

A Survey on Firms' Participation in Open Source Community Projects

Eugenio Capra, Chiara Francalanci, Francesco Merlo
Dipartimento di Elettronica e Informazione, Politecnico di Milano
{capra | francala | merlo }@elet.polimi.it
Cristina Rossi Lamastra
Dipartimento di Ingegneria Gestionale, Politecnico di Milano
cristina1.rossi@polimi.it

Abstract. The role of firms in commercial Open Source projects (e.g., former MySQL, EnterpriseDB, SugarCRM) is a consolidated and generally accepted fact. On other hand, community Open Source projects, which are built upon communities and not directly associated with firms, are commonly perceived to be based mainly on the work of volunteers. Up to now, firms' role in these projects has been poorly investigated. We conducted a survey on 1,302 *SourceForge.net* projects to inquire about the level and the typology of involvement of firms. We propose three different models for firm participation and provide empirical evidence on their diffusion in *SourceForge.net*.

Keywords: Community Open Source Projects, Firms' Participation

1 Introduction

Open Source (OS) was born as a strong ideological-oriented movement mainly supported by volunteer developers [21]. OS is commonly associated to peculiar project governance models, which promote openness and the freedom of developers [21,28] as opposed to closed software traditional governance models, which are focused on customers' needs, corporate hierarchies and tight planning.

Nowadays, firms' involvement in OS is a consolidated fact. The largest IT and software companies are investing in OS [3, 22]: Eclipse has been released under an OS license, MySQL has been acquired by Sun Microsystems. Moreover, the number of firms that are building their business on OS is increasing (e.g., Funambol, SugarCRM, Jaspersoft, Alfresco, Zimbra).

Plenty of evidence exists that firms actively participate in OS projects with different business goals and approaches [9, 22]. In this respect, an important distinction should be made between *i*) commercial OS projects, and *ii*) community OS projects [31].

Commercial OS projects (e.g., former MySQL, EnterpriseDB, SugarCRM, JasperSoft, Zimbra, Alfresco, Funambol) produce software released under open licenses according to the OSI standard [24] and encourage participation from the community. However, they are entirely led by a firm, which controls the access to the code base, defines the evolution strategy of the project, and sets the implementation roadmap [3,9,11]. These firms view OS as a new business model that leverages the community as a source of feedback, testing, and marketing, and aim at selling the whole range of services built around their software, such as customization, maintenance, training, and hosting.

Community OS projects are led by community members and are not directly associated to a firm. They are usually hosted on online software repositories, such as *SourceForge.net*, *Apache Foundation*, or *Tigris.org*.

A common misconception is that community OS projects are sustained only by volunteers. Although there is no single firm owning the project, evidence exists that in many projects some or most developers can be employed by a firm and paid to take active part in the community and develop specific components of the software program according to their firm's needs [31]. Code contribution is a common and effective strategy for firms to contribute to OS projects, as it allows them to directly impact on the final product. However, there is a number of different typologies of involvement other than code development activities. Firms may test the software, report bugs, contribute to packaging, suggest functional requirements, write documentation, or simply participate and animate forums dedicated to the project [11]. In addition, firms may provide financial, logistic, and marketing support, or generally coordinate the project [10].

A number of recent studies have aimed at understanding the benefits that the OS paradigm may bring to companies from a business perspective [9,12,22], but the role of firms in community OS projects is still poorly investigated.

This paper proposes different models for firms' participation in community OS projects and provides empirical evidence related to the *SourceForge.net* online software repository.

Our work is based on the results of a large-scale survey sent to the administrators and users of 1,302 *SourceForge.net* projects. Basing on our data, we propose three different models for firms' participation in community OS projects and provide evidence on their diffusion among the projects hosted on *SourceForge.net*.

The paper is organized as follows. Section 2 proposes some models for firms' participation in community OS projects that we have identified from the literature and case studies. Section 3 describes the empirical methodology adopted and the sample of projects that we considered. Section 4 reports our results. Finally, Section 5 concludes the paper and proposes future developments of our research.

