

Accepted for publication

Arazy O., Lifshitz-Assaf H., and Balila A. 2018. The Neither a Bazaar nor a Cathedral: the Interplay between Structure and Agency in Wikipedia's Role System, Journal of the American Society for Information Science and Technology (JASIST), forthcoming.

## **Neither a Bazaar nor a Cathedral: the Interplay between Structure and Agency in Wikipedia's Role System**

**Ofer Arazy**

Department of Information Systems

The University of Haifa

Phone: 972-4-8288504

E-mail: [oarazy@is.haifa.ac.il](mailto:oarazy@is.haifa.ac.il) (primary contact)

**Hila Lifshitz-Assaf**

Stern School of Business

New York University

E-mail: [hlassaf@stern.nyu.edu](mailto:hlassaf@stern.nyu.edu)

**Adam Balila**

Department of Information Systems

The University of Haifa

E-mail: [adambalila@hotmail.com](mailto:adambalila@hotmail.com)

## **Neither a Bazaar nor a Cathedral: the Interplay between Structure and Agency in Wikipedia's Role System**

### **Abstract**

Roles provide a key coordination mechanism in peer-production. Whereas one stream in the literature has focused on the structural responsibilities associated with roles, the another has stressed the emergent nature of work. To date, these streams have proceeded largely in parallel. In seeking to enhance our understanding of the tension between structure and agency in peer-production, we investigate the interplay between structural and emergent roles. Our study explored the breadth of structural roles in Wikipedia (English version) and their linkage to various forms of activities. Our analyses show that despite the latitude in selecting their mode of participation, participants' structural and emergent roles are tightly coupled. Our discussion highlights that: (I) participants often stay close to the "production ground floor" despite the assignment into structural roles; and (II) there are typical modifications in activity patterns associated with role-assignment, namely: functional specialization, multi-specialization, de-functionalization, changes in communication patterns, management of identity, and role definition. We contribute to theory of coordination and roles, as well as provide some practical implications.

**Keywords:** Peer-production; online community; Wikipedia; structure; agency; structural role; emergent role.

# **Neither a Bazaar nor a Cathedral: the Interplay between Structure and Agency in Wikipedia's Role System**

## Introduction

Online production communities, such as Wikipedia and open-source software development projects, represent an alternative form of organizing that is particularly suited for the co-creation of knowledge-based products (Benkler 2006; Shirky 2009). Volunteers in peer-production are responsible for project tasks (including both production and administrative duties). As participants become more involved in the project and gain the community's trust, they gradually move from the periphery to the community core, gaining access to more sensitive and influential decisions (Amrit and Van Hillegersberg 2010). The community works to attract and retain volunteers with diverse interests and skill-sets, and is responsible for governing the production project.

In his seminal essay, Raymond (1999) depicts traditional command-and-control organizations as cathedrals, managed through formal structures; in stark contrast, the self-organized community-based model that is common in open source software development is viewed as a bazaar (suggesting self-organization, where participants exert agency is determining their mode of participation). In reality, the participation trajectories taken by contributors to peer-production are the result of the interplay between structure and agency: online production communities, through structural (or governance) mechanisms (such as policies, and access privileges) channel the potential for participants' agency, thereby influencing their actions and behavior (Auray 2012; Butler et al. 2008; Forte et al. 2009; Schroeder and Wagner 2012). Conversely, participants apply their agency and take an active part in shaping the community's

structure (for example in co-developing policies and collaboratively defining roles' responsibilities) (Aaltonen and Lanzara 2015). Our focus in this study is on one particular angle of the structure/agency tension: contributors' roles within the community.

Roles represent a primary mechanism for coordinating work. Structural roles allow access to resources and decision making and reflect the participant's position in the community. Mature online production communities often develop intricate role schemes with specialized responsibilities, sometimes implemented through access privileges, which entail rights to participate in: production activities, community decision-making, and the development and enforcement of norms, policies and rules. Wikipedia, for example, has developed a complex organizational structure that includes over twenty different access privileges (i.e. structural roles), such as privileges associated with quality assurance work and with combatting deliberate manipulations attempts (Arazy et al. 2014).

