THE EVOLUTION OF LEADERSHIP STRUCTURES IN ONLINE COMMUNITIES: A SOCIAL NETWORK PERSPECTIVE

Mohammed Aael Harysi  
*University of Warwick*, phd16mh@mail.wbs.ac.uk

Bogdan Negoita  
*HEC Montréal*, bogdan.negoita@hec.ca

Joe Nandhakumar  
*University of Warwick*, joe.nandhakumar@wbs.ac.uk

Follow this and additional works at: [https://aisel.aisnet.org/ecis2019_rp](https://aisel.aisnet.org/ecis2019_rp)

Recommended Citation
[https://aisel.aisnet.org/ecis2019_rp/89](https://aisel.aisnet.org/ecis2019_rp/89)

This material is brought to you by the ECIS 2019 Proceedings at AIS Electronic Library (AISeL). It has been accepted for inclusion in Research Papers by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
THE EVOLUTION OF LEADERSHIP STRUCTURES IN ONLINE COMMUNITIES: A SOCIAL NETWORK PERSPECTIVE

Research paper

Harysi, Mohammed, The University of Warwick, Coventry, United Kingdom, phd16mh@mail.wbs.ac.uk
Negoita, Bogdan, HEC Montréal, Montréal, Canada, bogdan.negoita@hec.ca
Nandhakumar, Joe, The University of Warwick, Coventry, United Kingdom, joe.nandhakumar@wbs.ac.uk

Abstract

Online communities are increasingly seen as new forms of organizing. However, we have limited understanding of how leadership structures in an online community emerge. Prior literature focuses on leadership that is relatively concentrated in the hands of select individuals, once the online community has already been established. This paper therefore seeks to explore how leadership evolves throughout the history of an online community, from an embryonic stage, through the emergence stage to the establish stage. In the context of an online community built around a GitHub-hosted project called GitPoint, we carried out a content analysis of interactions conducted across social networks, including Twitter and Gitter, between its individual members. Based on this analysis, the paper offers insights into the emergence of leadership structures in an online community and makes key contributions to the literature on leadership in online communities.

Keywords: Online Communities, Online Leadership, Social Networks.

People enjoy interaction on the Internet and the feeling of belonging to a group that does something interesting: that’s how some software projects are born. — Linus Torvalds

1 Introduction

Online communities have rapidly emerged as a new organisational form (Resnick, Konstan, Chen and Kraut, 2012; Puranam, Alexy and Reitzig, 2014; Johnson, Safadi and Faraj, 2015). Because online communities represent new “ways of collaborating and organising knowledge work across boundaries”, their structures have been recognised as more “reconfigurable, boundary-less and flexible” (Kudaravalli and Faraj, 2008, p. 707) compared to traditional forms of organising, such as teams and groups (Wageman, 1995). Their shared goals range from software development (Singh and Phelps, 2013; Lindberg, Berente, Gaskin and Lyttinen, 2016) and healthcare support (Cobb, Graham and Abrams, 2010) to sports commentary (Ren, Kraut and Kiesler, 2007). Indeed, the group structure has been conceptualised as an emergent phenomenon in which the ongoing interactions between individuals shape the structural form that the group takes over time. Barley (1990, p. 6) noted that structure is “a global pattern that emerges from the relationships that exist among all members of a collective”. Similarly, in online communities, activities, such as communication, collaboration and leadership, become
structural patterns over time, and they manifest in networks of relationships and behavioural interactions that vary at different points in the team life cycle (DeRue, Nahrgang and Ashford, 2015). We argue that the leadership structure in online communities emerges over time and through a set of behaviours and ongoing interactions (Johnson et al., 2015). We extend this argument by explaining how individual roles (i.e., the function assumed or part played by an individual in a particular situation), behaviours (i.e., what the individual does in a particular situation), and expressions of leadership (i.e., the way in which an individual conveys leadership in a particular situation) engender and constitute the leadership structure throughout the online community lifecycle. The aggregation of these dyadic interactions and relationships between community members creates a group-level leadership structure that varies in two important configurations: centralised and decentralised (distributed). Prior research suggested that a centralised configuration of leadership emerges when one member becomes the “only person establishing goals and motivating members to accomplish those goals” (DeRue et al., 2015, p. 1194). In contrast, a decentralised leadership structure emerges when many or all group members perform the leadership’s diffusion of responsibilities (Huffaker, 2010; Bolden, 2011).

Indeed, several previous studies have enriched our understanding of the emergence of online leadership in mature online communities (O’mahony and Ferraro, 2007; Faraj, Jarvenpaa and Majchrzak, 2011; Collier and Kraut, 2012; DeRue et al., 2015; Johnson et al., 2015). However, much less is known about how leadership structures evolve before, during and after the existence of these communities. Further, these studies do not reveal the antecedents that led to the emergence of these structures.

Such limitations of the extant literature have many implications. The conclusions based on the partially structural perspective used in prior studies do not reveal the full dynamics of an online community, such as the interactions that occur during its birth and emergence phase. Therefore, although considerable progress has been made in understanding leadership in the context of an online community, the focus of the extant literature has been on leadership that is concentrated in the hands of select individuals when the online community has already been established. Hence, the aim of this paper is to address the following research question: How do leadership structures become established in online communities?