2 Models of firm participation in community OS projects

The incentives for a firm to participate in an OS project are different from those of volunteer developers. A whole stream of literature has focused on the analysis of why independent contributors might freely devote time and resources to community OS projects: apart from idealistic motives for participation [5,23], the most common incentives to participation are related to career concerns [18] and personal satisfaction [14]. Conversely, firms are profit-oriented agents looking for an economic return [1]: their primary goal is typically to get tangible benefits from the participation in OS projects. The business objectives related to the participation in OS projects can be achieved by performing a wide spectrum of activities: in the following we present three different models of firm participation in community OS projects derived from the analysis of the literature and of real cases. Along with the description of each model, we discuss the main activities that firms actually perform when involved in community OS projects. The models have been refined by analyzing the answers to the survey that will be described in Section 0.

2.1 The coding model

The most simple way a firm can get involved into a community OS project is by performing code-writing activities. In some cases, firms can be directly involved in the development of a project together with external voluntary contributors. For example, the development of the Asterisk VoIP telephony engine has been started and is currently performed by Digium Inc. together with the users community [24]. In other cases, firms tend to develop some components of an application internally as closed-source software, while they release to the community all the components that do not constitute a differentiation factor and are not critical to competitive advantage [14]. This process of selective revealing varies with the specific characteristics of each firm [15]. For example, a firm could decide to release the source code of a non strategic application in order to increase the demand of complementary services and tools [18]. Other firms often pay their employees to participate in OS projects in order to gain an influential position and drive future developments [5,18].

The main activities that firms perform when involved in participation by means of the coding model, obviously beside *writing code*, comprise *fixing bugs*, *customizing* the software for customers with specific needs, *providing packaged distributions* of the software by adding installation wizards, guided procedures for configuration, or providing compiled binaries instead of plain source code. This model of participation also takes into account situations in which firms *donate a pre-existing codebase* to the OS community to switch from closed-source to open, or *start from scratch a new project* by founding the community and subsequently performing one or more of the activities discussed above.

2.2 The support model

Beside the contributions that entail code-writing activities, a firm can be involved in many other ways into an OS project. Firms are often sponsors of OS projects, providing financial or logistic support [10]. For example, the Firefox Internet browser, as well as the Thunderbird email client, are developed by the Mozilla Foundation, which is in turn sponsored by many firms that are not directly involved in code-writing activities, such as Google. Also in this case, companies expect tangible benefits from their involvement in OS projects: as noted by West and O'Mahony [32], beside direct economic returns, motives to support a community OS project comprise the need to achieve a wider adoption of an application in order to benefit from direct and indirect network externalities or the need to get development help on areas of low priority for the firm (e.g., in particular dialects).

When involved in an OS project by means of the support model, the typical activities that a firm performs comprise providing direct or indirect *financial support*, *logistic support* by supplying hardware resources, Web hosting or rooms, *marketing support* by advertising the project either on its Website or on search engines (for example, by advertizing the project on Google AdSense). Other typologies of support that firms can provide to a community OS project include the *participation to online forums* in order to keep the community alive by answering to users and customers, provide explicit *user or customer support* after the application has been installed, *test the product by using it* in a real production environment, *provide training* courses for users, *report bugs* found during everyday usage, or *write documentation* such as user manuals or installation guides.

2.3 The management model

The last model of firm participation in a community OS project entails all those activities related to project administration and management. In many cases, a firm might not be interested to directly participate in the development by performing code-writing activities, nor be interested in providing support to the project: as noted by Krishnamurthy [18], in some cases companies can play the role of coordinators of new OS communities, especially during the initial phases of development and creation of the community itself. In other cases, firms may have interest in joining the community and being involved as project administrators and coordinators in order to drive the development of the product to gain competitive advantages, or being able to set cooperation agreements with competitors in strategically differentiated areas of business.

For example, Novell is involved in the MONO Project through its subsidiary Ximian Group (acquired in 2003). Although performing some code-writing development activities, Novell acts also as a coordinator of development: as reported in the official Website of the MONO Project [24], Novell has set an agreement with Microsoft in September 2007 in order to coordinate the development of some subsystems of the MONO Project. Moreover, as reported in the official homepage

[30], the Ximian Group “*plays a central role in the Open Source community, providing leadership and core technology to key Open Source projects and industry groups*”.

The main activities that firms perform when playing a management role while participating in community OS projects comprise the *suggestion of requirements and functionalities* to be added to the software (either to achieve competitive advantages or to implement needed functionalities), *planning and designing* the application, or simply *coordinating the development*, as seen in the case of Novell.

