Functional Roles and Career Paths in Wikipedia

Ofer Arazy
University of Alberta
University of Haifa
ofer.arazy@gmail.com

Felipe Ortega
University Rey Juan
Carlos
felipe.ortega@urjc.es

Oded Nov
New York
University
onov@nyu.edu

Lisa Yeo
Loyola University
Maryland
mlyeo@loyola.edu

Adam Balila
University of Haifa
adambalila@hotmail.com

ABSTRACT

An understanding of participation dynamics within online production communities requires an examination of the roles assumed by participants. Recent studies have established that the organizational structure of such communities is not flat; rather, participants can take on a variety of well-defined functional roles. What is the nature of functional roles? How have they evolved? And how do participants assume these functions? Prior studies focused primarily on participants' activities, rather than functional roles. Further, extant conceptualizations of role transitions in production communities, such as the Reader to Leader framework, emphasize a single dimension: organizational power, overlooking distinctions between functions. In contrast, in this paper we empirically study the nature and structure of functional roles within Wikipedia, seeking to validate existing theoretical frameworks. The analysis sheds new light on the nature of functional roles, revealing the intricate "career paths" resulting from participants' role transitions.

Author Keywords

Peer-production; Wikipedia; organizational structure; functional roles; role transitions.

INTRODUCTION

Recent years have seen the emergence of a community-based model for the production of knowledge-based goods such as Wikipedia and open-source software [44]. As participants become more involved in their projects and gain the community's trust, they gradually move from the periphery to the community core, gaining access to more sensitive and influential decisions [1].

The problem of motivation for participation in peer production communities has attracted significant attention at the academic community, and numerous studies have discussed the various motives driving volunteers to contribute their time, effort, and expertise. Less is known, however, about how these volunteers sustain and increase their participation, take on additional responsibilities, and

become involved in the project's administration. Although on the surface peer production projects may be thought to be non-hierarchical, in reality a core group of leaders usually emerges (often through formal election processes) to provide centralized coordination, manage quality control processes, mediate conflicts, and develop organizational policies [4, 16, 43].

Online production communities are characterized by a core-periphery structure, where the majority of contributors are involved in few tasks and participate at the community's periphery, while a relatively small portion of contributors take on additional responsibilities and constitute the core [12, 14, 20, 24, 30]. The core-periphery conceptualization concentrates on participants' power (or authority) within the community, emphasizing lateral movements from the community's fringes to positions of responsibility and influence [56]. Prior works in the area have tried to model the process of transitioning from the community's periphery to the core. For example, the 'Reader to Leader' (R2L) framework [48] attempted to synthesize these earlier works and provide a comprehensive conceptualization on the successive steps volunteers take on their way to community leadership. According to this framework, while the main path from the periphery to the core leads through these successive stages, alternative routes are possible, including those representing attrition (i.e. transitions from the core to the periphery). Notwithstanding the importance of organizational power in characterizing the composition of online communities, this conceptualization overlooks other important dimensions, such as the functional organization of community work. Thus, extant conceptualization lack the capacity to distinguish between different functions situated at the same level of the organizational structure. Traditional organizations are characterized by a functional organization, dividing work between departments such as: Finance, Operations, Marketing, etc., and the functional organization is at the core of scholarly literature on management and organizations [40]. However, the study of functional roles has been largely absent from the literature on online communities.

Therefore, the objective of this study is to investigate the functional organization of online production communities: the nature of functional roles and the way in which these roles are traversed when participants move from the periphery to leadership positions. In part, our study could

be viewed as an empirical investigation of existing conceptualizations of emergent leadership, namely the Reader-to-Leader (R2L) framework [48]. Despite the significant attention that the R2L framework has received within the research community, to date, this framework has not been empirically evaluated. In fact, Preece and Shneiderman [48] acknowledge that "Even though our [R2L] framework and discussion of its components are supported by research, the framework needs empirical testing, which is an obvious next stage of this work" (p. 24).

The setting for our study is Wikipedia, one of the most notable examples of peer production [9]. Wikipedia was able to recruit thousands of volunteers to produce millions of encyclopedic entries in 287 languages, and develop extensive policies and mechanisms for governing its collaborative authoring process. Wikipedia's success attracted the attention of scholars, who investigated Wikipedia's organizational model [4, 24, 49]. Our research methodology employs publicly available Wikipedia system logs to identify participants' access privileges. It builds on the framework from [3] to map access privileges onto organization roles, and analyzes role transition for a large sample of Wikipedia contributors. In addition, we perform a qualitative analysis of a selected subset of editors, adding depth and context to the quantitative data analysis.

Hence, the goal of this study is to address the following research questions:

RQ1: How are the functional roles of Wikipedia characterized and what activity profile is associated with each role?

RQ2: How has Wikipedia's functional composition evolved over time?

RQ3: How do contributors to Wikipedia transition between roles on the path from community's periphery to the core?

Our investigation yields important insights about the routes to leadership in peer production projects, informs extant theory in this area, and offers guidelines to designers and custodians of online communities.

SOCIAL ROLES AND ORGANIZATIONAL STRUCTURE
The concept of social role has been the subject of extensive analysis in sociology [10]. Social roles encapsulate the social context, history of actions, structures of interaction, and the attributes people bring to the interaction by providing a meaning system, which both constrains and enables action [37]. The importance of this concept lies in its utility: the classification of types of social relations and behaviors into a smaller set of roles reduces the analytic complexity of social systems and facilitates the comparative study of populations across time and settings [32]. Roles can be understood through two primary dimensions: structure and culture. The structural definition of roles pertains to commonalities in behavior patterns, while the cultural dimension refers to roles that are recognized in

social (or organizational) settings and differ in terms of their accessibility (the extent to which it is easy to accept a role) and situational contingency (the contextual factors that affect action). Roles are resources that help people accomplish their goals, and tools used in the establishment of social structure [8]. Hence, social (or organizational) structure can be viewed as an ecology of roles.

Formal organizations are generally understood to be systems of coordinated and controlled activities that arise when work is embedded in complex networks of technical relations and boundary-spanning exchanges [38]. Organizational structure affects organizational action in two primary ways. First, it provides the foundation on which standard operating procedures and routines rest. Second, it determines individuals' role assignment and their access to decision-making processes. Hence, the structure determines the extent to which individuals at different organizational positions can influence the organization's actions [39].

ORGANIZATIONAL ROLES IN ONLINE COMMUNITIES
Online communities, and in particular production communities, have been investigated extensively in recent years. Relevant to our inquiry of roles and organizational structure are prior works that studied community members' roles. Previous investigations have discovered that users often follow very distinctive patterns of activity, playing roles in their online community [58]. Production communities are often described in terms of a core-periphery structure, which entails a dense, cohesive core, and a sparse, unconnected periphery [11, 13, 20, 33, 35, 45, 52]. Contributors play different roles: the majority of participants, who are not very active, are situated at the community's periphery; and a small minority, who take on additional responsibilities and privileges, constitute the community's core. For example, Long and Siau [35] found that social interaction patterns in open source projects start with a single hub, and as the project matures the configuration evolves into a core/periphery structure. Most prior studies in the area have defined participants' position on the periphery-core continuum based on the quantity and types of activities they perform [5, 30, 34]. For example, [4] categorized contributors into two primary classes: administrative- vs. content-oriented; and [57] applied a combination of interpretive and network analysis methods to identify key roles. Much less work has been devoted to studying contributors' formal duties [16]. A few prior works provide a partial description of Wikipedia's formal roles [24, 42, 55]. Building on these earlier works, a recent study [3] has analyzed access privileges in Wikipedia and developed a comprehensive description of Wikipedia's organizational structure, including twelve roles (organized in a power hierarchy: *Unregistered Users, Registered Users, Manually Registered Users, Technical Administration, Border Patrol, Quality Assurance, QA Technicians, Administrators, Security Force, Directors, Privacy Commissioner,* and *Benevolent Dictator* (see Table 1).