Drawing deductively from extant literature on online leadership and more specifically on leadership in online communities, we built a theoretical argument that the emergence and evolution of leadership structures, composed of roles, behaviours, and expressions of leadership (Bass and Bass, 2008), is rooted in interactions between select individuals across social networks. Empirically, we carried out an extensive content analysis of digital trace data (Berente, Seidel and Safadi, 2018) of interactions between members of an online community (i.e., GitPoint users) across social networks (Twitter and Gitter), as their online community was coming into existence. Leveraging the research lens of distributed leadership (Yukl, 1999; Denis, Langley and Sergi, 2012), we sought to understand the emergence of leadership structures in the context of their online community.

2 Literature Review

2.1 Distributed Leadership

Online leadership differs from traditional leadership because of the lack of traditional structures (e.g., hierarchies) in online communities (Johnson et al., 2015). The analysis of online leadership has benefited from the previous works on organisational leadership. We argue that a deeper understanding of the emergence and evolution of leadership in online communities requires reviewing and synthesising the growing body of organisational research and theorising that leadership is a distributed collective phenomenon. Much of the previous literature on leadership has focused on describing the individual or heroic leader (Rost, 1991; Bono and Ilies, 2006), which has led to only a partial perspective on leadership. However, a review of the theories and future directions in the leadership literature revealed the emergence of research that attempts to reach beyond the “heroic” view of unitary leadership (Avolio, Walumbwa and
Weber, 2009). Thus, the fundamental view expressed in the extant literature is that leadership is neither the responsibility nor the monopoly of a single individual. Instead, it is a collective phenomenon that is distributed or shared among different people. Furthermore, it is potentially fluid and constructed in interaction (Yukl, 1999; Gronn, 2002; Denis et al., 2012). Therefore, the leadership’s influence is considered distributed across multiple individuals rather than based on the traits, behaviour or influence of a particular individual (Carson, Tesluk and Marrone, 2007). In the “plural” perspective, leadership is conceived as a collective concept involving multiple individuals who assume and relinquish leadership responsibilities in both formal and informal interactions over time.

Denis, Langley and Sergi (2012) reviewed the forms of leadership that in one way or other implied plurality or “leadership in the plural”. They identified four streams of scholarship on plural leadership. The first stream, “sharing leadership in teams”, emphasises the need for members to co-lead each other in achieving the group or organisational goals or both. Shared leadership does not require that one individual performs all leadership functions. Instead, it is a shared process that features “a set of people who can collectively perform them. Some leadership functions may be shared by several members of a group (...) and a particular leadership function may be performed by different people at different times” (Yukl, 1999, pp. 292–293). The second stream, “pooling leadership capacities at the top to direct others”, focuses on situations where two, three or more people work together as co-leaders of others who are outside the team (Denis, Lamothe and Langley, 2001). The third stream, “spreading leadership within and across levels over time”, considers leadership as it is dispersed up, down and across the organisational hierarchy (Spillane and Diamond, 2007). It considers that leadership may be passed “from one hierarchical level to another over time as well as across intra-organisational and inter-organisational boundaries” (Denis et al., 2012, p. 2013). This stream is associated with the term “distributed leadership”, which was developed extensively by researchers interested in inter-organisation (Huxham and Vangen, 2000).

Distributed leadership is not done by “an individual to others, or a set of individual actions through which people contribute to a group or organisation”; instead, it is a group activity that works “through and within relationships, rather than individual action” (Bennett, Wise, Woods and Harvey, 2003, p. 3). In this relational context of distributed leadership, Bass and Bass (2008, p. 17) presented leadership as the initiation of a structure, whereby leadership is seen “not as passive occupancy of position or an acquisition of a role but as of originating and maintaining the role structure– the pattern of role relationships”. Finally, the fourth stream, “producing leadership through interaction”, views leadership as a human social construction that is enacted through rich connections and interdependencies of organisations and their members (Crevani, Lindgren and Packendorff, 2010). Thus, the concept of distributed leadership extends beyond a view of leadership as a “property of individuals and their behaviours” to the perception of leadership as “a collective phenomenon” that is distributed and constructed through fluid interactions between individuals (Ropo, Salovaara, Sauer and De Paoli, 2015, p. 122). Writing about distributed leadership in the context of online communities requires a fundamental understanding of the dispersion of leadership roles across communities and beyond their boundaries as well as how a variety of individuals relay leadership responsibilities over time to achieve desired outcomes (Denis et al., 2012).

The four streams described here differ in identifying the leader and the follower. The streams of “sharing leadership” and “producing leadership” focus on the idea that because group members explicitly or implicitly lead each other, the followers and leaders are essentially the same people. Conversely, the other two streams, in which leadership is pooled or distributed “tend to retain the notion that some are leaders and some are followers”, for example in the pooling stream, leadership is coalitional, whereby small numbers at the top share roles and jointly “exert influence over others through the consolidation of power relations” (Denis et al., 2012, p. 269). Similarly, the distributed leadership stream implies that different leaders contribute at different times, thereby constituting an emergent coalition.
2.2 Leadership in Online Communities