The three participation models discussed above should not be interpreted as mutually exclusive. Clearly, a firm can be involved in a community OS project by performing more than one typology of activity at a time, mixing together the three models. As a result, basing on the firm's needs and expectations in participating a community OS project, each dimension can be more or less emphasized.

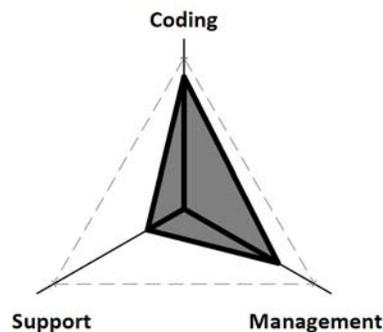


Fig. 1. Example of firm participation over multiple models.

For example, as in the case of Novell, a firm could be more interested in writing code and providing managerial and coordination support to an application in which it is involved more than providing explicit financial or logistic support. Fig. 1 provides an example of visualization of the discussed situation: the level of participation for each model is represented as an independent dimension, and the shaded area emphasizes the predominant dimensions of participation.

The different models of participation that a firm can adopt while involved in an OS project may change during the time. Depending on the stage of development, the feedback from the community, the evolution of business objectives and market shares, a firm can decide to switch between different models of participation in subsequent time instants. For example, consider the participation of IBM in the Jikes project, a high-performance Java compiler [24]. Initially, Jikes was started as an internal project of IBM, as a classic proprietary software development project. The binary form of the compiler was released in April 1997 on the AlphaWorks site, while in December 1998 an open version under the IBM Public Licence was released. Initially IBM maintained a central role, but with more and more contributions coming from voluntary developers, it started to slacken the managerial

pressure over the project. In March 2002 the project was migrated on the DeveloperWorks site, with IBM ceasing to perform developing activities, and providing only logistic support (basically, servers to host the community and the codebase). Finally, in 2005, IBM completely dismissed the project, which was moved to the *SourceForge.net* repository, and which now lives just as a community OS project. As stated in the Jikes project homepage [24], “*today there are no IBMers who work on Jikes as part of their job description*”.

3 Sample and methodology

Since our goal was to analyze whether and how firms participate in community OS projects, we conducted a survey on a large group of project administrators of OS community projects from the *SourceForge.net* repository.

We chose *SourceForge.net* because it is one of the most referenced online software repositories, both for number of projects (more than 136,000 as in October 2008) and for number of users. In addition to that, *SourceForge.net* projects are commonly associated to the most “open” side of the continuum between open and closed software development model approaches [2, 4, 19], and are rarely associated to firms. We chose to focus on only one repository because different repositories may have different rules and philosophies, whereas we aimed at analyzing comparable projects.

First, we defined the population of projects we wish to target in this study. Since mining online repositories (such as *SourceForge.net*) can lead to controversial results because of the varying quality of available data [17], we applied the following set of criteria:

- Project maturity: active and beta status or higher, according to the status reported by the home page of each project on *SourceForge.net*; inactive and less mature projects were excluded because of their instability and low significance.
- Development team: at least 2 team members, listed either as administrators or developers. One-developer projects were excluded as the community dimension is lacking.
- Programming language: Java; selected projects were restricted to the Java programming language since it is the most diffused on *SourceForge.net* (more than 25% of the total projects), and to avoid potential biases in the application sample due to differences in programming languages.

As in July 2007, we found 4,392 applications satisfying the selection criteria listed above. All the 7,760 *SourceForge.net* users listed as project administrators of our target population have been invited to participate to the survey. Table 1 provides some summary statistics on the number of administrators and developers of the applications of our population of projects. Whenever a user was found to be listed as administrator for more than one project, we required to compile a questionnaire for each project he or she was involved in. Considering multiple participation requests, a total of 8,780 inquiries about survey participation were sent.

Table 1. Summary statistics of projects sample (N=4,392)

Variable	Average	St.Dev.	Min	Max
Team members	5.07	7.00	2	141
Administrators	2.01	1.37	1	15
Developers	3.06	6.47	0	132

The questionnaire has been hosted on *SurveyMonkey.com*, a specialized website in online surveys. Before submitting the questionnaire to the selected administrators, we conducted a *pre-test* phase: we administered the questionnaire to a random sample of 195 administrators, stratified according to the size of the development team. The *pre-test* phase allowed us to identify a peculiar non-response bias, due to the fact that the e-mail by which we invited administrators to participate in the survey explicitly referred to firms' involvement and thus was considered more interesting by the administrators of the projects in which firms were actually involved. We eliminated every reference to the word "firm" in the cover e-mail and in the first page of the online survey, so reducing the impact of the bias.