Level on Org. Chart	Role	Description	Access Privileges
Level 0	*Unregistered Users*	Non-community members	*
Level 1	*Registered Users*	Newly registered users	user
			autoconfirmed
	Manually Registered Users	New users who had to be manually registered to bypass some restrictions	confirmed
			IPpblock-exempt
Level 2	*Technical Administration*	Privileged users responsible for the administration of the technical aspects (e.g. user accounts, files)	filemover
			accountcreator
	Border Patrol	Users responsible for fighting vandalism by reverting malicious edits	rollback
	Quality Assurance	Privileged users responsible for patrolling Wikipedia and ensuring content quality	reviewer
			autoreviewer
	QA Technicians	Users who develop automated tools (i.e. edit filters) to assist quality assurance work	abusefilter
Level 3	*Administrators*	Highly involved users that are responsible for the social administration of the English Wikipedia community	sysop
			bureaucrat
Level 4	*Security Force*	Highly trusted users who are working to keep malicious users out and combat intentional manipulations of content	oversight
			checkuser
Level 5	*Directors*	Key users responsible for oversight of the Wikimedia organization	steward
			importer & transwiki
	Privacy Commissioner	High-ranking users who investigate complaints about violations of privacy policy	ombudsman
Level 6	*Benevolent Dictator*	Jimmy Wales; responsible for defining high-level policies and norms and for overall direction of the community	founder

Table 1: The organizational structure of Wikipedia (adapted from [3])

ROLE TRANSITIONS IN ONLINE COMMUNITIES

Traditionally, when individuals advance their careers in organizations, they move through positions that provide them with increasing degrees of responsibility, often including vertical authority over others [14]. Recently, organizational theorists have argued that traditional notions of career progression make several assumptions that no longer reflect workplace realities, such as: mobility within a single large organization, stability in the organization and its environment, and hierarchical (i.e. vertical) progression [7]. Thus, there is a need to pay more attention to lateral (or horizontal) progression that relies on expertise, identity, achievement, and community involvement, rather than on hierarchy and formal authority [27]. Such lateral movements often involve transitions between functional areas.

Online production communities are characterized by movements from the community's periphery to positions of responsibility and influence at the community's core [21, 43, 56]. Although many studies have discussed the different ways in which users participate in online communities and the various activities they take part in, no study offers an overall framework of career paths in online production communities. Prior research has concentrated on descriptions of the core members' (leaders, owners) duties and activities [15, 46], the distribution of authority among these members [20], as well as on the factors affecting a participant's promotion to a leadership role [13, 17, 18, 33]. For example, [46] showed that the nature of work differs between Wikipedians and non-Wikipedians. Others have examined promotion procedures: [13] investigated the criteria for promotion decisions (i.e. the RfA process and adminship decisions); [19] studied the relationship between editors position within the social network and promotion decisions; and [33] demonstrated that promotion depends on the candidate's relative activity levels. These studies – although rich in detail – provide a restricted and simplified conceptualization of the transitions towards leadership.

Preece and Shneiderman [48] provide an overall framework of emergent leadership in online communities by

synthesizing prior works in the area. They present a nuanced classification of activities, organizing roles in successive levels of involvement, termed the "Reader to Leader" (R2L) framework. The *reader* category represents the lowest level of participation, and includes tasks such as browsing, surfing, and searching. Reading of user-generated content posted by other participants can be thought of as legitimate peripheral participation [31]; it is a necessary condition for knowledge reuse, and is a typical first step toward more active participation [47]. Next, the *contributor* category represents a gradual move towards larger and more frequent contributions as participants' confidence grows and they feel empowered and appreciated [59]. Activities in this category include: contributing comments and responses to others' postings, rating previous contributions, tagging and categorizing existing content, and making small knowledge contributions (e.g. correcting an error, adding a hyperlink). The next level of participation, *collaborator*, involves two or more contributors discussing, cooperating, and working together to create something or share information [22]. These activities allow participants to bring more from themselves and more fully express their opinions and ideas. Finally, the *leader* category represents the highest level of participation. In addition to being active contributors and collaborators, leaders (also referred to as "owners"), are responsible for social management tasks (e.g. establishing community norms and explicit policies or conflict resolution) and administration of the technology infrastructure [16]. A unique feature of the R2L framework is that it provides an intricate picture of role transitions, acknowledging both upstream and downstream transitions, as well as bypasses of intermediate stages (e.g. from *contributor* directly to *leader*, bypassing the *collaborator* stage). Notwithstanding these merits, the R2L conceptualization seems to collapse two important dimensions: organizational power and function, impeding the distinction between different functional roles that are situated at a similar point on the Reader–Leader continuum. We also note that, despite its influence in the CSCW community, the R2L framework has not yet been examined empirically. The objective of the current study is, thus, to try and fill in these gaps in the literature, empirically evaluate role transitions and assess the extent to which the R2L conceptualization captures the reality of career paths within production communities.

METHOD

Research Setting
The focus of this empirical investigation is the English Wikipedia and its community of editors. We base our data gathering and analysis on Wikipedia's system logs; harvesting these logs can reveal important insights about members' ongoing behavior in its natural setting.

Over time, the Wikipedia community has developed a comprehensive and detailed set of procedures for governing the collaborative editing process, including a well-defined scheme of roles access privileges [14]. We focus on human editors and exclude software bots [25] from this analysis. We note that in Wikipedia, anyone can contribute without registering. This results in two types of participants: anonymous (represented by the '*' tag), whose contribution is recorded and associated with an IP address, and registered users, whose contributions are associated with a user name. Since we cannot track subsequent role transitions for participants who initially contribute anonymously (it is not possible to univocally associate IP-based contributions with a user name when one starts as an anonymous contributor and later registers an account), anonymous contributors were excluded from our analysis.

Our mapping of access privileges onto organizational roles, illustrated in Table 1, was based on the earlier work of [3], employing a combination of top-down (review of relevant scholarly literature and Wikipedia's own definitions of roles and access privileges) and bottom up methods (using log data of all Wikimedia users and statistical analysis techniques). A key advantage of using the framework at [3] is that it arranges Wikipedia roles on a continuum from the community periphery to the core, thus allowing us to analyze movements up (or down) the organizational chart. Transitions between roles were defined as follows: when a participant is assigned an access privilege that is associated with a more advanced role, we designate her as being promoted to this role; and when a participant loses all privileges associated with the advanced role, we designate her as being demoted to a less advanced role.

Sample
Our sample included 2,174 Wikipedia editors who are all registered members. Since articles' topical categories may attract different types of contributors, we sought a sample of users who have contributed to a representative sample of Wikipedia articles. Therefore, our sampling procedure used a seed of 96 articles that provides a representation of Wikipedia's topical categories and which has been employed in earlier studies [4, 6]. This set of articles (created in January 2007) was selected based on randomization and a stratified sampling of Wikipedia's topics, congruent with Wikipedia's top-level classification[1] [29] (categories: culture, art, and religion; math, science, and technology; geography and places; people and self; society; and history and events). From this original set of 96 Wikipedia articles, 3 were discontinued and their edit history is no longer available, leaving us with 93 articles.

Our sample of Wikipedia participants included every editor who contributed at least one edit to any of these Wikipedia articles (prior to the 2007 cut-off). For the reasons mentioned earlier, we excluded anonymous contributors and software bots. Our procedure rendered a sample of 2,174 distinct

---

[1] For a list of Wikipedia top-level categories, please refer to http://en.wikipedia.org/wiki/List_of_overviews

participants. We followed these participants from the time they registered with Wikipedia until June 2012, recording every change in their Wikipedia roles during this period.

Data Collection
In order to address RQ1, we employed data harvested from Wikipedia logs. To profile the activities of the various roles, we retrieved data regarding the contributions of our user sample across all Wikipedia namespaces, including: *main article* pages (production work); *talk pages* (coordination work); *user* and *user talk* pages (communicating with others); Wikipedia and other pages (community work) [17]. For the second RQ, we gathered data regarding the evolution of roles within Wikipedia from an internal report by a senior Wikipedian[2]. Finally, for tackling RQ3 and collect data regarding role transitions, we first queried the Wikipedia API (http://en.wikipedia.org/w/api.php) to determine whether a participant is a registered user; we then used the API to receive the list of role changes for each participant in our sample, resulting in 7,563 role-change events. Organizing the role change history, we determined the participant's role at each point in time, and recorded role transitions.