In previous research, online leaders were identified based on distinctive characteristics and strategies, such as those who have formal roles or high network centrality (Gary, 2010). In another definition, a leader is a participant who is “present, setting agendas, moderating interactions, and keeping the group on track with its goals” (Huffaker, 2010, p. 595). Johnson, Safadi and Faraj (2015, p. 166) defined “an online community leader as a participant recognised by other participants as influential in what the community does or how it does it”. Even if these definitions seem to have the same meaning, scholars have interpreted them differently. Individual perceptions of leadership vary due to the idea that leadership is “phenomenological process whereby individuals attribute leadership to certain others” in a social network (Mehra, Dixon, Brass and Robertson, 2006, p. 68). Thus, leadership’s definition, is not limited to the network position or the membership in the core. Previous studies acknowledged that use of straightforward and concise language (Johnson et al., 2015) and the believe in the usefulness of the community (Koh, Kim, Butler and Bock, 2007) are signs of leadership as well. Further, in production-based online communities, continuous participation and high engagement are linked to the strength of the bonds and ties developed in social practices. These practices reflect the communal activity of such communities (Ren et al., 2012). Bonds and ties can be built through community leaders because sociability and gregariousness are characteristics associated with leadership (Huffaker, 2010). The research findings on leadership in the context of online communities are mixed. For example, Ganley and Lampe (2009) found that high network centrality (i.e., brokering) is beneficial for newcomers to an online community, but less so for others with high ratings (i.e., very active users). Other studies of online discussions and online communities showed that there was no relation between brokering and overall participation (Toral, Rocío Martínez-Torres, Barrero and Cortés, 2009). In contrast, Koh et al. (2007) found that leaders in successful online communities had high network centrality and spent time motivating the participation of other members to build social identity within the group. Consistent with these findings, Fleming and Waguespack (2007), O’Mahony and Ferraro (2007) and Collier and Kraut (2012) showed that degree centrality (i.e., network position) predicted leadership. Moreover, (Johnson et al., 2015, p. 180) observed that leaders “are not just among the most central, but also post a large number of positive, concise posts with simple language familiar to other participants”.

In summary, in the existing literature, scholars studying leadership in online communities either focus on the structure of the network (indegree, outdegree and betweenness centrality) or on individual traits and characteristics (positive posts, higher contribution, experience, etc.) to understand who is a leader. They do so primarily in contexts where the online community is mature and well established. This points to a possible reason for the existence of this paradox – that leadership in online communities is shaped by structure or competence. The reason may rest with when in the lifetime of an online community studies have focused on leadership. So, a partially structural perspective may not reveal the full dynamics of an online community, such as interactions that occur during the birth phase of the online community (Faraj, von Krogh, Monteiro and Lakhani, 2016).

3 Methods

3.1 Research Design

In this paper, we investigate the formation of leadership structures in online communities by examining the dynamic properties of users’ interactions prior to the existence of the online community (GitPoint). We adopt a content analysis approach and focus on understanding the emergence of leadership structures throughout the history of GitPoint’s repository (Gioia, Corley and Hamilton, 2013; Williams and Shepherd, 2017). GitPoint (https://gitpoint.co) is a public repository on the GitHub platform, which was created on 26 March 2017. This repository was created by an amateur programmer who intended to build an open source application that enabled GitHub’s developers to view repositories, user information, review codes and manage
their projects through smartphones. GitPoint has 142 members and is self-managed in allocating tasks, fixing bugs and maintaining codes. Interestingly, after its creation, GitPoint’s repository was idle until the original owner shared it through Twitter. The original owner used Twitter to generate and screen ideas for the GitPoint project. Notably, the owner encouraged participation, information and opinions about GitPoint’s repository through Twitter.

In the present study, this repository was selected to investigate the emergence of leadership structures from distributed interactions across social networks (Twitter and Gitter) for three reasons. The first reason is that GitPoint is seen as a virtual space for social aggregation, where technology enables individuals to interact, manage, share and build ideas digitally. The second reason is that GitPoint’s repository arose without traditional organisational structures. For example, complex and interdependent tasks were accomplished without traditional leaders and subordinates. The third reason is that GitPoint complies with Kim’s (2000) proposal regarding the strategies that constitute successful and sustainable online communities: a clear vision, flexible spaces, individual roles, leadership and events.

### 3.2 Data Collection

The empirical aspect of this study required three phases of data collection, which relied primarily on digital trace data capturing and analysis. In the first phase, digital data were collected from GitHub (GitPoint’s repository) through its automated API functionality. The data included the profiles and exchanges of all users. GitPoint’s repository included 77 contributors, 65 users, 732 open and closed issues, 358 forks, 416 pull-requests and 594 commits. A contributor to GitHub contributes to changes to the repository’s original files. Users are members of the repository who participate, but their participation has not been approved for inclusion in the original files. Issues in GitHub are used to keep track of changes (e.g., tasks, enhancements and bugs). The issues file contains GitPoint’s member exchanges, such as issues that need to be addressed, authors, assignees and labels. Building on these exchanges, we considered that close attention to team effectiveness (i.e., the number of open/closed issues and the diversity of authors and assignees) would help us to identify the explanatory mechanisms underlying the emergence and evolution of the leadership’s structure. Forking was another important aspect of our data collection. A fork is a copy of a repository. Forking a repository gives anyone the ability to experiment with changes without affecting the original repository’s files. We considered that a higher number of forks was a sign of GitPoint’s effectiveness. Additionally, GitHub pull-requests allow users to alert each other about the changes that they have made to a project (repository). When a pull-request is sent, interested users can review and discuss the request. However, only users who have the authority can accept and add changes to the original files of a repository. The main reason for including the pull-request files is that there are leadership functions that GitPoint’s users execute once a pull-request is opened. These leadership functions and interpersonal perceptions of others vary in terms of the presence and absence of ties or weighted value ties. The commits file represents any individual’s changes that have been added to the original files of GitPoint’s repository. This file contains commits messages, which are brief clarifications of the changes that have been made. Indeed, the Twitter handles of GitPoint users were an important attribute of the data in the first phase of the data collection. These handles constituted the corpus of the Twitter data that were related to the production-based online community (GitPoint).