The survey started on November 27th, 2007. After the first e-mails, two follow-ups were conducted, starting from December 5th, 2007 and December 18th, 2007, respectively. The online questionnaire has been available for completion for 8 weeks after the first e-mails were sent. A total of 1,833 questionnaires were completed and submitted (corresponding to 1,411 unique projects), resulting in a response rate of approximately 21%. Table 2 presents a summary of the three phases of the online survey, while Fig. 2 shows the temporal distribution of the received answers.

Table 2. Summary statistics of the online survey phases.

	First mailing	Follow up 1	Follow up 2	Total
Sent e-mails	8780	6855	4429	-
Answered	726	882	225	1833
Declined	27	141	30	198
Error/out of office	1172	1403	933	3508
No answer	6855	4429	3241	-

Whenever inconsistencies were found (e.g., inconsistent replies from different administrators of the same project), we contacted the respondents by e-mail and asked for clarification, even though we did not always received answers. We excluded from the sample projects for which we were not able to solve inconsistencies, thus resulting in a final sample of 1,302 projects.

The questionnaire has been structured in three parts. The first part addressed general questions about the project and questions to verify the information gathered through the *SourceForge.net* home page on the activity status of the project. The second part of the questionnaire focused on evaluating the number of firms involved in the project (if any) and their approaches of participation. We asked users and administrators if they are paid by a firm to specifically work on the project, either as employees or as external professionals or community members hired *ad-hoc*. In particular, we asked respondents to indicate which were the three most

significant (i.e., time consuming) activities they perform within the project on behalf of their firm. We inserted multiple choice questions with all the different activities described in Section 0, i.e. code development, testing, bug fixing, writing documentation, writing in forums, managing and animating forums, supporting users, managing the website, planning and designing, defining and proposing requirements, coordinating. We also included free text fields to be sure to encompass all the possible activities performed in a project. This helped us to integrate and refine the models proposed in Section 0.

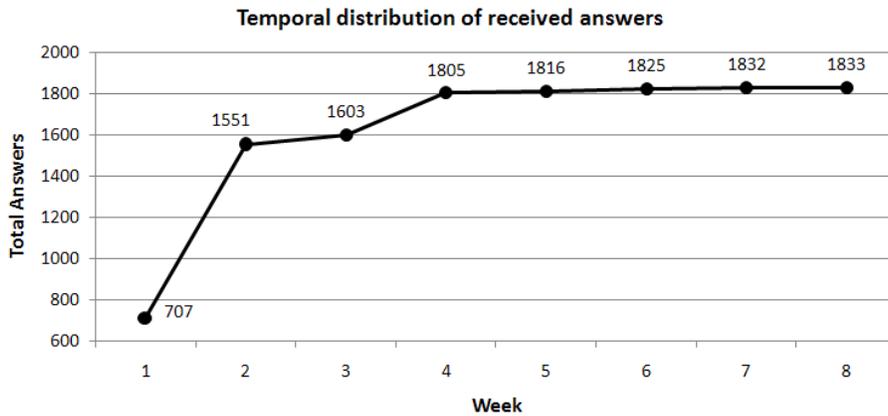


Fig. 2. Temporal distribution of received answers to the online survey.

In the third part of the questionnaire we asked users and administrator to express their opinion on the percentage of code developed by developers hired by firms. We also asked if they were aware of other kinds of support from firms, e.g. donating code, funding, marketing, logistic and infrastructural support, and to rank them by importance.

By analyzing the multiple choice answers we associated each project participated by a firm to one of the models proposed in Section 0. As we already discussed, a firm may contribute to an Open Source project according to more than one model, thus leading to blended participation approaches. However, we asked to rank the activities performed on behalf of a firm by significance, and were thus able to identify the prevailing model of participation for each project.

4 Results of the survey on Firms participation in OS projects

Administrators and users of 409 out of 1,302 projects (31%) have declared that one or more firms are somehow involved in their project. This confirms the leading role played by for-profit firms not only in commercial OS projects, but also in community OS projects, such as those hosted on *SourceForge.net*.