In order to verify the accuracy of the user-role data, we compared it to other sources: the rights log history[3] and alternative sources archiving role transitions at the early days of Wikipedia[4]. In cases of mismatch (7% of the events), a discrepancy was flagged, and was resolved manually. Some of the more common discrepancies included software bots that were initially registered as regular users (and once identified as bots, were excluded from the sample) and administrators who lost their privileges due to inactivity (in such cases, we tracked the missing date for this event). In addition, in a few cases transitions that were recorded properly reflected a technical issue, rather than a role change (e.g. a particular privilege was added, removed a few minutes later, and then a different privilege was assigned). After a careful manual inspection of short-duration event sequences, we excluded those events that did not seem to represent an actual role transition. Once we compiled a reliable list of access privilege transitions, we mapped them onto Wikipedia roles (as specified in Table 1 above) to arrive at role transitions. Many of the access privilege transition events did not induce changes in Wikipedia roles (e.g. a 'filemover' taking on the additional privileges of an 'accountcreator'; both corresponding to the *Technical Administrator* role). We focused our attention on those access privilege transitions that entail a move between role categories.

In order to provide richer context and ground the overall analysis in additional data sources and individual user cases, we identified a list of seventeen editors that capture the various role transition patterns, focusing on unique cases (e.g. editors that transition particularly quickly or slowly; extremely active editors, cases of demotions). Altogether, we investigated 17 editors with IDs: 11; 96; 20134; 29678; 44020; 82835; 124324; 126457; 274040; 889851; 1403682; 1812441; 1862829; 2164608; 2267145; 3162157; and 3138762.We gathered information about these editors from a variety of Wikipedia sources, including editors' personal pages and talk pages; requests for adminship (RFA)[5]; requests for de-adminship[6]; the logging table (e.g. checking who is operating a bot)[7]; and pages tracking dispute resolution cases. This selected set of editors captured cases to exemplify the various role-transition dimensions; in terms of *seniority*: 8 veterans (joining Wikipedia up to 2004) and 9 that joined later; *editing activity*: 7 with relatively low activity patterns across the various namespaces (no more than few thousand edits), 5 moderate (tens of thousands of edits), and 5 highly active editors (more than one-hundred-thousand edits); *rank*: 1 editor that never made a role transition, 14 that have gained Level 2 privileges, 13 Level 3 cases, and 1 editor at Level 5; *pace*: 5 that made fast transitions (less than a year), 5 that transitioned at a regular pace (one-three years), and 6 slow movers (transitioning after more than three years); and *controversial*: 5 editors that have been blocked (some only temporarily and others for prolonged periods) and 7 that have been demoted (either voluntarily or forced by the community). We used these cases to illustrate role transitions and career paths from a micro perspective.

RESULTS
Out of the 2,174 editors in our set the majority were from the community's periphery (i.e. auto and manually-registered users), relatively few were at the intermediate levels, and no instances were recorded for the top levels (namely, the 'steward', 'importer & transwiki', 'ombudsman', and 'founder' privileges)[8]. In total, we recorded 21 *Manually Registered Users* (1 confirmed and 21 IPpblock-exempt); 12 *Technical Administrators* (9 filemovers and 5 accountcreators); 127 users with *Border Patrol* (i.e. rollback) privileges; 287 at the *Quality Assurance* role (244 reviewer and 150 autoreviewer); 49 *QA Technicians* (i.e. abusefilter); 436 *Administrators* (435

---

[2]http://en.wikipedia.org/wiki/User:NoSeptember/The_NoSeptember_Admin_Project

[3]https://meta.wikimedia.org/w/index.php?title=Special%3ALog&type=rights

[4] Namely the Meta:Bureaucrat log (https://meta.wikimedia.org/wiki/Meta:Bureaucrat_log), which records privilege changes prior to 10 December 2004.

[5]en.wikipedia.org/wiki/Wikipedia:Successful_requests_for_adminship

[6] en.wikipedia.org/wiki/Wikipedia:Requests_for_de-adminship

[7] For members joining Wikipedia prior to September 7, 2005, some information is missing from the logging table; in those cases we used the time of first edit as a proxy for registration date.

[8] Across the entire Wikipedia community (all languages), there are only 46 participants with the *'steward', 'importer & transwiki', 'ombudsman',* or *'founder'* access privileges. This explains why they were almost absent from our sample.

sysops and 10 bureaucrats); 11 users as *Security Force* (9 oversight and 9 checkuser); and 1 *Director* (with importer privilege).

To address RQ1 concerning the activity profile of each role, we partitioned each editor's trajectory into different temporal sections, each listing the roles held in that time period, as well as the length and activities across the various Wikipedia namespaces. Then, for each role, we aggregated the activity of all editors. In terms of overall activity, *Technical Administrators* and *QA Technicians* have the highest average daily activity, with 113 and 42 daily edits respectively. Level 2 and 3 roles make on average 11-14 edits; while those at the entry (Level 1) and top levels (Levels 4 and 5) make less than 5 daily edits. In terms of effort across the various Wikipedia namespaces, although all roles focus primarily on editing main pages, each role is characterized by a distinct activity profile, as illustrated in Figure 1. For example, *New Registered Users* focus primarily on editing main pages (74% of their edits); *Technical Administrators* and to a lesser extent *Border Patrol* are relatively inactive on main pages, but instead concentrate on coordination work (40% and 21% respectively in talk pages); *Security Force* is characterized by little production work (35% on main pages), and greater focus on pages associated with policy creation and enforcement (37% and 22% respectively on user talk and Wikipedia namespaces); and *Directors* are highly active on Wikipedia and Wikipedia talk namespaces (22% and 19% respectively). For example, editor #3162157, who joined on January 2007 and never transitioned beyond the Registered User role, made 301 edits to main pages; editor #274040, who joined on May 2005 and gained various Level 2 (*Technical Administrations, Quality Assurance, Border Patrol*) and Level 3 privileges, made roughly 30,000 edits to main pages, 3,700 in talk pages, 25,000 in user and user talk pages, and 14,000 in other namespaces; and editor #96, a Director that joined at 2001, made 4,200 edit to main pages and 1,000 edits across other namespaces.
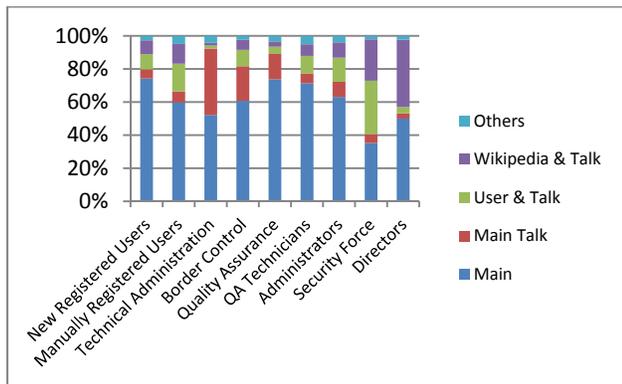

Figure 1. Activity profiles associated with functional roles (as proportion of roles overall activity).

Overall, we notice that editors at Level 3 and above consistently present a significant number of contributions to meta pages which pertain to community activity, such as talk, user, user talk, Wikipedia or Wikipedia talk pages. Level 2 roles, on the other hand, present very high levels of activity at very specific namespaces linked to their particular duties (e.g. editors in Technical Administration role spend substantially higher percentage of their work on category pages, and QA Technicians focus more on template pages).

To control for the effect of topical categories, we partitioned the sample into sub-samples (by building on the topical organization of the articles that were used to seed the editor sample), and analyzed the activity patterns of each topical sub-sample. A multivariate analysis of variance (MANOVA) revealed very similar activity patterns across topics – for all Wikipedia namespaces – suggesting that the results are independent of particular topical domains.

