

Company Contributions in the Linux Kernel Community: A look at collaboration, innovation, and organizational learning

Dawn M. Foster
November 5, 2014

Abstract

Open source projects, and the Linux kernel in particular, have large numbers of contributions from individuals being paid by organizations to contribute as part of their employment. However, very little research has been done to study the interaction and collaboration of companies and other organizations within these open source communities. This research proposal is designed to fill this gap by studying Linux kernel code contributions from individuals who are employed by these organizations using network analysis and interviews to identify the relationships between organizations. The collaboration between companies within the Linux kernel would also be considered in the context of organizational learning and innovation. The Linux kernel was selected as the study project because it is a very large, neutral project with contributions from many different companies.

Background and Literature Review

The motivation for individuals to contribute to open source projects has been studied extensively with a focus on looking at why so many people contribute without receiving any compensation (Hars and Ou, 2002; Herrmann, Hertel and Niedner, 2003; von Krogh, Haefliger, Spaeth, and Wallin, 2012). An early survey conducted in 2000 on motivation for contributions to the Linux kernel was designed to understand why developers participate “for free”. The study found that 20% of the developers were paid to contribute as part of their regular job, and another 23% were sometimes paid for their Linux work (Herrmann et al, 2003). Contributors who are paid to contribute as part of their employment at a company are often motivated to make more contributions to the project (Roberts et al 2006), have a stronger desire to get their code incorporated into the project, and some eventually become committers (Shah 2006). In a survey of open source

developers, Lakhani and Wolf (2005) found that 40% were being paid to contribute to open source projects and that some of the leading motivations for participating included creativity, intellectual stimulation and improving programming skills. This suggests that innovation and learning are important considerations for people working within open source projects.

Organizational learning, innovation and work are often thought of as separate and conflicting topics, but within communities, these three concepts are strongly interrelated as people learn from their peers and innovate in a collaborative way as a part of their work with others in the community. The learning and innovation that come out of the work being done within a community tends to be stronger and more applicable to the day to day work of participants than most formal corporate training programs, and much of the learning occurs from watching communication between more experienced, competent practitioners (Brown and Duguid, 1991). In open source communities, collaboration between companies within a project facilitates learning and innovation, and it represents a unique blend with aspects of both of the traditional innovation models, private and collective innovation (Grand, von Krogh, Leonard, and Swap 2004). The traditional approach is that innovation is either a private investment with the goal of increasing profit within a firm or a collective action where the output is a public good, but Eric von Hippel and Georg von Krogh (2003) referred to the blended approach to innovation within open source communities as the private-collective innovation model. Within the Linux kernel, this unique form of innovation happens in an environment where learning tends to be more ad hoc and informal, which is at odds with the planned and structured training that happens within traditional organizations (Ciborra and Andreu, 2001).

Studying organizational participation within open source communities is becoming increasingly important. Open source projects are seeing increased numbers of participants who are employed by corporations and non-profit organizations to contribute to open source projects (Jensen and Scacchi, 2007; Roberts, Hann, and Slaughter, 2006). For example, most contributions to the Linux kernel are made by developers who are paid by organizations, and in the most recent Linux kernel report, only 13.6% of the contributions to the Linux kernel are made by unpaid developers (Corbet, Kroah-Hartman, and McPherson, 2013). The number of unpaid developers is down from 18.9% in 2010 (Corbet, Kroah-Hartman, and McPherson, 2010).

Research Questions

Despite the increased participation from organizations, there has been very little research into the topic of the complexities introduced by company involvement, aside from motivation to participate in open source projects (Iivari 2011). Conaldi and Lomi (2013) did an interesting network analysis of individual contributors linked to problems represented by software bugs that provided some of the inspiration for this proposal. However, rather than looking only at networks of individual participants, one element of this proposed research would investigate networks of participants employed by companies linked to code contributions to better understand the relationships and collaboration between participants at different companies.

This research would also look at company participation through two related theoretical frameworks, organizational learning and innovation, to increase the relevance beyond those interested in open source projects and into the broader business audience. In particular, the private-collective innovation model (von Hippel and von Krogh, 2003)

would be used to investigate how innovation occurs within the Linux kernel and understand the benefit to companies for sharing their innovations with the broader community. Brown and Duguid (1991) suggest that within communities of practice, people learn by being on the periphery and observing behavior of experienced, competent practitioners. This research would study how informal, observational organizational learning occurs in a similar manner within the Linux kernel community.

The Linux kernel would be used as the primary study project for several reasons. First, the Linux kernel is a very large open source project; the current release contains almost 17 million lines of code (Corbet et al 2013). Second, the majority of participants are employed by companies who sponsor their work on the project. Since 2005, almost 10,000 developers representing over 1000 companies have contributed to the project, and the company affiliation for these individuals is well documented (Corbet et al 2013). Third, the Linux kernel is a neutral open source project where no one company has a controlling influence, so it is an ideal environment to study cross-company collaboration, innovation and learning in an open source project. Fourth, while this author is not a kernel developer, she has extensive expertise working with Linux and open source projects and finds the Linux kernel to be an interesting project for further study.