Table 3 reports some descriptive statistics on the characteristics of the projects participated by firms. On average, each project was participated by 2 firms, with more than 60% of the projects participated by only one company. In most cases (59.4%), firms have entered the project from its beginning.

Table 3. Descriptive statistics of projects participated by firms

Variable	No. obs.	Min	Max	Mean	St. Dev	p50	p75	p95
No. of firms per project	409	1	35	2.17	3.18	1	2	5
Years of involvement of firms within the project	404	0	12	3.16	2.48	3	5	8

Figure 3 summarizes the results on the prevailing model of firms' participation in OS projects. In 68% of the cases firms mainly support the project in non-development activities, such as testing, animating forums, writing documentations and providing financial and logistic support (support model). In 30% of the cases firms' main activity within the project is the contribution of code (coding model), whereas only in 9 cases (2%) firms coordinate the projects without performing other major activities (management model). This last result is in line with the spirit of *SourceForge.net*, which mainly hosts community-based projects [4, 19]. The governance of these projects is based on the concept of *do-ocracy* [29], i.e. decisions are made by the developers who contribute more actively to the project. Consequently, a firm can have a role in defining requirements and evolution strategies only if it actively contributes to the project.

Figure 4 shows the distribution of the proportion of code contributed on a voluntary basis in the projects where at least one firm was involved. It is interesting to note that in 26% of these projects the percentage of code contributed on a voluntary base is less than 10%. On average, when firms actively contribute to the codebase of a project, they contribute more than half of the code (57%). If the same percentage is computed for the code of all the projects in the sample (i.e., including the projects with no firms involved), approximately 20% of the code is contributed by hired developers.

As we noted in Section 0, firms may contribute to the code base of a project in a number of different ways. According to the results of the survey, the most common situations are that:

- They ask some of their employees to develop code for the project during their working time;
- They hire some developer already working in the community of the project and ask him/ her to develop specific functionalities;
- They release part of their proprietary code under an OS license and donate it to the project.

Actually, 45% of the projects with at least one firm involved had been founded by the leading firm, which released part of its proprietary code and published it on *SourceForge.net*.