Our next analysis focused on the evolution of roles, addressing RQ2. Figure 2 illustrates the temporal evolution of roles within Wikipedia. As the figure shows, in the early days of Wikipedia, the entry level positions (Levels 0 and 1: unregistered and registered members) were joined by 'core' (or leadership) roles: Level 3 (sysop; 2002); Level 4 (checkuser; 2005); Level 5 (steward; granted to the founder Wales in 2002); and Level 6 (founder; 2002). In 2004 steward privileges were granted to other beyond the founder, whereas in 2006 the privileges of oversight (Level4), ombudsman, importer, and transwiki (Level 5) were introduced. Only at the next phase of Wikipedia's evolution - in 2008 and 2009 (with later additions in 2011) was the intermediate functional layer introduced and Level 2 roles emerged.
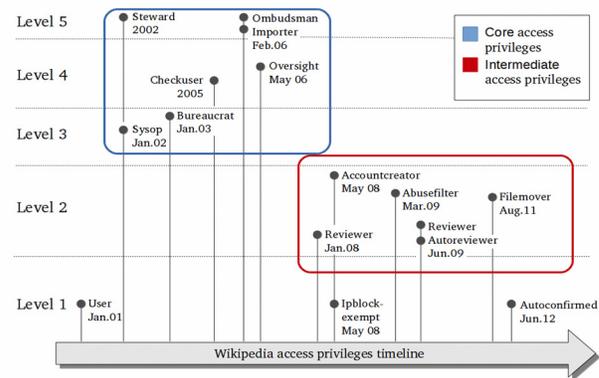

Figure 2. The Evolution of Wikipedia's access privileges

In order to understand how these various roles were populated as Wikipedia evolved, we studied the distribution of access privileges by plotting the levels reached by the users in our sample. In the early years of Wikipedia, the vast majority of users were al Level 1; by 2005, as the second phase of Wikipedia's life began, the project increasingly gained traction, and the number of new editors joining rapidly grew; realizing the need to curate the

immense amounts of new content being constantly created, Level 2 functions were introduced and gradually populated. Figure 3 shows the changes in the proportion of editors (y-axis) in each of the top five levels. We can see that the proportion of Level 3 users (sysop, bureaucrat) started to represent a noticeable fraction of the total population circa April 2005, and from January 2008 it has remained quite stable. Furthermore, since the introduction of levels in the functional layer (Level 2), starting in January 2008, the proportion of users taking roles from this layer has rapidly increased, reducing the proportion of entry level users (Level 1). This finding offers evidence of the emergence of a group of intermediate level users focusing on tasks such as content curation and fighting vandalism.
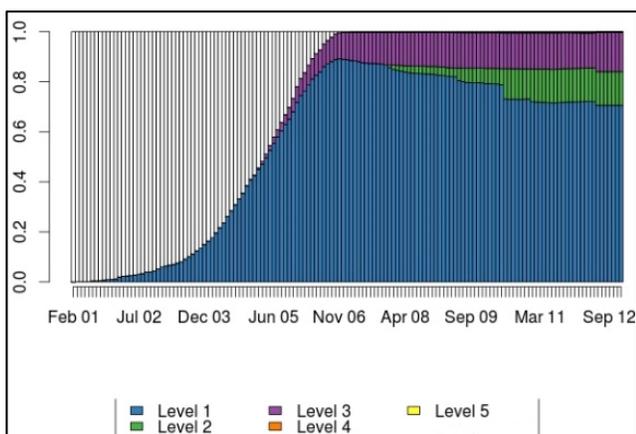

Figure 3. The proportions of role population over time.

An example from our selected subset of editors could illustrate this evolution. Consider editor #126457, a librarian who worked at Cambridge College and the British Antarctic Survey, and primarily contributes to pages covering historical figures, events, and places (in addition to his community activity). Editor #126457 became a member on October 2004, an *Administrator* (Level 3) on November 2005, and added the role of *QA Technician* (Level 2) on July 2009.

Once we established the nature and evolution of Wikipedia's functional roles, we turned our attention to investigating editors' role transitions. In order to address RQ3, we first counted the times each role transition type occurred in our sample. In total, we recorded 1,103 role transitions (see Table 2). Out of the 2,174 participants, 732 (34%) transitioned to various organizational roles, while the remainder 1,442 (66%) never moved beyond the entry levels (i.e. registered users)[9].

We observed both horizontal and vertical transitions. We note horizontal transitions at the intermediate level (Level

2), where there are transitions between the *Border Patrol* and *Quality Assurance* roles. For example, editor #20134 registered with Wikipedia in August 2003, gained the *Quality Assurance* role in June 2010, and six months later added the *Border Patrol* role. Vertical transitions from one organizational level to another include transitions such as the one from the *Registered Member* (Level 1) to *Quality Assurance* (Level 2) roles, or from *Border Patrol* (Level 2) to *Administrators* (Level 3) (for example, editor #1862829 registered on July 2006, gained the *Border Patrol* role on April 2008, and became an *Administrator* on October 2008). Overall, the majority of transitions were vertical (995 instances; 90%). For the horizontal transitions, it is interesting to note that while there are some roles with little to/from transition (e.g. *QA Technicians*), other roles are characterized by frequent horizontal movement (e.g. *Border Patrol*)[10].

For the vertical transitions, while in some cases the transition to the community's core is sequential. (e.g., 346 transitions from Level 1 to 2 and 21 transitions from Level 2 to 3), the majority of vertical transitions skip intermediate levels (e.g., 421 transitions from Level 1 to 3). Examples from our selected set of editors could illustrate the upward transition process: editors #2267145, #11, and #126457 who made the transition to administrative role (i.e. Level 3) prior to the introduction of the functional layer (at 2007, 2006, and 2005 respectively) moved directly from being *Registered Users* to administrative position, the latter two editors adding on Level 2 roles at a later time; in contrast, *editors* #1862829 and #44020 attained their administrative position later (at 2008 and 2010 respectively) and have made linear transitions: Level 1 => Level 2 => Level 3.

Despite the majority of vertical transitions are upward (792 cases), a relatively large number of transitions were downward (195 cases). In particular, we noted 81 cases where core community members (Level 3) shed their special privileges and become regular members. For downward transitions, we distinguish between cases where an editor was first granted the higher level role and later added a lower level role (as in the case of editor #126457), and cases of demotion where an editor loses the higher level privilege (for example, editor #29678, who gained *Administrator* privileges on March 2009, lost them due to inactivity on October 2011). Table 2 presents these various role transitions.

In order to represent the intricate role transition dynamics, Figure 4 illustrates promotion and demotion role transitions (for clarity, we excluded from the figure same-level transitions, as well as cases of downward transitions that do not entail demotion).

---

[9] Note that after becoming an *Administrator*, if one adds a series of Level 2 roles, each such addition is recorded as a transition from Level 3 to the Level 2 role.

[10] Cases where an editor played multiple Level 2 roles and lost one are not captured in our analysis.

Level	From \ To	Registered Users	Manually Registered Users	Tech. Admin.	Border Patrol	Quality Assurance	QA Tech.	Administrators	Security Force	Directors
		1		**2**				**3**	**4**	**5**
1	Registered Users		10	+2	+109	+217	+16	+420	+4	+1
1	Manually Registered Users	7				+2		+1		
2	Technical Administration	-1								
2	Border Patrol	-14	5	6		66		+16		
2	Quality Assurance	-14	2	6	12			+4		
2	QA Technicians	-2	1			1		+1		
3	Administrators	-81	6	4	8 / -7	4 / -3	40 / -2		+7	
4	Security Force						1			
5	Directors									