While structural roles constrain to some extent participants' engagement, contributors to peer production still have substantial latitude in determining how, when, and what to work on (Benkler 2006; Lakhani and Panetta 2007; Puranam and Hakonsson 2015; von Krogh et al. 2003). Recent years have seen increased scholarly interest in the nature of emergent work and in particular in the processes by which prototypical activity patterns organically emerge in knowledge co-production (Faraj et al. 2011; Kane et al. 2014). Staying with the Wikipedia example, despite the structural role system, production activity is to a large extent free from workflow constraints, such that contributors are autonomous to choose whether they want to make small changes such as fix typos, add new content, delete content, or reorganize others' prior postings (Arazy et al. 2011). Furthermore, not only that production processes are open to all, participation in coordination and policy work, too, is largely unconditional on access

privileges (Kane 2011). In line with recent studies, we use the term “emergent roles” to refer to prototypical activity patterns that organically emerge from participants’ activity choices (Arazy et al. 2016; Arazy et al. 2017).

The structural perception of roles highlights the importance of organizational design, while the emergent view of roles stresses contributors’ agency to enact particular actions at the moment. The organizational literature (Bechky 2006) suggests that structural roles and their enactment are interwoven. However, to date, research investigating participants’ emergent role behavior in peer-production has proceeded largely independent of scholarly work on communities’ structural role system. Conversely, conceptual frameworks regarding emergent role behavior (Arazy et al. 2016) have been divorced from models of the transition between structural roles (e.g. the Reader-to-Leader framework; (Preece and Shneiderman 2009)).

To investigate the tension between emergent and structural roles in peer-production, we wish to understand the extent to which the assignment into a structural role constrains behavior (i.e. leads to a change in the contributor’s self-enacted activity profile). When structural mechanisms are powerful – as in traditional command-and-control organizations - we would expect sharp behavioral changes to accompany role transitions. In contrast, in fluid setting with weak structural constraints, role transitions events are likely to have little impact on behavior patterns. We perceive peer-production initiatives as organizations that blend limited structure with ample opportunities for contributors to express their agency, and thus conjecture that transitions into structural roles will be associated with moderate changes in activity patterns.

More specifically, our objective is to try and identify common patterns in behavioral changes that are associated with these structural role-assignment events. If structural role-transitions are to be viewed as lateral movements – as proposed by Dahlander and O'Mahony

(2011) – we are to expect a shedding of one emergent role and taking on a different activity profile. If, on the other hand, we are to view the role structure as a hierarchical power relationship (Arazy et al. 2014), a promotion into a higher-level role entails assuming additional responsibilities and thus is likely to be associated with adding of a novel dimension to the existing activity profile. Our investigation has the potential to yield important insights about the way in which structure and agency interplay in peer-production participation trajectories, inform research on online knowledge collaboration, and offer guidelines to custodians of online communities.

The setting for our empirical investigation is the English Wikipedia. Wikipedia was able to recruit thousands of volunteers who contribute their knowledge and develop extensive governance mechanisms for coordinating the collaborative authoring process (Aaltonen and Lanzara 2015; Arazy et al. 2013; Forte et al. 2009; Ransbotham and Kane 2011). This setting is particularly suitable for investigating the tension between structural and emergent of roles: on one hand Wikipedia has developed an extensive access privilege system (Arazy et al. 2015), while on the other hand the various forms of participation within Wikipedia are largely open to all, regardless of their access privileges.

In what follows, we provide an overview of related work on roles in organizational theory and in peer-production, highlighting the gap this study tries to address. We then proceed to describe our research methodology and the study's findings. We conclude by discussing the implications of our findings – to both theory and practice – and by pointing to a few possible future research directions.

## Related Work

### *Theory of Roles in Organizations*

For decades, the tension between social structure and human agency has attracted the attention of social scientists. Agency refers to an actor's ability to make a choice to act, whereas social structure is the patterned social arrangements in society that are both emergent from and determinant of individuals' actions (Fuchs 2001). In organizational settings, the same tension between agency and structure is manifested: while the structure of organizational design constrains members' actions, individuals within an organization still exert free choice in determining how to perform their job.