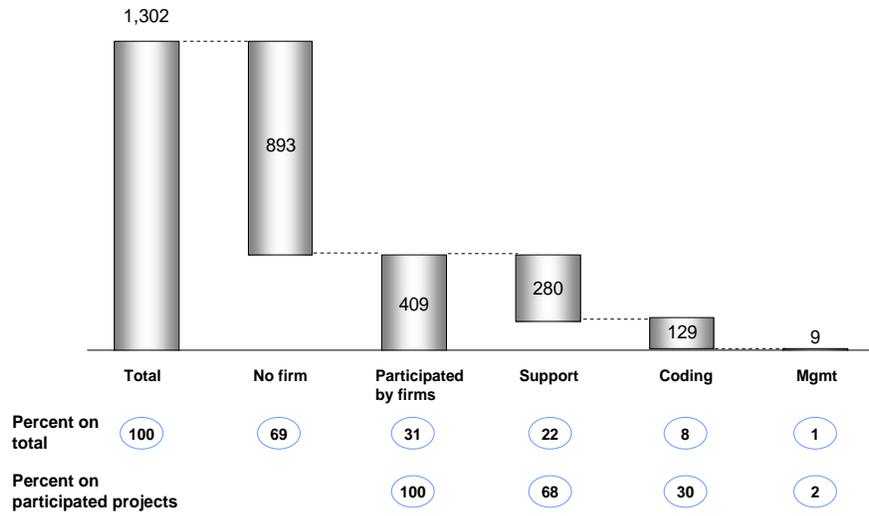


Fig. 3. Prevailing models of participation of firms in Open Source projects

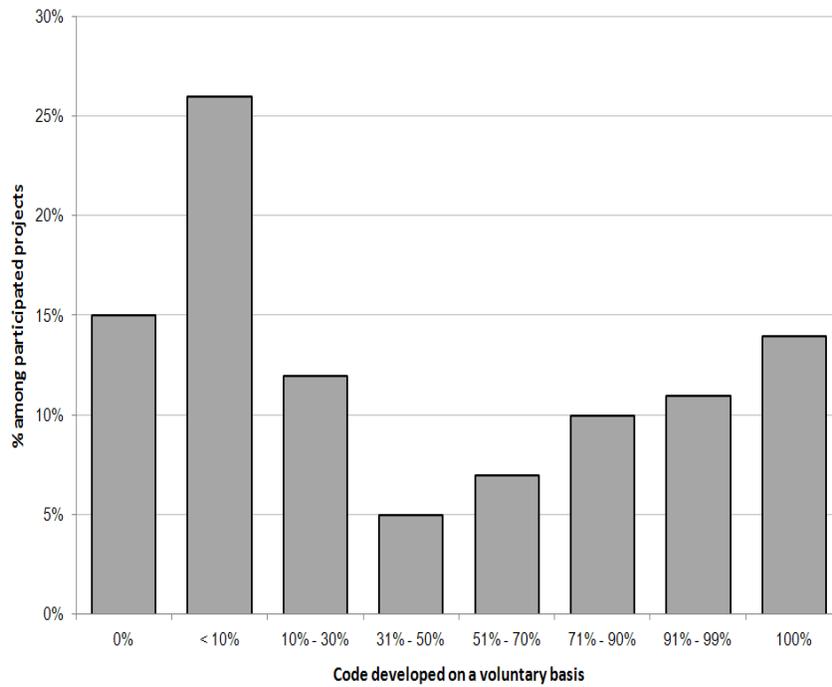


Fig. 4. Proportion of code developed on a voluntary basis, among projects with firms involved

5 Conclusion and future work

Our survey shows that firms have a significant role even in those OS projects that are commonly believed to be mostly based on the work of volunteers, such as those hosted on *SourceForge.net*. Notwithstanding that these projects are usually based on informal governance structures, our results witness that firms support these projects by contributing code, but also – and mostly – by performing non-development activities.

The issue of commercial firms' involvement in the OS arena is increasingly attracting the attention of scholars as it witnesses the metamorphosis of OS from a strongly ideological oriented movement into a more mainstream and commercially viable form [9, p. 587]. However, up to now, contributions addressing firms' participation in OS community projects have been mainly based on descriptive and anecdotal evidence [7, 10, 22], while the firms' role within these projects has been poorly investigated.

This paper contributes to the current debate on the topic by providing systematic quantitative evidence on the fact that commercial firms do actively contribute to OS community projects, even if with different intensities and undertaking different activities. We distinguished among a coding, a support, and a management model of participation and, for each of them, we found results that have connections with the overall research in the area.

Specifically, our findings on the coding model are in line with the stream of literature [14, 15] highlighting that firms do not only act as takers, by simply adapting the OS code downloaded from the Internet to their customers' needs, but they are also givers, thus directly undertaking code-writing activities. At the same time, our data show that companies support the OS communities in several ways (support model) and, in some cases, they even sponsor and set up OS projects, as underlined, for instance, by O'Mahony and West [32]. Finally, our results on the management model contribute to the growing body of research exploring governance mechanisms in the OS framework [6]. We found that only in few cases companies directly coordinate the projects in which they are involved. This seems to provide empirical support to the stream of literature observing that the adoption of formal and hierarchical forms of governance by companies is likely to crowd out the intrinsic motivations of developers [6].

This paper is only a first step towards a comprehensive understanding of the many issues related to commercial firms' involvement in OS projects. First, the empirical results should be verified also in other online repositories apart of *SourceForge.net*. Second, more research effort should be directed to investigate how companies' participation affects the evolution of the OS projects: as a further development of this research, we are planning to study the impact of firms' presence on projects' quality, size, and success.

Acknowledgments

We would like to thank Dario Lorenzi and Sara Contini for helping us to conduct the survey.