In the second phase, having identified Twitter accounts for the same individuals who participated in GitPoint’s repository, we were able to collect their microblog posts, which are public tweets that anyone can trace, using Twitter’s premium search API (i.e., full-archive endpoint). Twitter’s premium API provided not only complete and instant access to tweets since March 2006 but also the functionality to search historical data (tweets) in varying periods of time. Hence, we had access privileges to the full history of GitPoint members’ activities on the Twitter platform. Using the historical research privileges, we accessed data, such as #hashtags, @mentions, tweets, retweets and links, before the emergence stage and the established stage of GitPoint’s repository. To clarify, the interactions of GitPoint members on
Twitter were collected individually; each user had his or her own set of tweets. By integrating Twitter’s premium API configurations in a Python script, we traced all the GitPoint user accounts that we identified in the first phase of the data collection on GitHub. Although this dataset of individuals contained an extensive amount of data, we focused on tweets that were made before or during the emergence stage of the GitPoint repository. However, even after we sorted the Twitter data in a specific period, an unrelated set of data existed, which consisted of tweets that were not related to GitPoint’s repository. Consequently, in this broad database, we read the entire dataset of posts to identify the tweets, retweets, @mentions, links and #hashtags that were related to GitPoint (Vaast, Safadi, Lapointe and Negoita, 2017). We sought to isolate the unrelated data; hence, we proposed the following criteria: 1) exclude tweets and participation dated after the emergence stage; 2) data must be relatively relevant to the GitPoint repository (direct posts about GitPoint or about something that is indirectly related to GitPoint, such as #ReactNative, the framework used to build GitPoint); 3) users (Tweeters) must be GitPoint members or users (Tweeters) who directly Tweeted about #GitPoint. Through reading the dataset mentioned above, we identified Gitter as a communication platform that was used in addition to Twitter. Gitter is a chat and networking platform that helps to manage, grow and connect communities through messaging, content and discovery. Therefore, Gitter was the third phase of the data collection. The main reason for including Gitter was to triangulate the data points from different sources and to create as complete a story as possible regarding the emergence of leadership structures in the GitPoint repository. The data combined from multiple sources assisted in determining the variables of interest, including changes in the network structures, varying behaviours and the emergence of new roles (Williams and Shepherd, 2017, p. 274).

3.3 Content Analysis

In this study, we adopted a content analysis approach (Gioia et al., 2013). We conducted an extensive content analysis of the interactions in the online community of GitPoint users across the social networks of Twitter and Gitter before and during the emergence stage. In recent years, there has been an increasing amount of literature on social networks (e.g., microblogging). Numerous studies have attempted to explore the use and nature of the interactions on social networking sites (Boyd and Ellison, 2007; Vaast and Walsham, 2013; Leonardi, 2014; Williams and Shepherd, 2017), while others presented them as sources of data (Vaast et al., 2017). Currently, social networks are conceptualised as key locations for networking, sharing, news and marketing. These social network sites hold vast quantities of naturally occurring data, which provides a tremendous opportunity for researchers to enrich their theoretical repertoire. Our analysis is based on three foundations. First, “a majority of social network research has been quantitatively focused” (Williams and Shepherd, 2017, p. 269). However, although these quantitative techniques create value, they also have limitations. Second, the actual events in our proposed phenomenon cannot be determined only by quantitative means, such as measuring the correlation between variables (Zachariadis, Scott and Barrett, 2013). For instance, measuring the attitudes of potential leaders through quantitative analyses and correlations between variables could lead to neglecting the benefits of inductive research, which reveals the causal mechanisms of actual events. Hence, research that relies only on quantitative methods in generalising knowledge claims has been described as “unsatisfactory and problematic” (Zachariadis et al., 2013, p. 862). Third, the wide use of content analyses in information systems research motivated our use of that a qualitative methodology to understand the phenomenon of the emergence of leadership structures in the history of GitPoint’s repository by examining the distributed interactions across the social networks of Twitter and Gitter (Gioia et al., 2013; Vaast et al., 2017; Williams and Shepherd, 2017).