Roles are the basic units of socialization; they provide a meaning system by encapsulating the social context, history of actions, structures of interaction, and the attributes people bring to the interaction (Blau et al. 1995). Two streams of role theory, the structural and the interactionist, provide different approaches to studying roles in organizations. From a structural perspective, a role is defined in terms of the expected behaviors that arise from norms and others' demands (Hughes 1958; Thomas and Biddle 1966). Structural role theory focuses on the ways in which role expectations constrain and circumscribe individuals' behavior. Interactionist frameworks to understanding role behavior (Blumer 1986; Mead 1934), on the other hand, focus on the ways individuals can dynamically construct and reconstruct social arrangements through the behaviors they enact. Put together, role structures provide a general framework that constrains and enables individuals' action, yet individuals have much leeway in enacting their own roles (Turner 1978).

Studies of self-organizing teams demonstrate how structural roles and enactment are interwoven. Brown and Eisenhardt (1997), for example, have studied projects teams in an engineering firm and found that neither mechanistic nor organic structures produced optimal

results; instead “...successful multiple-product innovation blends limited structure around responsibilities and priorities with extensive communication and design freedom to create improvisation within current projects” (P. 1). Other examples include high-response teams (e.g. firefighting, emergency-response), who rely on the role structure to coordinate complex and interdependent activities under extreme time constraints, yet at the same time organize their behavior by adjusting in relation to one another (Bigley and Roberts 2001). In the film projects studied by Bechky (2006), for instance, the generalized role structure and the role enactments were found to mutually support one another, establishing a means for almost immediate coordination on each new project.

### *Roles in Online Production Communities*

Contributors in online production communities play different roles, with the majority of participants, who are not very active and often do not assume a structural role, at the community’s periphery, and a small minority who take on additional responsibilities and privileges constituting the core. In reviewing the literature on roles on peer production, we attend to both streams in the literature: studies that have emphasized the structural perspective and works that have adopted the emergent lens.

**Structural Roles in Online Production Communities.** Prior works that are grounded in the structural view of roles stress the importance of access to resources and decision-making processes. Studies of Wikipedia adopting the structural perspective provide a description of the formal role structure, by building on the community’s extensive access privileges system (Arazy et al. 2014; Forte et al. 2009; Stvilia et al. 2008). The majority of these prior studies have primarily been interested in a particular Wikipedia mechanism (e.g. (Forte et al. 2009)). Other



relevant works in this area have investigated the functions and responsibilities associated with structural roles (Arazy et al. 2015; Butler et al. 2007), and have studied the process by which individuals are promoted from one structural role to another (Burke and Kraut 2008; Collier et al. 2010).

**Emergent Roles in Online Production Communities.** Recent conceptualizations provide an alternative view and depict online communities as flexible entities, governed through emergent coordination mechanisms, where participants' roles are enacted at the moment (Arazy et al. 2016; Faraj et al. 2011; Kane et al. 2014). Faraj et al. (2011) argue that the absence of deep social relationships and the lack of traditional structural mechanisms, free the collaboration from concerns of social conventions, ownership, and hierarchy. Few prior works have empirically explored the nature of these emergent roles. Empirical studies in the area have profiled contributors' activities, and then clustered them to identify the "bundles" of activities that represent their prototypical activity patterns (Arazy et al. 2016; Arazy et al. 2017; Liu and Ram 2011; Welser et al. 2011; Yang et al. 2016).

#### *Relationship between Structural and Emergent Roles in Online Production Communities*

There are only a handful of studies that have investigated the relationship between contributors' emergent roles (as reflected in their activity profiles) and structural positions within the online community. Dahlander and O'Mahony (2011) investigated the early years in the GNOME open source software development community. They analyzed the activity profiles (in terms of: "technical contributions", "technical communications", and "coordination work") leading to and resulting from being accepted into two privileged groups (foundation member and board director). They found that "project members contributed significantly more effort to the

project after progressing to lateral authority roles, most notably with respect to coordination work” (p. 962). A few studies have investigated promotion decisions in Wikipedia, focusing on the criteria employed by the community in accepting (or rejecting) a candidate’s nomination (Burke and Kraut 2008; Collier et al. 2008). More relevant to our investigation are studies that considered the tension between contributors’ agency and the community’s structural role system. These studies have tended to focus on the activity performed by contributors once already in a structural position, often investigating the differences in activity profiles between different types of roles in Wikipedia (Collier et al. 2010; Zhu et al. 2011) or in open source software projects (von Krogh et al. 2003).