References

1. A. Bonaccorsi, and C. Rossi. "Comparing Motivations of Individual Programmers and Firms to Take Part in the Open Source Movement: From Community to Business, Knowledge", *Technology and Policy*, vol. 18, no. 4, pp. 40-64, 2006.
2. A. Capiluppi, and M. Michlmayr. "From the Cathedral to the Bazaar: an empirical study of the lifecycle of volunteer community projects", in "Open Source Development, Adoption and Innovation", Springer Boston, pp. 31-44, 2007.
3. E. Capra, and A.I. Wasserman. "A framework for evaluating managerial styles in Open Source projects", *Proceedings of Open Source Systems Conference*, pp. 1-14, 2008.
4. E. Capra. "Software design quality and development effort: an empirical study on the role of governance in Open Source projects", Ph.D. Dissertation, Dept. of Electronics and Information, Politecnico di Milano, Italy, 2008.
5. L. Dahlander, and M. W. Wallin. "A man on the inside: unlocking communities as complementary assets", *Research Policy*, vol. 35, no. 8, pp. 1243-1259, 2006.
6. L. Dahlander, L. Frederiksen, and F. Rullani, "Online communities and open innovation: governance and symbolic value creation", *Industry and Innovation*, vol. 15, no. 2, pp. 115-123, 2007.
7. L. Dahlander, M. Magnusson, "How do firms make use of Open Source communities?", *Long Range Planning*, vol. 41, pp. 629-649, 2008.
8. P. B. De Laat. "Governance of open source software: state of the art", *Journal of Management and Governance*, vol. 11, no. 2, pp. 165-177, 2007.
9. B. Fitzgerald. "The transformation of Open Source software", *MIS Quarterly*, vol. 30, no. 2, pp. 587-598, 2006.
10. K. Fogel. *Producing Open Source software*, O'Reilly, Sebastopol (CA), 2006.
11. R. Goldman, and R.P. Gabriel. "Innovation happens elsewhere: Open Source as business strategy", Morgan Kauffmann, April, 2005.
12. G. Goth. "Open Source business models: ready for prime time", *IEEE Software*, vol. 22, no. 3, pp. 99-100, Nov/Dec 2005.
13. A. Hars, "Working for Free? Motivations for Participating in Open-Source Projects", *International Journal of Electronic Commerce*, vol. 6, no. 3, pp. 25-39, 2002.
14. R. E. Hawkins. "The economics of open source software for a competitive Firm. Why give it away for free?", *NETNOMICS*, vol. 6, no. 2, pp. 103-117, 2004.
15. J. Henkel. "Selective revealing in open innovation processes: The case of embedded Linux", *Research Policy*, vol. 35, pp. 953-969, 2006.

16. G. Hertel, S. Niedner, and S. Herrmann. "Motivation of software developers in Open Source projects: an Internet-based survey of contributors to the Linux kernel", *Research Policy*, vol. 32, no. 7, pp. 1159-1177, 2003.
17. J. Howison, and K. Crowston. "The perils and pitfalls of mining SourceForge", *Proceedings of International Workshop on Mining Software Repositories*, pp. 7-12, 2004.
18. S. Krishnamurthy. "A managerial overview of Open Source software", *Business Horizons*, vol. 46, no. 5, pp. 47-56, 2003.
19. S. Krishnamurthy. "Cave or Community?: An Empirical Examination of 100 Mature Open Source Projects". *First Monday*, Available at SSRN: <http://ssrn.com/abstract=667402>, 2002.
20. J. Lerner, and J. Tirole. "The Economics of Technology Sharing: Open Source and Beyond", *Journal of Economic Perspectives*, vol. 19, no. 2, pp. 99-120, 2005.
21. E. Raymond. "The Cathedral and the Bazaar". Springer, 1999.
22. D. Riehle. "The Economic Motivation of Open Source Software: Stakeholder Perspectives", *IEEE Computer*, vol. 40, no. 4, pp. 25-32, 2007.
23. J. A. Roberts, I. H. Hann, and S. A. Slaughter. "Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects", *Management Science*, vol. 52, no. 7, p. 984, 2006.
24. The Asterisk Project. Website <http://www.asterisk.org/about>.
25. The Jikes Project. Website, <http://jikes.sourceforge.net>.
26. The MONO Project. Website, http://mono-project.com/main_page.
27. The Open Source Initiative. Website, <http://www.opensource.org>.
28. R.M. Stallman. "The GNU manifesto". Oxford University Press, NY, USA, 1990.
29. A. Tille. "Custom Debian distributions". Website, <http://people.debian.org/~tille/debian-med/talks/paper-cdd/debian-cdd.en.pdf>, January 25th, 2007.
30. The Ximian Group. Website, <http://www.novell.com/linux/ximian.html>.
31. A.I. Wasserman, and E. Capra. "Evaluating software engineering processes in commercial and community Open Source projects", *Proc. Int'l Workshop Emerging Trends in FLOSS Research and Development*, 2007.
32. J. West, and S. O'Mahony. "Contrasting Community Building in Sponsored and Community Founded Open Source Projects", *Proceedings of 38th Annual Hawaii International Conference on System Sciences*, 2005.