and early contributors leading to the emergence of regular contributors, who, at the same time, are increasingly socially involved.

This argument is, however, theoretically and practically disputable, at least with reference to the knowledge production, which is what we are looking at here. Although epistemic communities and communities of practice are probably similar regarding the socialization process, they are quite different, regarding knowledge production. Theoretically, LPP is considered in communities of practice, to be like professional/specialized forums (online or, as in Wenger (1998)'s studies, local, geographically situated), where people exchange primarily about their "practices" and build their knowledge of those practices. Participation is a process, where people first observe, then make minor contributions, and gradually increase their engagement and the complexity of their contributions. Conversely, communities of creation are (virtual) epistemic communities, or task-oriented groups, which brings experts together around a common goal (Amin and Roberts, 2008)[2]: the building of (new) knowledge. People, even newcomers, are evaluated on their capacity to produce this knowledge,[3] on the basis of competencies acquired mainly outside the community and before starting to contribute (Edwards, 2001). In other words, from the beginning, some contributors are more able to propose major contributions than others, and go on to make significant contributions (there is no gradual engagement).

The argument about the differences between big and small contributors being an indicator of a learning period can be overturned: those who are the most willing to become regular contributors have, from the beginning, different capacities and goals, which are simply not statistically discernible because of the mass of lurkers. The studies comparing current small contributors and current big

[1] On that matter, the MIT Technology Review proposes a good introduction of the recruiting problem and of its consequences: http://www.technologyreview.com/featuredstory/520446/the-decline-of-wikipedia/

[2] On that matter, we follow Amin and Roberts (2008)'s analysis of the different forms of "knowing in action", except on one point: they distinguish between virtual communities and epistemic ones, on the single basis of the nature of the communication between people (face-to-face or virtual, see Table 2, page 257), when the other communities are segmented by the type of knowledge and the competences of the people involved in. In fact, Cohendet et al. (2001), Rullani and Haefliger (2013) who agree on the other parts of the analysis do not support that particular part. Following their example we will rely on Amin and Robert's analysis of the knowledge production, and will not discuss the impact of the mean of communication on the exchange, which lies beyond the scope of this study.

[3] For Wikipedia, when projects have rules for running for the position of administrator, they are about knowing the rules, but also about the number of edits of articles (more than 3000 and more of one year of activity for the French Wikipedia, https://fr.wikipedia.org/wiki/Wikipédia:Candidature_au_statut_d'administrateur). O'Mahony and Ferraro (2007, part II), on Open Source projects, showed that "developers who were making greater technical contributions (in terms of impact but not effort) and who were more engaged in organization building were more likely to become members of the leadership team". (p. 1096). Fleming and Waguespack (2007) found the same result in their study of the Internet Engineering Task Force community.