### *Gaps in the Current Literature*

The question of how the structural role held by a contributor affects the emergent role she volunteers to play remains unanswered. The organizational literature, too, provides little insight in this regard. In traditional command-and-control organizations employee activity is largely determined by role prescription, and thus predicts behavior to substantially change with the assignment into a new role. In stark contrast, if we accept the view that online production communities are highly fluid and structure-free organizations (Faraj et al. 2011), then changes in structural positions are likely to have little effect on contributors’ activity profiles.

Furthermore, the few studies showing a linkage between structural roles and activity patterns, disagree on the *nature* of structural role transitions. One strand in the literature has conceptualized role transitions in online communities as lateral movements (Dahlander and O’Mahony 2011; Preece and Shneiderman 2009), thus implying that upon a transition from one structural position to the other a participant would shed his current activity profile and assume a

new behavioral pattern. Another strand has made the case for power relations between structural roles, whereby one role subsumes a lower-level role (Arazy et al. 2014); if this conceptualization is accurate, then when moving up the role hierarchy, a participant is expected to expand his activity profile and complement his activity profile with new forms of action.

The diverging views in previous research on roles, coupled with insufficient empirical validation, stress the need for a more nuanced understanding of the behavioral dynamics that is associated with the transition into structural roles. The goal of this study is, thus, to address these gaps and enhance our understanding regarding the relationship between Wikipedia's structure (in terms of the structural roles employed by a contributor) and contributors' agency (manifested in the profiles of activities they voluntarily enact). Our research questions concern the extent to which role-assignment events are associated with changes in contributors' activity patterns. In particular, we ask:

- *RQ1: How do Wikipedians' activity profiles change as a result of being assigned structural roles?*
  - *RQ1a: Which types of activities are most likely to change as a result of these role-assignment events?*
  - *RQ1b: The assignment into which structural roles is most likely to engender changes in activity profiles?*
- *RQ2: Across the various structural roles and types of activities, do behavioral changes resulting from role-assignment events follow distinctive patterns?*

## Research methodology

### *Sample and Recording of Detailed “Production” Activity*

The setting for our study is Wikipedia, which hosts many different projects, each defined as the co-production of a particular encyclopedic article (Halavais and Lackaff 2008). Given Wikipedia’s collaborative-authoring process, our sampling strategy centers on sets of contributor groups co-producing articles, and investigates all of members of each of the groups studied. Our starting point is the data produced by (Arazy et al. 2016), which includes 1000 articles from the January 2012 dump of the English Wikipedia and was generated through a double-stratified sampling procedure (based on articles’ maturity and topical domains). Our sample includes only participants that have made at least one activity to these 1000 articles: altogether 222,119 contributors making 721,806 activities (i.e. article revisions). An advantage of this data set is that editing activities have been annotated, employing a combination of manual and machine learning approaches (Arazy et al. 2016). The unit of analysis for the annotation was at the revision level, where each revision could contain multiple types of “editing work”.

Category
Create a New Article
Fix Typo(s)/Grammatical Errors
Rephrase Existing Text
Add Substantive New Content
Delete Substantive Content
Hyperlinks
Add Citations
Add Wiki Markup
Reorganize Existing Text
Add Vandalism
Delete Vandalism
Miscellaneous

Table 1. *The edit categories used to annotate the revisions in the dataset from (Arazy et al. 2016)*

Table 1 lists the twelve categories of wiki work that we have used. An evaluation on a test set showed that this approach yielded good results. Please refer to Arazy et al. (2016) for

additional details regarding the data set and the procedure for annotating the wiki editing activities.

### *Recording Activity across Wikipedia Namespaces*

We also recorded data regarding the activities of the 222,119 contributors outside the set of 1000 articles. Wikipedia maintains a series of wiki pages dedicated to various forms of engagement (Kittur and Kraut 2010). Overall, Wikipedia maintains 29 areas of content (i.e. namespaces): the Main namespace is dedicated to co-authoring encyclopedic entries, and the other namespaces for coordination, communication and policy. Here we focus on the 16 core namespaces, which are grouped in subject/talk pairs<sup>i</sup>. Table 2 lists the Wikipedia namespaces included in our analysis and briefly describes each. For each activity made to these namespaces by contributors in our samples, we recorded the time of activity and its size (in bytes).