Our first reading of the datasets (Twitter, Gitter and GitHub) was aimed to provide a comprehensive understanding of how to identify the embryonic, emergence and established stages. We started by carefully observing how users’ activities, roles, behaviours and membership evolved over time. We alternated between qualitative and quantitative measurements; for instance, we traced how participation differed, how roles emerged and how
behaviours varied from one user to another among all three stages. The history of GitPoint was divided into the three stages (embryonic, emergence and established). The embryonic stage represents the stage before the existence of GitPoint (i.e., individuals’ interactions on Twitter), which we identified as the period from 11 July 2017 to 31 August 2017 based on the users’ activity on Twitter. Second, the emergence stage is the stage between the embryonic and the established stages. We considered the first two months after the embryonic stage (01 September 2017 to 31 October 2017) as the emergence stage of GitPoint for three reasons. The first reason was that our dataset (Twitter and Gitter) showed an increase in the number of participation (posts) at both the member and the stage levels. The second reason was that new roles emerged for members, and the third reason is that the membership was steadily increasing. The established stage represents the maturity of GitPoint’s repository, wherein the leadership structures were fully developed. The rationale for dividing the data into three successive periods was not based on a particular theory. Instead, it was a practical way to structure the flow of events in each period. Moreover, as structuring and sorting the data proceeded, we divided each stage (GitHub, Twitter and Gitter) into two categories of issues: social and technical. We then aimed to interpret with caution the potential patterns and instances of what we called instantiations of leadership: the roles, behaviours and expressions of leadership. Furthermore, these instantiations were divided into nodes using Nvivo (e.g., roles: “owner”, “developers”, “amplifiers” and “visionaries”). The rationale for focusing on these instantiations of leadership in each stage (i.e., embryonic, emergence and established) was to gain an insightful insights into how the leadership structures emerged and evolved over time (Gioia et al., 2013). When we had identified these instantiations and stage “histories” of GitPoint in GitHub, Twitter and Gitter, we conducted a new systemic data collection to gather information on the instantiations in leadership (i.e., the roles, behaviour and expressions of leadership) in each stage (Yin, 2009). This data collection process involved reviewing GitPoint’s members’ interaction at each stage and then sorting them according to our analyses. These instantiations were coded for social and technical issues throughout the three stages.

4 Findings: The Stages of Leadership

4.1 The Embryonic Stage

Prior to this stage, we collected the profiles of all users who participated in GitPoint. One important aspect of the collected data was the participants’ Twitter handles. Thus, we identified this stage based on the earlier information exchange reflected in the individual tweets in relation to what was to become the GitPoint project. We paid careful attention to the corpuses of these tweets to understand the type of roles that were emerging and the behaviours of users during their online interactions. Initially, as we noted earlier, this stage was divided into two categories: social issues and technical issues. For example, we attempted to investigate the use of social psychology (e.g., the use of social psychology to motivate contributions) in social issues (Beenen et al., 2004). We also attempted to investigate how GitPoint’s users interacted in terms of technical issues (e.g., suggestions for using Gitter as a communication platform for the group and opinions about extending GitPoint’s support for Android phones). The roles of “owner” and “developers” were the only roles that we were confident about before we started exploring the data. The “owner” role represents the original owner of the GitPoint and the “developers” role represents users who participated mainly in the technical issues category such as suggesting technical solutions or using certain tools e.g., @alejandronanze: @RolfKoenders @hdjirdeh Slack has a message limit rate. That sucks. Gitter integrates good with 3rd parties too. I say gritter +1. As the analyses progressed, we found an explicit announcement of GitPoint by the original owner via Twitter: “@hdjirdeh: Couldn’t find a @github iOS app that had everything I needed, so I built one GitPoint - made with @reactnative; and one welcoming everyone to participate, regardless of their experience”). Therefore, as expected, the original owner triggered the idea of GitPoint through Twitter.

Notably, these tweets that were initiated by the original owner led to spikes of activity, which gained momentum through retweets and tweets by others. In these spikes of activity, we
observed that the spikes for technical issues were higher than the spikes for social issues. The original owner played more than one role in the embryonic stage; for example, the original owner expressed his role as an owner and developer in the technical issues category. Motivating contributions through goal-setting and expressions of leadership were observed in the social issue category. Moreover, as the qualitative analyses proceeded, we observed that the first tweets about GitPoint (i.e., the announcement) gained momentum immediately after the original owner posted them. This momentum led us to discover a new role, which we named “amplifiers” (see Figure 1). Amplifiers were users who promoted GitPoint and brought others’ attention to participating in GitPoint by reformulating or retweeting the original owner’s tweets e.g., @hosmelq (RTWEET): @hdjirdeh: Couldn’t find a @github iOS app that had everything I needed, so I built one.

**THE EMBRYONIC STAGE (ROLES)**

![Diagram of roles](image)

*Figure 1. shows illustrative summary of the coded roles in the technical issues’ category, where the owner role was the most frequently coded role, followed by the developers, amplifiers and visionaries roles.*