Table 2. Role transitions count; a plus sign ("+") represents promotions, and minus sign ("-") demotions; no sign indicates a same-level transition, or cases where higher-level roles were retained while picking up lower-level privileges.
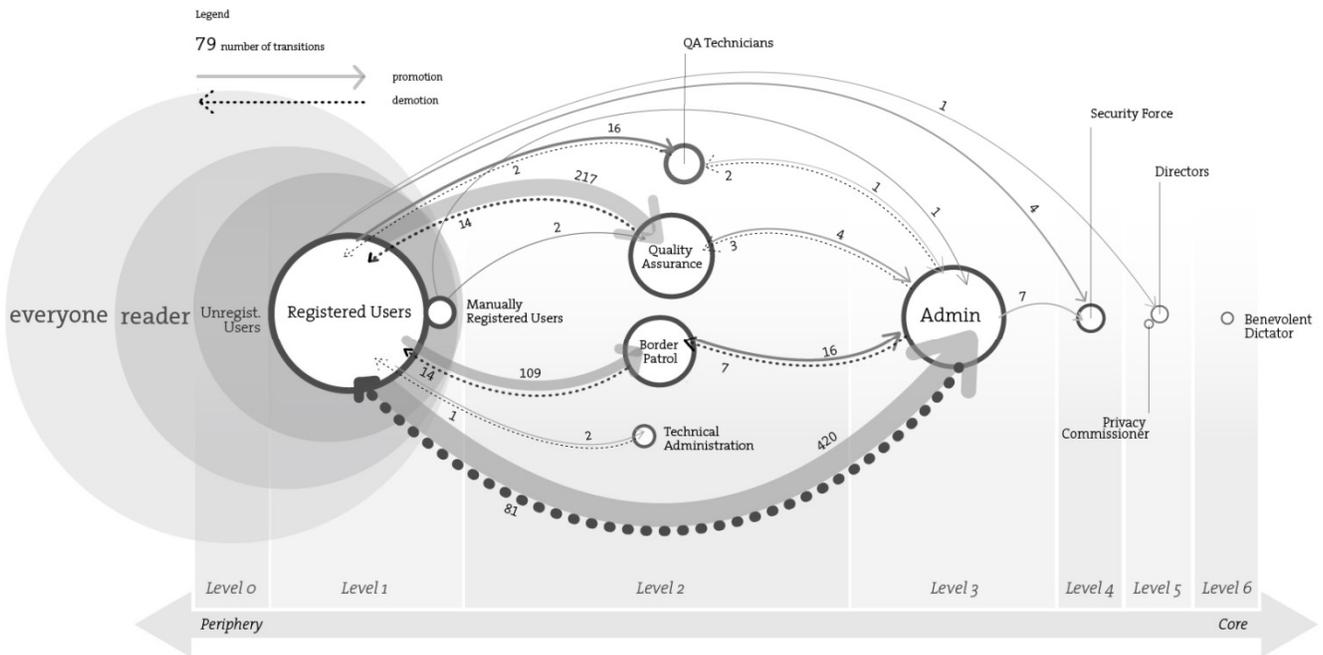

Figure 4. Promotion and demotion transitions. Circle size corresponds to the number of editors in the sample; upward solid arrows represent transition to a higher level; downward dotted arrows represent cases where editors lost higher level privileges. Width of arrows represents number of role transitions count. For clarity, same-level transitions, as well as downward transition that do not entail demotion (i.e. editors kept their higher-level privilege), are excluded from the figure.

The analyses described above highlight the importance of the mid-level strata of contributors. These contributors are characterized by functional focus and diversity and represented by Level 2 roles in Wikipedia's organizational structure. In order to fully understand how these roles are attained, we performed an analysis of incoming transitions into Level 2 roles (see Figure 5).
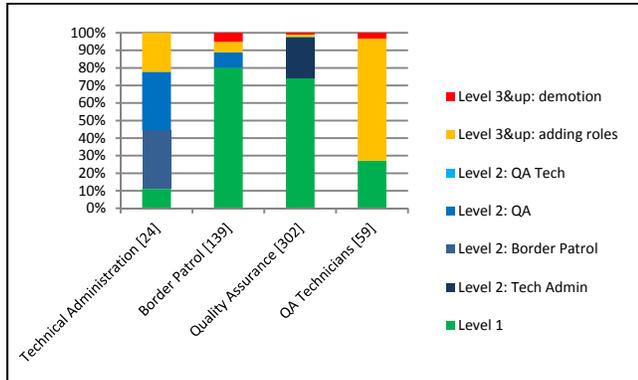

Figure 5. Sources of incoming transitions into Level 2 roles.

This analysis reveals that in addition to the differences in terms of activity profiles reported earlier, the various Level 2 roles seem to represent different career paths. For example, transitions into *Technical Administration* are primarily through other Level 2 roles (namely, *Border Patrol* and *Quality Assurance*); progression into *Border Patrol* is directly from Level 1; and the role of *QA Technicians* is often appended to *Administrators*.

In addition to the patterns emerging from our quantitative analysis (as discussed above), the qualitative analysis of the selected 17 editors revealed some interesting insights. Although we cannot validate statistically the significance of these findings, they do highlight some interesting patterns. The discussions associated with promotion decisions reveal that editing activity is a key consideration (see [13, 33]). As we have shown earlier, while all editors focus their editing activities on main pages (see Figure 1), administrators consistently present high activity levels in that namespace. The administrators in our representative set - #1862829, #44020, #2267145, #11, #889851, #126457, #29678, #274040, #124324, #82835, #1403682, #2164608, and #1812441 - had an average of 60,000 edits in main (ranging 5,000-120,000) up to our study's cutoff. In addition, contributions to other namespaces, which represent communications between editors and community activity, also seem to be relevant for promotion. For example, editor #2267145 has made over 50,000 edits to talk pages; #889851 has edited user and user Talk pages over 36,000 times; and editor #82835 made over 82,000 edits to talk pages, 38,000 to user and user talk pages, and roughly 85,000 other meta pages. This data suggests that while a significant editing activity in main pages is required for promotion to administrative position, active participation in a variety of other namespaces is also essential to achieve a promotion. Timing also seems to have influenced the promotion process, and our qualitative analysis reveals the very low volume of participation in older voting processes for *Administrator* (in contrast with the high community participation from 2007 onwards). An extreme example is editor #96, who was made an *Administrator* on Feb. 2004 with just 6 positive votes.

The in-depth examination of 17 editors also shed additional light on the process of downward transitions. As revealed through the statistical analysis, a downward transition can entail a demotion (when higher-level privileges are shed; e.g. editor #29678) or a case where a higher-level role - typically, Level 3 - adds lower-level privileges - commonly Level 2 (e.g. editor #126457). When studying the reasons for demotions, we found evidence for both community-initiated and self-demotions. Reasons for community-initiated demotions were prolonged inactivity period (as in the case of editor #29678) or abuse of Wikipedia norms (often abuse of rollback privileges). For example, editors #889851, #274040 and #124324 lost administrative privileges due to misuse of administrative tools, failing to respond to community concerns, or inappropriate off-wiki behavior. Self-demotions are often the result of continuous controversies and confrontations (e.g. editors #82835, #1403682, and #2164608). We note that demotion cases are often linked to blocking incidents, where an editor is prevented from editing activities for some period (in few cases, indefinitely). For instance, editor #274040, who was forced to give up administrative privileges, was banned 12 different times (due to page-move vandalism, blocks and violations of earlier bans). In fact, regaining administrative privileges required a clean record of 'blocks'. It is interesting to note that extremely active users who contribute across the various namespaces, are often involved in controversies and disputes, and in multiple occasions have been banned.

Our qualitative analysis also revealed that some of the users at higher ranks make use of advanced tools for automating editing tasks (that is, they operate 'bots'), leading to unusual peaks in editing activity. Most notably, we can identify the use of the AutoWikiBrowser tool as one of the main triggers for an unusually high number of edits. For example, editor #1862829 is not only one of the developers of AutoWikiBrowser, he is also one of the 20 most active Wikipedians and runs a bot to carry out tasks that would otherwise be tedious.

Our study of the selected set of editors also examined participation in additional community activity. When analyzing participation in WikiProjects – projects typically dedicated to improving articles in a particular domain – we were not able to find evidence linking WikiProject participation to any particular user profile.

DISCUSSION

Recent years have seen increasing interest in investigating organizational roles within online communities and the process by which contributors transition between roles. Conceptualizations such as the Reader-to-Leader framework [48] have received significant attention. Yet, to date, there has been no quantitative empirical evaluation of role transition processes. Our study of career paths in Wikipedia over a period of over ten years provides insights into the dynamics of transitions between the community's periphery and core. The results highlight the significance of functional roles for understanding how online production communities organize their work. In particular, our findings call into focus the functional middle organizational level (i.e. Level 2) of Wikipedia.