### *Recording Participants Structural Roles*

Of the 222,119 contributors in the data set, 8,247 held special access privileges. 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 access privileges (Arazy et al. 2014; Butler et al. 2008). Whereas access privileges are an important mechanism for governing, by and large, the editing activities across the various Wikipedia namespaces are unconstrained by structural access privileges, with a few exceptions (for English Wikipedia privileges): (a) the creation of a new wiki page in any of the namespaces is not available to unregistered contributors<sup>ii</sup>; (b) The access privileges of *sysop*, *rollbacker* and *abusefilter* are

provided with special tools to more effectively revert pages back to their prior revisions<sup>iii</sup>; and (c) contributing to the Mediawiki namespace pages is only available to *sysop*<sup>iv</sup>.

As of April 2017, the Wikipedia community in the English language organized structural roles through over thirty distinct access privileges, where a contributor can hold multiple access privileges. We focus on human editors (excluding from this analysis software bots (Niederer and Van Dijck 2010)) and access privileges belonging to the English Wikipedia. Data regarding contributors' access privileges was retrieved through Wikipedia's API. We restricted the data gathering to only those users who held special access privileges, excluding anonymous contributors and non-privileged registered members. Also excluded from our analyses were access privileges that are no longer in use (e.g. related to the now inactive User Feedback feature: *afptest* and *afptest-hide*), as well as special access privileges given to external parties such as researchers, educators or developers of the wiki software (*epadmin*, *epcampus*, *epcoordinator*, *epinstructor*, *eponline*, *epstaff*, *developer*). Finally, we limit our analysis to access privileges associated with the English Wikipedia, excluding 'global' access privileges such as *founder*, *importer*, and *steward* (Arazy et al. 2014). Table 3 provides a list of the structural Wikipedia roles. To extract information regarding role-transition event (i.e. granting or revoking of access privileges) date, we used the WikiDAT tool (<http://glimmerphoenix.github.io/WikiDAT/>).

### *Linking Role-Transition Events to Activity Profiles*

In order to test whether the assignment into a structural role is associated with a change in activity profile, we decomposed each contributor's activity array, based on the access privileges she has received. That is, for each of her special access privileges (i.e. structural role), we created a list of the contributor's activity in the period before and after the date of structural role

assignment, covering edit activity across all Wikipedia namespaces<sup>v</sup>. Each of the activity lists was normalized, recording for each activity types the percent it represents out of the total activity for that list (instead of the number of edits). A similar list was created for representing a contributor's detailed editing activities within the set of 1000 articles.

### *Statistical Methods*

Focusing on the patterns that reoccur frequently in our sample, we proceeded to analyze the access privileges with at least 30 instantiations, as described in Table 2. For these access privileges and for each activity type independently, we compared contributors' profiles pre- and post- role-assignment events. This analysis was repeated for both global activities (i.e. activity type entails the particular namespace) and detailed production activities. For both the global and detailed activity, we repeated the analyses employing two alternative metrics: activity count and size (in bytes). To test the differences pre/post role-assignment, we performed a Mann-Whitney U tests (aka Mann-Whitney-Wilcoxon test), which is a non-parametric analysis that checks whether one sample tends to have larger values from another and is recommended for non-normal distributions. Our null hypotheses were that distributions of both pre- and post-event lists are equal; we used a 2-tailed test employing the statistical significance threshold of  $p < 0.05$ . Tests were repeated and performed independently for the various access-privileges activity-types pairs. When calculating statistical significance levels, we applied the False Discovery Rate (FDR) correction: a statistical method for counteracting the problem of multiple comparisons (i.e. familywise error rate), which focuses on the proportion of falsely significant tests (i.e. Type I errors in the null hypothesis testing) (Benjamini and Hochberg 1995). Our analysis of the extent to which structural role transitions are linked to changes in activity profiles was based on a quantitative summary of the Mann-Whitney U test results.