Methodology

One methodology for this study would be an organizational network analysis to identify the relationships between organizations by analyzing Linux kernel code contributions from individuals who are employed by these organizations. Research by Herrmann et al (2003) found that some Linux kernel development is done in spontaneous teams that form around work on Linux subsystems, so this proposed research would also include

network analysis of interactions between organizations as it relates to teams working together on subsystems. Company contributions are quantified every year using the “gitdm” tool for the Linux kernel report, and the full data are also made available in a git repository (Corbet et al 2013), which would be used as an input into this proposed research.

In addition to the quantitative data from online sources, a longitudinal survey would be conducted as a series of interviews over one and a half to two years to collect qualitative data that supplements and adds context for the online data sources. One option is to focus interviews on several of the top companies contributing to the Linux kernel and do in-depth interviews with both individual contributors and management. Individual contributors to the Linux kernel would be interviewed to find out who they say they collaborate with the most often, which would be compared to the network analysis of the git data described above. The interviews would also include questions about innovation and organizational learning topics. The initial interviews would be longer and more comprehensive with shorter update interviews every six months over a period of one and a half to two years. These shorter follow-up interviews would focus on who they collaborate with to look at how collaboration with people working for other companies changes or stays the same over time. Interviews with management would follow a similar cadence, but they would be focused more on organizational learning, innovation and other benefits to the company for contributions to the Linux kernel.

The University of Greenwich was selected primarily because of the work and expertise within the Centre for Business Network Analysis. The research project listed on the website, “Re-defining the Space for Companies-Communities Interaction” by Dr.

Riccardo De Vita and Dr. Guido Conaldi along with Dr. Guido Conaldi's previous research on open source projects (University of Greenwich 2013) are particularly interesting and relevant to this research proposal. The current plan is for this research to be self-funded by the author.

References

- Brown J S and Duguid P 1991, 'Organizational Learning and Communities-of-Practice: Toward a Unified View of Working, Learning, and Innovating', *Organization Science*, 2, 1, pp. 40-57.
- Ciborra C and Andreu R 2001, 'Sharing knowledge across boundaries', *Journal of Information Technology*, 16, pp. 73-81.
- Conaldi, G and Lomi, A 2013, 'The dual network structure of organizational problem solving: A case study on open source software development', *Social Networks*, 35, 2, pp. 237-250.
- Corbet, J, Kroah-Hartman, G, and McPherson, A 2010, *Linux Kernel Development: How Fast It is Going, Who is Doing It, What They are Doing, and Who is Sponsoring It*, The Linux Foundation, viewed 15 September 2014, https://www.linuxfoundation.org/sites/main/files/lf_linux_kernel_development_2010.pdf
- Corbet, J, Kroah-Hartman, G, and McPherson, A 2013, *Linux Kernel Development: How Fast It is Going, Who is Doing It, What They are Doing, and Who is Sponsoring It*, The Linux Foundation, viewed 15 September 2014, <https://www.linuxfoundation.org/publications/linux-foundation/who-writes-linux-2013>
- Grand s, von Krogh G, Leonard D, and Swap W 2004, 'Resource Allocation Beyond Firm Boundaries: A Multi-Level Model for Open Source Innovation', *Long Range Planning*, 37, pp. 591-610.
- Hars, A and Ou, S 2002, 'Working for Free? Motivations for Participating in Open-Source Projects', *International Journal of Electronic Commerce*, 6, 3, pp. 25-39.
- Herrmann, S, Hertel, G. & Niedner, S 2003, 'Motivation of software developers in open source projects: an Internet-based survey of contributors to the Linux kernel', *Research Policy*, 32, 7, pp. 1159-1177.
- Iivari, N 2011, 'Participatory design in OSS development: interpretive case studies in company and community OSS development contexts', *Behaviour & Information Technology*, 30, 3, pp. 309-323.
- Jensen, C and Scacchi, W 2007 'Role migration and advancement processes in OSSD projects: A comparative case study', *Proceedings of the 29th international conference on Software Engineering*, ICSE, Minneapolis, pp. 364-374.
- Lakhani, K R and Wolf, R G 2005 'Why Hackers Do What They Do: Understanding Motivation and Effort in Free/Open Source Software Projects' in *Perspectives on Free and Open Source Software*, eds. J Feller, B Fitzgerald, S Hissam, and K R Lakhani, MIT Press, Cambridge, MA, pp. 3-22.

- Roberts, J, Hann, I, and Slaughter, S 2006, 'Understanding the Motivations, Participation, and Performance of Open Source Software Developers: A Longitudinal Study of the Apache Projects', *Management Science*, 52, 7, pp. 984-999.
- Shah, S 2006, 'Motivation, Governance, and the Viability of Hybrid Forms in Open Source Software Development', *Management Science*, 52, 7, pp. 1000-1014.
- University of Greenwich 2013, *Centre for Business Network Analysis: Research Projects*, University of Greenwich, viewed September 27, 2014, <http://www2.gre.ac.uk/about/faculty/business/research/centres/cbna/research>
- von Hippel E and von Krogh G 2003, 'Open Source Software and the "Private-Collective" Innovation Model: Issues for Organization Science', *Organization Science*, 14, 2, pp. 209-223.
- von Krogh, G, Haefliger, S, Spaeth, S, & Wallin, M 2012, 'Carrots And Rainbows: Motivation And Social Practice In Open Source Software Development', *MIS Quarterly*, 36, 2, pp. 649-676.