Distributed Leadership in OSS
Nora McDonald
Drexel University
2100 Delancey Pl
+16102569632
norakmcdonald@gmail.com

ABSTRACT
Open-source software (OSS) is software whose source code is available to view, change, and distribute without cost, and is typically developed in a collaborative manner that has captured the imagination of those who view the web as enabling more “democratic” models of governance. Researchers have, for years, debated the social structure of OSS projects – in particular, the extent to which they represent decentralized forms of organization. Many have argued that the significant concentration of code development responsibility raises doubts about whether the level of power-sharing truly qualifies as “distributed” in the way early observers predicted. This research will investigate how changes in the technology that supports these projects – specifically the greater visibility that characterizes the GitHub workspace may lead to a more broadly and quantifiably distributed leadership. Over the course of several studies employing several methodologies, it will examine leadership in OSS projects when visibility is a feature of the workspace.

Categories and Subject Descriptors
H.5.3 [Information Interfaces and Presentation]: Group and Organization Interfaces --- collaborative computing; computer-supported cooperative work; theory and models; web-based interaction.

General Terms
Management, Human Factors

Keywords
Open-source software; distributed leadership; social computing

closer to an “eight-twenty rule,” in which a relatively small number of participants shoulder most of the weight. Some OSS projects, like Linus Torvalds’ well-known Linux project, are depicted as something closer to a “benevolent dictatorship” – permissive but fundamentally centralized – potentially establishing a pattern widely adopted by other OSS communities that followed. But as infrastructures that support and house OSS development become more transparent, flexible, and informal, new opportunities for more distributed forms of leadership emerge. That evolution in structure, and the potential changes in behavior that may result, invite new ways of measuring leadership in these communities which can take account of new patterns of distributed interactions.

One such infrastructure is, arguably, GitHub, a code-hosting repository that uses Git version control system and makes the level of participation on OSS projects readily visible. With the introduction of GitHub, the level of participation on OSS platforms is also broader. Perspectives on GitHub’s more transparent and collaborative interface, and its potential to transform collaboration, were first introduced to the literature by Dabbish, Stuart, Tsay and Herbsleb [1]. The focus of their research has been on individual user processes and motivations, rather than project-level participation or success, and they have demonstrated the relationship between transparency and collaboration. The implications of transparency for leadership structure are equally important. Transparency not only enables participants to co-create work-product; it also supports development of new leadership behaviors such as monitoring, planning, and ad hoc coordination – leadership behaviors that are distinct from code and talk, and not necessarily measured by it.

2. RESEARCH GOALS
Our current understanding of leadership in the context of OSS communities is limited in at least two ways. First, leadership theories tend to conceive of leadership as emanating from a notional top, or as a set of qualities possessed by a single or a set of individuals – not something that is shared collectively. In a virtual “community” defined entirely by shared focus on tasks and goal-related outcomes, with no explicit leadership structure, qualities traditionally associated with leadership (e.g., charisma, influence, persuasion) are not present or take different forms. The second limitation flows naturally from this first – the inevitable tendency to measure what we can readily see. Thus, the metrics traditionally used to measure leadership in OSS communities are actually better suited to measuring productivity, which is clearly a different idea, than they are to measuring leadership. When there is no established “oversight” or known leadership structure, one is tempted to use as a proxy for influence the available measures of activity. Leadership, however, is a qualitative idea, and quantity of behavior does not necessarily do it justice.
In order to understand how platform visibility can influence patterns of leadership, it is necessary to employ a broader set of metrics that can produce a more nuanced understanding of leadership and community member interactions. The goal of this dissertation research is to answer the question, *How does technology support and transform leadership in OSS?* by focusing specifically on leadership in communities on the GitHub platform, using metrics and methods that allow for a broader conception of leadership than has previously been employed.

3. THEORETICAL FRAMEWORK

Distributed Leadership (DL) [2] offers a theoretical lens through which to think about the dispersion of leadership. DL is a theory that suggests leadership is shared among members of an organization; it argues that members of a community who do not occupy any formal leadership role can and do perform leadership activities or behaviors. It also attempts to understand how, why, and when participants demonstrate leadership behaviors, and the relationship between different forms of distributed leadership and community effectiveness. This more fluid, ad hoc conception of leadership lends itself well to technology-based *sui generis* communities that have no established leadership structure, although it has not traditionally been propagated or analyzed in a technology context.

Distributed Cognition (DCog) is another theory that may help fill in the gaps left by DL. DCog advances the view that knowledge resides not only with individuals, but also with the social and physical environment. The theory, along with actor-network theory (ANT), considers non-human objects to perform social roles, setting the stage for assignment of leadership behaviors to both people and technology objects.

4. RESEARCH APPROACHES

GitHub, a successful code-hosting repository, provides a fertile ground for an investigation of technology platform and leadership because its features expose the activities of community members to one another. My research encompasses qualitative interviewing, conversation analysis, observation, study of electronic archival data, and finally, a series of quantitative surveys.

First, exploratory research, including conversation analysis of pull request communications (Figure 1), walk-throughs, analysis of IRC channels, was conducted to support hypothesis development and study design. Next, 28 semi-structured interviews were conducted with respondents from eight different communities. In addition to asking respondents to describe how their roles and responsibilities had evolved throughout the course of their participation, interviews explored how the project had evolved on the platform; how the respondent/community uses its features and coordinates work; and how participants measure project “success.” Preliminary findings from this research were published [3]. Based on these analyses, I developed a set of metrics for use in making systematic project comparisons between projects on GitHub. A survey instrument for measuring success was also fielded to contributors on five OSS projects.

Results of an early pilot study focused on exploring ways to operationalize distributed leadership behaviors as an increase in awareness and monitoring of others using GitHub features. An example from this early study is shown in Figure 2, in which participants were asked to consider their awareness of what others are doing on GitHub versus other code hosting platforms.

5. CONTRIBUTION

Findings from this dissertation will demonstrate both the relevance of transparency to the emergence of leadership behaviors and potentially broaden our conception of how leadership is appropriately defined in virtual collaborative workspaces.

6. REFERENCES