#	Namespace	Edit Count		
		Mean	Median	STD
0	Main	4587	1217	21582
1	Talk	713	94	4790
2	User	265	45	3727
3	User Talk	1015	177	9347
4	Wikipedia	524	90	2442
5	Wikipedia Talk	104	11	348
6	File	62	1	977
7	File Talk	12	0	296
8	MediaWiki	1	0	16
9	MediaWiki Talk	2	0	28
10	Template	143	9	1297
11	Template Talk	45	1	603
12	Help	1	0	17
13	Help Talk	1	0	12
14	Category	115	1	1856
15	Category Talk	76	0	2360

Table 2. Contributors' activity across the various Namespaces, averaging the number of revisions each contributor has made.

## Results

### *Descriptive Statistics*

Our first descriptive statistics focuses on the revisions contributors make across the various namespaces (in terms the count and size of these revisions). Contributors' activity data exhibited a power law distribution, where on average they make substantial changes to articles in some namespaces (the highest activity recorded in 'Main'), while the average activity in other namespaces is quite low (namely, the primary and talk pages for the namespaces of : 'File', 'Mediawiki', 'Help' and 'Category').

The usage of structural access privileges, too, showed an uneven distribution, where some access privileges represent a meaningful structural role (the primary roles include: *rollbacker*, *reviewer*, *autoreviewer*, and *sysop*), while other access privileges are assigned to few contributors (e.g.: *oversight*, *checkuser*, *bureaucrat*, *templateeditor*, *OTRS-member*, *massmessage-sender*). Table 3 describes each role in terms of the percentage of contributors receiving this role out of all special-privilege contributors in the sample. In total, the number of



role-assignment (or revoke) events in our sample was 16,475. The Wikipedians with special access privileges made relatively few role transitions, taking on average 3 roles (i.e. 2 transitions). Here, too, activity displayed a power law distribution.

Role	Description	Contributors	
		%	No.
<i>Reviewer</i>	Able to review edits to articles placed under ‘pending changes protection’	31.9%	4613
<i>rollbacker</i>	May revert revisions using the rollback feature	28.9%	4169
<i>autoreviewer</i>	Could be trusted not to submit inappropriate material	15.7%	2270
<i>Sysop (administrator)</i>	Have access to a number of tools and additional functions (e.g. page deletion, page protection, blocking).	11.0%	1586
<i>ipblock-exempt</i>	Were manually exempt from a block	3.5%	512
<i>abusefilter</i>	Can create, modify, and delete edit filters	2.3%	325
<i>accountcreator</i>	Can create accounts for others	2.1%	307
<i>filemover</i>	Allowed to rename files	2.0%	289
<i>confirmed</i>	Require some intervention in order to be approved	0.7%	103
<i>templateeditor</i>	Can edit protected pages and edit edit-notices.	0.5%	66
<i>oversight</i>	Can hide revisions of pages from all users	0.3%	49
<i>checkuser</i>	Can view the IP addresses used by other users	0.3%	48
<i>bureaucrat</i>	Allowed to perform certain actions on other users' accounts (e.g. add other users to the ' <i>sysop</i> ' and ' <i>bureaucrat</i> ' groups)	0.3%	37
<i>OTRS-member</i>	Answer most e-mail sent to Wikipedia	0.2%	35
<i>massmessage-sender</i>	May send messages to multiple users at once	0.2%	30

Table 3. *Wikipedia’s access privileges and the number of privileged contributors in our sample.*

### *Changes across Wikipedia Namespaces Associated with Role-Transition*

In describing the results of our analyses below, for brevity, we only present details regarding the edit count metric. When interpreting the results, however, we mention important findings related to both the edit count and size metrics.

Table 4 below report on the results for all access privileges and all namespaces, for the activity count (240 statistical tests), showing both the statistical significant (in asterisks) and directionality of effects (green represent an increase in activity after role assignment event; ‘-’, red represents a decrease). Overall, we see a significant change in contributors’ activity patterns that is associated with the assignment of a new access privilege. We have observed these changes

in activity patterns in terms of both the number of edits performed across the various namespaces and the scope of activity (for both types of analyses, 14 out of the 15 access privileges showed a significant difference in activity for at least one namespace). We observe that receiving an additional access privilege may be associated with an increased or reduced editing activity. Thus, when being assigned a special access privilege, contributors concentrate their efforts in certain namespaces, while investing less effort and reducing their editing activities in other namespaces. When considering the counts of activities, in 42 of the 240 tests we saw a statistically significant increase in activity, while in 42 cases activity was decreased. For example, we see a significant decrease in the editing of the ‘Main’ namespace after the assignment of a *sysop* access privilege ( $p < 0.001$ ). Similarly, the assignment of various access privileges (*rollbacker*, *reviewer*, *autoreviewer*, *confirmed*, *bureaucrat*) is associated with a significant decrease in ‘Talk’ editing activity.