An interesting finding was that most of the references that were coded for the “amplifiers” node (27 of 42) were not members of the GitPoint repository. Hence, another role emerged: the “visionaries”. The “visionaries” role represents users who participated in drawing the roadmap and helping to formulate the goals of the GitPoint repository e.g., @mac_hour: Any plans on supporting Github Enterprise, and an iPad app? in that order I would not mind paying for it if needed. It was clear that in addition to the original owner, the users (Tweeters) were involved in expressing visions and setting goals. All the visions that the original owner enacted, such as support for iPad, an Android version as well as support for the GitHub Enterprise, were suggestions made by GitPoint’s users. As we continued the analysis of the tweets in an “open-coding” fashion, we sought to identify the prevalent user behaviours and the ways in which leadership was exhibited in the early days of the GitPoint project. We named these new concepts “benevolence” and “liking”. Regarding benevolence, we searched for tweets and retweets that indicated kindness, goodness and goodwill and monitored their influence on all users. Regarding liking, we searched for tweets and retweets that showed agreement, enjoyment or satisfaction with others’ tweets. Benevolence and liking seemed to be the most frequent in exchanges related to the social issues in managing GitPoint. They were less frequent in discussions about the technical aspects of the project. Interestingly, the original owner was the only coded reference for benevolence, and he was the most frequently referenced user regarding benevolence and liking among all users (15 of 43 references in both social and technical issues). Although the network position of the original owner resulted from a high number of interactions (i.e., @hdjirdeh was the highest referenced node), we realised that at this stage, leadership emerged as a social process, not as a fixed attribute (e.g., high participation). We also noticed that the original owner’s personal characteristics and experiences informed a set of behaviours...
that motivated others to participate. We perceived that this set of behaviours was correlated with the structure of leadership because the original owner was the most referenced user in the “Roles” node. In summary, the findings for this stage were contrary to our earlier predictions. We assumed that leadership would be distributed at the birth of the online community. However, the data analysis showed that at this stage, the leadership was centralised in the original owner. On one hand, the original owner (@hdjirdeh) was the most frequently coded reference in both categories (i.e., social issues and technical issues) at this stage (see Figure 2).

Figure 2 shows the number of coding references per user.

On the other hand, the “developers” role was the second most frequently coded node, followed by the “amplifiers” and “visionaries” roles (see Figure 1). Most importantly, the configuration of the leadership was highly centralised in the original owner. Furthermore, we realised that the tensions among ideas, passion, time and social ambiguity ebbed and flowed (Faraj et al., 2011). One generative fact was that individual characteristics and experiences shaped individual contributors’ behaviours during these intense fluctuations. For instance, we found that the original owner engaged in targeted interactions with everyone based on their potential contributions to the GitPoint whether in organising, sharing or coding e.g., @hdjirdeh: Don’t care whether you have 0 experience in dev or are an expert in @reactnative, - everybody is always welcome to learn.

Further, in exploring the users’ profiles, we observed that individuals with unique characteristics and previous teamwork experience behaved more positively than others did. They showed more appreciation for other contributors’ efforts; they expressed benevolence at higher rates than the other participants did, and they accepted others’ opinions regardless of their experience.

4.2 The Emergence Stage

In the emergence stage, the total participation was 1,159 posts. We observed that this increase in participation was comparable to that in the embryonic stage (n = 350 posts). Moreover, the participation was not only at the group level; there was also an increase in participation at the individual level. The steady membership (i.e., the high enrolment of users) in GitPoint during the emergence stage was another aspect that we considered. We observed that 14 users participated through Twitter during the embryonic stage (July–August 2017) compared to 16 new memberships during the emergence stage (September–October 2017) and 13 new memberships during the established stage (November–March 2018). Therefore, the embryonic and emergence stages were the periods of the highest membership enrolment compared to the established stage. Furthermore, in addition to the four roles identified in the embryonic stage
(i.e., owner, developers, amplifiers, and visionaries), new roles emerged. The movements of existing users between the roles were other indicators of the emergence stage. As in the embryonic stage, we observed the key characteristics of individual participation in online activities that were related to the GitPoint project. Statistically, we first sought to determine whether there was any change in the network structure.

We found a significant shift in the membership of the online community, which we considered a sign that the leadership configuration had changed. For instance, the original owner was no longer the highest participant in terms of content generation (i.e., the total number of posts). Instead, user @machour was the highest participant in this stage. Moreover, the number of active users was higher than in the previous stage. Hence, the original owner no longer seemed to dominate the project’s ownership regarding the level of participation. This stage differed from the embryonic stage in the following key aspects. The “amplifiers” role that emerged in the embryonic stage faded away at this stage. Following the increase in the rate of participation, we realised that user participation was focused on technical issues (e.g., coding, new releases, etc.) rather than on the idea-sharing (e.g., promoting and encouraging participation in GitPoint) that occurred during the embryonic stage. A new role emerged from the change in the nature of participation (i.e., increased focus on coding). The findings of our analysis indicated that the “developers” role was the most frequently referenced node, and the “organisers” emerged as a new role in this stage. Further, the “developers” role was prevalent considering the frequency with which users shared their original code contributions, their reflections on the challenges encountered during the development process, and how others reviewed these technical contributions. Among the developers, certain individuals used their technical contributions to signal or broadcast their competencies to the wider online community membership. The “organisers” role represents the users whose participation did not necessarily involve coding solutions but was focused on organising, guiding and relating challenges e.g., @andrewda: What do y’all think about adding some detox tests? #575. Therefore, these organisers engaged in managerial tasks rather than coding. Notably, developers and organisers were the highest coded reference nodes compared with the owner node.