The findings show that formal roles determine users' actual activity patterns across a variety of editing tasks (i.e. namespaces). To date, much of the research on the organizational structure within production relied on activity counts for determining one's position on the periphery-core continuum [5, 30, 34].

A second contribution this study makes is in delineating the evolution of Wikipedia's organizational structure. Earlier studies have noted Wikipedia's increased bureaucracy [14] and have associated the growing complexity of Wikipedia's organizational structure with decreased effectiveness [26, 54]. Our findings, on the other hand, suggest that the introduction of additional functions (namely, Level 2 roles) was necessary for curating the mounting number of new contributions. This finding is in line with recent works which suggest that online production communities that fail to create such organizational structures are not sustainable [41].

We also make a contribution in describing role transition dynamics within Wikipedia. Our empirical evaluation validates prior knowledge in the area [13, 23, 35, 43] and demonstrates that the number of community members decreases as we grow closer to the core: only a relatively small number of contributors proceed to Level 3 and even fewer continue to becoming a core member (i.e. Levels 4 and higher). Level 2, because of its late introduction, is still not fully developed, but we expect that the number of contributors at this level would grow and surpass the number of administrators (Level 3). Our results (see Table 2) clearly show that not all paths are equally traversed; in fact, some paths are never traversed (at least in our sample), suggesting that – although not formally articulated – there are de-facto career paths within Wikipedia.

A key finding of our study is that the strength of paths (in terms of the frequency in which participants traverse them) does not correspond to the predictions of extant conceptualizations. In particular, against the predictions of the R2L model, which suggests a primary path to leadership characterized by a linear sequence from the periphery to the core of the community (with some additional, non-linear secondary paths that may skip a particular stage), findings from the present study suggest that these non-sequential paths are the rule rather than the exception. Namely, in most cases participants move directly from the entry levels to the community's core (skipping intermediate functional roles). This may be an artefact of the late introduction of Level 2 roles, although even after 2009 most of the Wikipedians arriving at Level 3 did so directly, bypassing Level 2 roles. The implication of this finding is that Level 2 roles must not be seen as a step towards becoming a community leader; instead, they represent functional positions that are important in their own right.

While most of the prior studies in the area emphasize upward transitions towards leadership positions, our study provides evidence for the existence of transitions from the core back to the community's periphery. In some cases this may simply reflect community regeneration, where old leadership makes way for new generations of contributors that step-up to take additional responsibilities (as has been observed in the context of open-source software development [52] and Wikipedia [49]). In fact, in mid-2011 Wikipedia developed a policy to retire inactive administrators[11]. In other cases, however, these 'downward' transitions are the result of heated confrontations between community members. For example, consider the case of editor #124324, who engaged in arguments with a few other editors, eventually became frustrated with his treatment (both directly and through the formal arbitration/incident handling process), and voluntarily gave up his administrator privileges. In the words of editor #124324:

"In response to the Committee's decision to declare finding ... "[#124324] ... has used his administrative tools while involved ...)", I've requested a desysopping"

"anybody who cares: my self-block wasn't some kind of "tantrum" or strategy; it was a genuine attempt to get the hell away"

In other cases, editors are forced to give up privileges, as in the case of the editor #82835, one of the most active Wikipedians. The case involves accusations of disruptive editing against the editor. The demotion decision centers around principles of collegiality and the use of automation tools. The community decided to revoke administrator privileges and ban editor #82835 for up to one year.

Interestingly, as with the upward transitions, we found that when core members (at Level 3 of the organizational chart) move down the organizational ladder they most often transition directly to become a regular member, rather than traversing through the intermediate Level 2.

---

[11] http://en.wikipedia.org/w/index.php?title=Wikipedia:INACTIVITY

To the best of our knowledge, demotions were not previously investigated in the context of production communities and no prior studies have investigated the departure – whether voluntary or compulsory – of key community members (aside from the well-known case of Wikipedia cofounder Sanger's exit of Wikipedia [53]). While such demotions are the exception, rather than the norm, exposing this pattern enriches our understanding of career paths within online production communities.

Another insight from our study is that the dynamics of the roles transitions are more complex and intricate than originally perceived: we found evidence for both horizontal and vertical transitions; both upward and downward movements; sequential progression through intermediate phases, as well as direct transitions that bypass intermediate levels. Preece and Shneiderman [48] have already postulated that role transition is not one-directional, and that users can move up and down the leadership ladder. However, the evidence presented here points to complex patterns of transitions, whereby participants switch between organizational roles (a pattern brought to the extreme by one editor whose career spanned six role transitions). For example, we observed repeated cases where those who directly progressed to higher-level administrative roles (i.e. Level 3) choose to later add intermediate functional roles (Level 2); a pattern not discussed in prior works. Not only do we expose previously unobserved transition patterns, our findings also offer empirically-validated weights (or probabilities) for the various role transitions.

In sum, our study builds on and extends prior knowledge in the area in several important directions. In recent years, researchers in the area have begun to recognize the need to move beyond the simplistic conceptualization of core-periphery. One line of research has focused on the process of promotion within Wikipedia. These studies tend to focus on a particular promotion decision, often to 'sysop': a single transition in the organizational chart and a single event in a user's trajectory. For example, [13, 17, 18] studied how activity patterns affect promotion to administrator position. While such studies are useful in shedding light on why and how people move to administrative positions, they tell us little about the organizational structure in its entirety; further, they do not capture contributors' full role transition trajectories. In addition, most of these studies view career path on a single dimension - promotion to position of power – not distinguishing between the functional roles played at each position. Because of the focus on promotion to higher-order positions (i.e. *Administrator*), they miss on the complex functional organization at the middle level of the organizational chart (i.e. Level 2). Another interesting line of research has been investigating the nature of leaders (and leadership) within online communities, moving beyond the single-dimension characterization of leadership (power or authority; i.e. core-periphery continuum). For example, [60] investigated editors leadership styles (transactional, person-focused, aversive, and legitimate leadership), and showed that these leadership behaviors are not restricted to those in formal administrator positions; and [61] developed a machine learning technique to automatically identify these leadership behaviors. [36] have analyzed leadership behaviors in an online community, argued that these behaviors should be distributed between many community members (as opposed to centralized leadership), and developed a tool to assist in leadership tasks. Our study extends these lines of investigation in four important ways: (a) our study focuses contributors' functional roles (as opposed to leadership behaviors), and brings into attention the intermediate "professional" level that is often absent in studies of leadership; (b) we study Wikipedia's entire functional organization (and transitions between functional roles), rather than focusing on a particular type of transition, thus enabling us to validate conceptualizations such as the Reader-to-Leader framework, and revealing unpredicted patterns (e.g. editors moving directly to administrative position and later assuming functional lower-level roles; (c) we demonstrate that despite the loose governance of Wikipedia and the fact that editing activities are open to all, the functional role seems to determine the type of activities editors engage in, such that each functional role is associated with a distinct activity profile; and (d) we shed new light on downward transitions out of the community's core, a topic that has been under-investigated.

Finally, we make a secondary methodological contribution regarding the usage of system logs as a tool for answering organizational research questions. Our experience has taught us that beyond the standard data cleaning that is necessary, there is a need to triangulate data from multiple sources to ensure its reliability. Some of the examples we have encountered include: automated bots not marked as such; incomplete registration date for early entrants; inconsistencies between two sources recording granting of access privileges; missing data on role granting in privileges logs (as evident by comments on editors' personal pages); and irrelevant privileges granting data in the logs that do not reflect role changes (i.e. a series of erroneous transactions and their corrections). We would like to offer a word of caution to researchers relying on log data in studying online communities, and to stress the importance of data triangulation.