Compared with the embryonic stage, the owner role continued to exist, yet it reflected a diminished position. Indirectly, his Gitter handle (@HousseinJirdeh) appearing less and less frequently in other individuals’ communications in the emergence stage. This finding suggests that the leadership was distributed in the emergence stage. Regarding the social and technical issues addressed during the emergence stage, the focus of participation was primarily on technical issues, yet the frequency with which the social aspects related to the activities of the GitPoint community were discussed was similar to the activity during the embryonic stage. We attributed the higher prevalence of discussion on social issues in the embryonic stage to the fact that, regardless of the tasks and roles of the users, most of their communication concerned promoting the project, connecting everybody and encouraging participation. This emphasis led most of the users in the embryonic stage to become more social at some point. However, the empirical results showed that the warmth of social interactions in the emergence stage was more pronounced, which indicated bonding between GitPont’s members. Such interaction encouraged others to participate and try to solve issues regardless of their experience and knowledge. In summary, our analysis of the emergence stage revealed a change in the leadership structure after the early embryonic stage. We associated the less frequent appearance of the owner node and the higher diversity and density of technical and managerial competence (i.e., skilled solutions were put forward by everyone) as indicators of the decentralisation of the leadership structure in this stage. For instance, we observed that GitPoint’s members grouped themselves based on their interests and skills (e.g., @alejandroanonez worked for GitPoint’s design and style, @mac_hour reviewed issues and testing, and @hdjirdeh and @lex11 worked on GitPoint’s releases). Interestingly, we continued to observe the effects of the individuals’ characteristics and experiences on their behaviours. We noticed that users who were kind, helpful and friendly in their communication were not the only users who influenced others; users who had previous experience working in teams, as evidenced in the digital trace data related to their professional history, gained increased influence through maintaining and
focusing on the work, providing solutions and reviewing others’ posts. Finally, in comparing this stage with the embryonic stage, we found that certain individual participants took on situationally specific roles over time. For instance, someone that acted primarily as a developer changed this role to that of an organiser and vice versa. Thus, these findings indicate that the roles enacted in the context of an online community during the emergence stage may last only as long as they are needed, which supports findings in the extant literature regarding the dynamics and flow of membership (Faraj et al., 2011).

4.3 The Established Stage

In the embryonic and emergence stages, we observed a significant shift in the leadership structure from centralised to decentralised. Similarly, based on our analyses of these previous stages, we expected new emergent roles and the explicit use of authority in the established stage. However, the “owner” and the “developers” roles continued to exist in all three stages. Because the leadership structure was fully developed by the established stage, two new roles emerged, which we called “committers” and “reviewers”. The “committers” role represented users who had the authority to accept other participants’ work changes to be added to GitPoint’s original files e.g., @TheCodeTalker now that the code is working, let’s move to the build failure and the commit workflow Bare with me on this one, so that your next contributions will be smoother. The “reviewers” were users who were assigned to certain issues or to review pull-requests to review e.g., @Chinesefan: Good. Can you update it? Then I think there are no problems from my side. Thanks for your patient revisions. The reviewing criteria were based on the users’ knowledge, abilities and interests. Users with some authority (e.g., owners and committers) could assign themselves or others specific tasks. The owner role continued to fade through all stages. Initially, in the embryonic stage, the owner role was the highest coded node, but it diminished during the emergence stage (i.e., decentralised leadership), where it represented only the ownership of the repository (i.e., a user with all privileges). Technical issues were higher than social issues because the interactions in this stage were more code oriented, organised and directed than in the previous stage, most of the issues involved enhancing performance, fixing bugs or reviewing pull-requests. Indeed, the developers and committer roles were the most frequently coded in this stage. Interestingly, in the established stage, some users (newcomers) with high participation did not participate in the embryonic or the emergence stages. Moreover, the original owner was not the highest contributor, and the variance in contributions was not high in the incumbent users, that is, those who participated in the embryonic and emergence stages, compared with the newcomers who participated only in the established stage).

In summary, we found that GitPoint members’ interactions in the emergence stage had significant implications on the leadership emergence in the established stage. The norms of leadership and the enactment of roles (e.g., gaining the role of administrator) did not depend on psychological attributes (e.g., knowledge and skills). Indeed, in the established stage, the leadership roles were enacted based on the group’s structure in the embryonic stage. This finding led us to conclude that, situationally, roles and predictors could be applied only at some points in a team’s lifecycle of work. For instance, higher participation by newcomers at the established stage did not engender new leaders, except that newcomers were seen as active users because the norms and concept of the leadership had already been established through the group’s structure in the emergence stage. Finally, at the established stage, the leadership configuration has changed. At the previous stage (the emergence stage) leadership was distributed. It was handed over between some members from one role to another e.g., new leaders beside the owner have emerged such as developers and organisers. Even the same leaders who had been identified during the emergence stage continued to lead at the established stage, the leadership configuration has changed from distributed to pooled leadership. In the emergence stage, leadership was associated with more democracy, empowerment, and participation between the GitPoint’ members who mutually lead each other, while in the established stage the leadership was clearly more structurally embedded (leadership is between the members that we have identified as leaders in the embryonic stage).
5 Discussion of Conceptual Development

Over time and through social interaction, leadership structures emerge in online communities. In addressing our key research question of how leadership structures evolve in online communities, we investigated an established online community (GitPoint) and studied its members’ previous interactions on social networks before it emerged into a GitHub-hosted project. We analysed the data collected from nine months of GitPoint members’ interactions on Twitter, Gitter and GitHub. Our findings in the embryonic stage showed that the level of the members’ activity was affected by the strength of their interpersonal bonds. For example, the findings showed that the original owner engaged in targeted interactions with the group members regardless of their experience and expertise. The leadership behaviours of the original owner during the emergence stage led to spikes of activity that indicated the gaining of momentum through retweets or original tweets. For example, through an explicit announcement of GitPoint by the original owner via Twitter, the original owner invited others to participate regardless of their knowledge and skills.

This finding is important not only for the configuration of leadership (centralised in the hands of the original owner) but also for the engendering of new roles, such as the amplifiers. Without the activity of the amplifiers, GitPoint’s content would not have been circulated. The findings also suggested that those viewed as leaders in the embryonic stage were not only among the most central participants, but they also engaged in targeted interactions with each team member based on their potential contribution to the team task. Furthermore, the findings showed that the leaders’ ability to mobilise resources and generate a distributed leadership structure in the emergence stage depended mainly on the quality of contributions (i.e., competence) in the emergence stage and the quality of interactions that occurred in the embryonic stage (O’Mahony and Ferraro, 2007). Thus, our findings allowed us to evaluate the interactions that occurred throughout the tipping points of the GitPoint lifecycle. Moreover, the ongoing interactions in the emergence stage had significant implications for the emergence of leadership in the established stage. As summarized in Table 1, the findings showed that the leadership structure was centralised at the embryonic stage, with the owner playing a key role in articulating the value proposition of the software development project at the heart of the online community and relying on amplifiers to reach out across their own social network to advertise the incipient project and its core idea. During the emergence stage, leadership became increasingly distributed with the owner letting go of certain content and administration-related responsibilities to newcomers which contributed specialist know-how (technical and management) that was conjointly enacted to develop and administer the budding online community. Lastly, at the established stage, leadership structures seem to have become increasingly formalized, resembling leadership structures centred around core-periphery arguments (Collier and Kraut, 2012) previously observed and reported in extant literature.

<table>
<thead>
<tr>
<th>Embryonic stage (centralised leadership)</th>
<th>Emergence stage (distributed leadership)</th>
<th>Established stage (pooled leadership)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Twitter</strong></td>
<td><strong>Amplifiers</strong></td>
<td><strong>Owner</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Developers</strong></td>
<td><strong>Visionaries</strong></td>
</tr>
<tr>
<td><strong>Gitter</strong></td>
<td><strong>Organisers</strong></td>
<td><strong>Developers</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Owner</strong></td>
<td><strong>Visionaries</strong></td>
</tr>
<tr>
<td><strong>GitHub</strong></td>
<td></td>
<td><strong>Committers</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Developers</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Owner</strong></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Reviewers</strong></td>
</tr>
</tbody>
</table>

Table 1. Shows the roles at each stage, where the darker the colour the higher the number and the lighter the colour the lower the number of referenced roles.

Regarding the theoretical implications of our findings; in previous studies (Butler, Sproull, Kiesler and Kraut, 2005; O’Mahony and Ferraro, 2007; Huffaker, 2010; DeRue et al., 2015;
Johnson et al., 2015), leadership structures were examined when the online community was mature and well established. In contrast, our study was focused on the emergence and evolution of leadership structures at different points in GitPoint’s lifecycle. Our findings showed that the leadership structures emerged during the emergence stage (distributed leadership) and became stable during the established stage (pooled leadership). Moreover, our findings indicated that the evolution of leadership structures involved a power transition among the members (Aime, Humphrey, Derue and Paul, 2014). For example, the original owner was not the most active (i.e., referenced) member during the emergence stage; others gained power based on the tasks that the community required. The findings of this study have important implications for the understanding of complex social phenomena, in which individuals interact to co-produce collective behaviours and structures. Path dependency is important in understanding and accounting for the history of the underlying phenomenon (Faraj et al., 2016). Furthermore, the present study offers a platform for considering not only the importance of the path dependency of complex social phenomena but also the importance of the effects of group-level structures on the emergence of other group-level structures. For example, in our study, the leadership structure in the emergence stage influenced the evolution of leadership structures in the established stage, such as how incumbent users preserved roles as leaders and newcomers adopted subordinate roles as followers.

The findings of this study have implications for those who manage or are responsible for structuring teams in online communities. The findings showed the importance of understanding the birth of a team and that team-building activities are the most effective in the early stages of a team’s life cycle (DeRue et al., 2015). Thus, managers who aim to enhance activity and viability in the early life cycle of a team might consider involving group members who are flexible and have the ability to engage in interactions with other group members regardless of their experience, knowledge and skills (Furst, Reeves, Rosen and Blackburn, 2004).

The present study has the following limitations. We aimed to explore how leadership evolved in the history of an online community from the embryonic stage to the emergence stage and to the established stage. We collected our data on leadership structures through Twitter, Gitter, and GitHub. However, GitPoint’s members may have used other social networking platforms, which were not included in our data collection. Nevertheless, we are confident that the empirical data that we explored form the embryonic stage and the emergence stage to the established stage allowed us to model what happened before, during and after the existence of the online community of GitPoint.

Finally, future research should examine the evolution of leadership roles and the shift in power relations among group members at different points in the team’s life cycle. The findings of our analysis showed that the individual participants adopted situationally specific roles at several points in the online community’s lifecycle (e.g., the amplifiers role in the embryonic stage, the organiser role in the emergence stage, and the reviewer role in the established stage). Therefore, the emergence and fading away of roles could depend not only on members’ characteristics but also on the time and the space of their participation and such dynamics could prove generative for subsequent research.
References