Implications for design and management
The findings of this study have important practical implications for designers and administrators of online communities. These communities strive to sustain contributors' participation, engagement and commitment. An understanding of the paths contributors take could help to develop diverse "career paths" within the community, such that contributors with different skill sets and interests could find suitable avenues for channeling their energy. An implication would be to place more emphasis on early detection and encouragement of contributors who seem to

be suitable (as indicated by their editing profiles) for particular roles. Community owners could provide more structure – through design and communication – to career paths, such that prior to advancing to core administrative duties, contributors should first serve on functional roles such as *Border Patrol* or *Quality Assurance*. Notwithstanding the recommendation above for "professional" career paths, it may be useful to create – in parallel - direct paths to leadership. Results from our study suggest that this path may be attractive to many Wikipedians (as indicated by the large number of editors who bypass Level to roles on their progression to becoming an *Administrator*). A design implication would be to develop tools that track contributors' activities and offer to them career guidance, including suggestions for roles and functions that best match their profile.

Our findings are also relevant to managers in firms who explore new collaborative production strategies. For example, some technology companies participate in open source software development [51], whereas others have adopted the practices of the open source movement for their internal software development projects [50]. In a similar vein, many companies use wiki technology as a knowledge management tool [2], and in particular for developing a Wikipedia-like organizational encyclopedia [28], adopting (at least in part) the organic processes that typify wiki-based collaboration over the Internet. Insights from this study would be valuable to such firms, help structure the community of practice, organize work more effectively, design career paths, and ensure the community's sustainability.

CONCLUSION
Taking the results of this study as a whole, we make a case for investigating the functional organization of online production communities. It is interesting to note that only now researchers of online communities are turning their attention to the functional organization of production communities, given that this topic is at the core of the organizational literature [40]. Existing conceptualizations of emergent leadership within peer production have primarily focused on the dimension of authority (i.e. proximity to the community core) and prior studies have largely relied on contributors' activity for estimating the position on the periphery-core continuum. We, however, argue than an understanding of the formal functions in the community – as well as the way in which contributors traverse these functional roles – is paramount for explaining how online production communities organize their work.

Given that this is only a first attempt to empirically analyze and quantify the process of role transition along participants' career paths in online production communities, our study has several limitations which we hope to address in future research. First, in terms of method, building on [3] we defined organizational roles

within Wikipedia based on editors access privileges. This approach has advantages over the method used in earlier studies where one's position on the core-periphery continuum was determined based on a simple quantification of his activity [48][12]. Nonetheless, we propose that future studies employ a multi-method approach, determining one's role based on access privileges, activity count, and additional recognition one receives (e.g. in Wikipedia: 'barnstars'[13] , service awards[14] , and other personal awards[15] [30, 33]). Second, our analysis of Wikipedia's organizational structure has focused on human agents, whereas automated software bots, too, carry out important tasks [42]. We, thus, propose that future studies provide a more complete view of work organization by also analyzing the various tasks performed by bots. Third, we acknowledge that the role transition patterns reported in this study have been influenced by the process by which the Wikipedia community introduced access privileges. For example, in the early stages of Wikipedia development, administrators were responsible for many of the functional roles; however, after the introduction of Level 2 privileges (2008-2011) many of these functions were delegated to the *Technical Administration*, *Border Patrol*, *Quality Assurance* and *QA Technicians* roles. This implies that the nature of *Administrator* role (as well as patterns of transitions into this role) have changed over time, calling for a future research that would re-examine role transition patterns once roles stabilize. Fourth, in order to increase the generalizability of our findings, in future research we plan to expand the analysis and perform a larger-scale evaluation (more participants, longer time period). In particular, the sample we employed excluded non-member contributors and late joiners (becoming active after 2007), and we call for future research that would repeat our study on a more comprehensive sample. Lastly, in future research we plan to extend our investigation to other online communities, in particular those that differ from Wikipedia in terms of the motivational drivers, the governance structure, or the enabling IT platform. Moving the investigation beyond Wikipedia is of special importance, given that the distinct role-transition patterns we have identified – and in particular the importance of the mid-level functional layer – may be relevant to only production communities that have grown beyond a certain size. We call for future research that would explore the phase in a community's life that call for the establishment of this kind of functional layer.

---

[12] Preece and Shneiderman note that "…these metrics [employed in previous studies] only capture activity. The development of more potent measures of efficacy in achieving personal and community goals would be a big breakthrough".

[13] http://en.wikipedia.org/wiki/Wikipedia:Barnstars
[14] http://en.wikipedia.org/wiki/Wikipedia:Service_awards
[15] http://en.wikipedia.org/wiki/Wikipedia:Awards

REFERENCES

[1] Amrit, C. and van Hillegersberg, J., "Exploring the impact of socio-technical core-periphery structures in open source software development," *Journal of Information Technology,* 25, 216-229, 2010.

[2] Arazy, O. and Gellatly, I., "Corporate Wikis: The Effects of Owners' Motivation and Behavior on Group Members' Engagement,," *Journal of Management Information Systems,* 29, 87-116, 2013.

[3] Arazy, O., Nov, O., and Ortega, F., "The [Wikipedia] World is Not Flat: on the organizational structure of online production communities," in *The 22nd European Conference on Information Systems (ECIS'2014)*, Tel-Aviv, Israel, 2014.

[4] Arazy, O., Nov, O., Patterson, R., and Yeo, L., "Information quality in Wikipedia: The effects of group composition and task conflict," *Journal of Management Information Systems,* 27, 71 - 98, 2011.

[5] Arazy, O., Stroulia, E., Ruecker, S., Arias, C., Fiorentino, C., Ganev, V*., et al.*, "Recognizing contributions in wikis: Authorship categories, algorithms, and visualizations," *Journal of the American Society for Information Science and Technology,* 61, 1166-1179, 2010.

[6] Arazy, O., Yeo, L., and Nov, O., "Stay on the Wikipedia Task: when task-related disagreements slip into personal and procedural conflicts," *Journal of the American Society for Information Science and Technology (JASIST),* 64, 1634–1648, 2013.

[7] Arthur, M. B., "The boundaryless career," *Journal of Organizational Behavior,* 15, 295-306, 1994.

[8] Baker, W. E. and Faulkner, R. R., "Role as resource in the Hollywood film industry," *American Journal of Sociology,* 97, 279-309, 1991.

[9] Benkler, Y., *The wealth of networks: How social production transforms markets and freedom*: Yale Univ Press, 2006.

[10] Blau, J. R., Coser, R. L., and Goodman, N., *Social Roles & Social Institutions*. New Jersey, USA: Transaction Publishers, 1995.

[11] Borgatti, S. P. and Everett, M. G., "Models of core/periphery structures," *Social networks,* 21, 375-395, 2000.

[12] Bryant, S., Forte, A., and Bruckman, A., "Becoming Wikipedian: transformation of participation in a collaborative online encyclopedia," in *International ACM SIGGROUP Conference on Supporting Group Work*, 2005, 1-10.

[13] Burke, M. and Kraut, R., "Mopping up: modeling wikipedia promotion decisions," in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW'08)*, San Diego, California, USA, 2008, 27-36.

[14] Butler, B., Joyce, E., and Pike, J., "Don't look now, but we've created a bureaucracy: the nature and roles of policies and rules in wikipedia," in *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI'08)*, Florence, Italy, 2008, 1101-1110.

[15] Butler, B., Sproull, L., Kiesler, S., and Kraut, R., "Community effort in online groups: Who does the work and why?," in *Leadership at a Distance: Research in Technologically-Supported Work*, Ed. Hillsdale, NJ, USA: L. Erlbaum Associates Inc., 2002, 346-362.

[16] Butler, B., Sproull, S., Kiesler, S., and Kraut, R. E., "Community effort in online communities: who does the work and why," in *Leadership at a Distance Research in Technologically-Supported Work*, S. Weisband, Ed., ed London UK: Lawrence Erlbaum Associates, 2007, 171-194.

[17] Collier, B., Burke, M., Kittur, N., and Kraut, R., "Retrospective versus prospective evidence for promotion: The case of Wikipedia," in *2008 annual meeting of the Academy of Management*, 2008.

[18] Collier, B., Burke, M., Kittur, N., and Kraut, R., "Promoting Good Management: Governance, Promotion, and Leadership in Open Collaboration Communities," presented at the *International Conference on Information Systems*, St. Louis, USA, 2010.

[19] Collier, B. and Kraut, R., "Leading the Collective: Social Capital and the Development of Leaders in Core-Periphery Organizations," presented at the *Collective Intelligence*, MIT, Boston, MA, USA, 2012.

[20] Crowston, K., Wei, K., Li, Q., and Howison, J., "Core and Periphery in Free/Libre and Open Source Software Team Communications," *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS'2006)*.

[21] Dahlander, L. and O'Mahony, S., "Progressing to the center: Coordinating project work," *Organization science,* 22, 961-979, 2011.

[22] Denning, P. J. and Yaholkovsky, P., "Getting to 'We': Solidarity, not software, generates collaboration," *Communications of the ACM,* 51, 19-24, 2008.

[23] Forte, A. and Lampe, C., "Defining, Understanding, and Supporting Open Collaboration: Lessons From the Literature," *American Behavioral Scientist,* 57, 535-547, 2013.

[24] Forte, A., Larco, V., and Bruckman, A., "Decentralization in Wikipedia governance," *Journal of Management Information Systems,* 26, 49-72, 2009.

[25] Geiger, R. S., "Are computers merely supporting cooperative work: towards an ethnography of bot development," in *Proceedings of the 16th Conference on Computer Supported Cooperative Work (CSCW'13)*, San Antonio, Texas, USA, 2013, 51-56.

[26] Halfaker, A., Geiger, R. S., Morgan, J. T., and Riedl, J., "The rise and decline of an open collaboration system: How Wikipedia's reaction to popularity is causing its decline," *American Behavioral Scientist,* 57(5), 664-688, 2013.

[27] Ibarra, H., *Working identity: Unconventional strategies for reinventing your career*. Harvard Business Press, 2003.

[28] Jang, S., Arazy, O., Nov, O., and Brainin, E., "'Crowding Out' in Corporate Wikis: the Effects of Job Responsibility and Motivation on Participation," in *6th Mediterranean Conference on Information Systems (MCIS'2011)*, Cyprus, 2010.

[29] Kittur, A., Suh, B., and Chi, E. H., "What's in Wikipedia? Mapping topics and conflict using socially annotated category structure," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'09)*, Boston, MA, 2009.

[30] Kriplean, T., Beschastnikh, I., and McDonald, D. W., "Articulations of wikiwork: uncovering valued work in wikipedia through barnstars," in *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work (CSCW'08)*, San Diego, CA, USA, 2008, 47-56.

[31] Lave, J. and Wenger, E., *Situated Learning: Legitimate Peripheral Participation*. Cambridge, UK: Cambridge University Press, 1991.

[32] Lerner, J., "Role assignments," in *Network analysis*, U. Brandes and T. Erlebach, Eds., Springer, 2005, 216-252.

[33] Leskovec, J., Huttenlocher, D., and Kleinberg, J., "Governance in Social Media: A case study of the Wikipedia promotion process," in *AAAI International Conference on Weblogs and Social Media (ICWSM'10)*, 2010.

[34] Liu, J. and Ram, S., "Who does what: Collaboration patterns in the Wikipedia and their impact on article quality," *ACM Trans. Manage. Inf. Syst.,* 2, 1-23, 2011.

[35] Long, Y. and Siau, K., "Social network structures in open source software development teams," *Journal of Database Management,* 18, 25-40, 2007.

[36] Luther, K., Fiesler, C., and Bruckman, A., "Redistributing leadership in online creative collaboration," in *Proceedings of the 2013 ACM Conference on Computer Supported Cooperative Work (CSCW'13)*, 2013, 1007-1022.

[37] Merton, R. K., *Social theory and social structure*. New York, NY, USA: Simon and Schuster, 1968.

[38] Meyer, J. W. and Rowan, B., "Institutionalized organizations: Formal structure as myth and ceremony," *American Journal of Sociology,* 83, 340-363, 1977.

[39] Miles, R. E., Snow, C. C., Meyer, A. D., and Coleman, H. J., "Organizational strategy, structure, and process," *Academy of Management Review,* 3, 546-562, 1978.

[40] Mintzberg, H., "The structuring of organizations: A synthesis of the research," *University of Illinois at Urbana-Champaign's Academy for Entrepreneurial Leadership Historical Research Reference in Entrepreneurship,* 1979.

[41] Morell, M. F., "Governance of online creation communities: Provision of infrastructure for the building of digital commons," PhD, European University Institute, 2010.

[42] Niederer, S. and Van Dijck, J., "Wisdom of the crowd or technicity of content? Wikipedia as a sociotechnical system," *New Media & Society,* 12, 1368-1387, 2010.

[43] O'Mahony, S. and Ferraro, F., "The emergence of governance in an open source community," *Academy of Management Journal,* 50, 1079-1106, 2007.

[44] Oreg, S. and Nov, O., "Exploring motivations for contributing to open source initiatives: The roles of contribution context and personal values," *Computers in Human Behavior,* 24, 2055-2073, 2008.

[45] Ortega, F., Izquierdo-Cortazar, D., Gonzalez-Barahona, J. M., and Robles, G., "On the analysis of contributions from privileged users in virtual open communities," in *Proceedings of the 42nd Hawaii International Conference on System Sciences (HICSS'2009)* Waikoloa, Big Island, HI, USA, 2009, 1-10.

[46] Panciera, K., Halfaker, A., and Terveen, L., "Wikipedians are born, not made: A study of power editors on Wikipedia," presented at the *GROUP, 2009.*

[47] Preece, J., Nonnecke, B., and Andrews, D., "The top five reasons for lurking: improving community experiences for everyone," *Computers in human behavior,* 20, 201-223, 2004.

[48] Preece, J. and Shneiderman, B., "The Reader-to-Leader Framework: Motivating technology-mediated social participation," *AIS Transactions on Human-Computer Interaction,* 1, 13-32, 2009.

[49] Ransbotham, S. and Kane, G. C., "Membership Turnover and Collaboration Success in Online Communities," *MIS Quarterly,* 35, 613-627, 2011.

[50] Riehle, D., Ellenberger, J., Menahem, T., Mikhailovski, B., Natchetoi, Y., *et al.*, "Open collaboration within corporations using software forges," *IEEE Software,* 26, 52-58, 2009.

[51] Roberts, J., Hann, I., and Slaughter, S., "Understanding the Motivations, Participation, and Performance of Open Source Software Developers," *Management Science,* 52, 984-999, 2006.

[52] Robles, G., Gonzalez-Barahona, J. M., and Herraiz, I., "Evolution of the core team of developers in libre software projects," in *Proceedings of the 6th IEEE International Working Conference on Mining Software Repositories (MSR'09)*, Vancouver, BC, Canada, 2009, 167-170.

[53] Sanger, L., "The early history of Nupedia and Wikipedia: a memoir," *Open sources,* 2, 307-338, 2005.

[54] Shaw, A. and Hill, B. M., "Laboratories of Oligarchy? How the Iron Law Extends to Peer Production," *Journal of Communication,* 64, 215-238, 2014.

[55] Stvilia, B., Twidale, M., Smith, L., and Gasser, L., "Information quality work organization in Wikipedia," *Journal of the American Society for Information Science and Technology,* 59, 983-1001, 2008.

[56] Von Hippel, E. and Von Krogh, G., "Open source software and the" private-collective" innovation model," *Organization Science,* 14, 209-223, 2003.

[57] Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., *et al.*, "Finding social roles in Wikipedia," in *Proceedings of the 2011 iConference*, Seattle, WA, USA, 2011, 122-129.

[58] Welser, H. T., Gleave, E., Fisher, D., and Smith, M., "Visualizing the signatures of social roles in online discussion groups," *Journal of Social Structure,* 8, 1-32, 2007.

[59] Wenger, E., *Communities of practice: Learning, meaning, and identity*. Cambridge, UK: Cambridge university press, 1999.

[60] Zhu, H., Kraut, R., and Kittur, A., "Effectiveness of shared leadership in online communities," in *Proceedings of the 2012 ACM Conference on Computer Supported Cooperative Work (CSCW'12)* , 407-416.

[61] Zhu, H., Kraut, R. E., Wang, Y.-C., and Kittur, A., "Identifying shared leadership in Wikipedia," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI'11)*, 3431-3434.