#### *Changes in Detailed ‘Production’ Activity Associated with Role-Transition*

A complementing view of contributors’ activity pertains to the nature of edits performed in the co-production of the encyclopedic entries. Table 5 below reports on the results of comparing pre- and post-role-assignment-events activity profiles for the contributors (in terms of activity count). When turning our attention to the nature of activities within the ‘Main’ namespace, we note differences in pre- and post-event activities for 8 of the 12 access privileges in this analysis (37 out of 132 = 28% of access-privileges / edit-types combinations tested with a significant difference when using the edit count metric, and 71% when using edit size). In an attempt to analyze the linkage between ‘production’ activity profiles and functional roles through different lens, we associated each activity profile with one of the “emergent roles” identified in Arazy et

al. (2016) (based on the minimal distance between activity profile and the vector representing each emergent role) and studied changes in these emergent roles. Corroborating our earlier findings, our results show that in 27% of the cases role-assignment events are associated with a change in emergent role (i.e. in the remaining cases – 73% - contributors stick to their emergent role in spite of being assigned a new functional role).

Role	Main	Talk	User	User Talk	Wikipedia	Wikipedia Talk	File	File Talk	Mediawiki	Mediawiki Talk	Template	Template Talk	Help	Help Talk	Category	Category Talk
<i>rollbacker</i>	.040	<span style="color:red">-0.686</span> <span style="color:red">***</span>	<span style="color:red">-0.833</span> <span style="color:red">***</span>	<span style="color:green">1.077</span> <span style="color:green">**</span>	-0.001	.165	<span style="color:green">.054</span> <span style="color:green">*</span>	.009	<span style="color:red">-0.002</span> <span style="color:red">***</span>	-0.002	.123	.013	.002	.000	.036	<span style="color:red">-.003</span> <span style="color:red">***</span>
<i>reviewer</i>	<span style="color:green">.710</span> <span style="color:green">*</span>	<span style="color:red">-0.471</span> <span style="color:red">***</span>	<span style="color:green">.163</span> <span style="color:green">***</span>	<span style="color:red">-0.539</span> <span style="color:red">***</span>	<span style="color:red">-0.117</span> <span style="color:red">***</span>	<span style="color:green">.014</span> <span style="color:green">***</span>	<span style="color:green">.042</span> <span style="color:green">***</span>	<span style="color:green">.028</span> <span style="color:green">***</span>	<span style="color:red">-0.001</span> <span style="color:red">***</span>	-0.002	.211	<span style="color:red">-0.053</span> <span style="color:red">***</span>	<span style="color:green">.001</span> <span style="color:green">***</span>	.001	<span style="color:green">.005</span> <span style="color:green">***</span>	<span style="color:red">-.010</span> <span style="color:red">***</span>
<i>autoreviewer</i>	-0.826	<span style="color:red">-0.185</span> <span style="color:red">**</span>	<span style="color:green">.235</span> <span style="color:green">**</span>	.145	<span style="color:green">.118</span> <span style="color:green">**</span>	.048	<span style="color:green">.023</span> <span style="color:green">***</span>	<span style="color:green">.022</span> <span style="color:green">***</span>	<span style="color:red">-0.002</span> <span style="color:red">***</span>	-0.002	.264	<span style="color:red">-0.044</span> <span style="color:red">***</span>	<span style="color:green">.002</span> <span style="color:green">*</span>	.003	<span style="color:green">.115</span> <span style="color:green">***</span>	<span style="color:green">.086*</span> <span style="color:green">**</span>
<i>accountcreator</i>	-1.29	.177	.737	<span style="color:red">-2.78</span> <span style="color:red">**</span>	1.70	.822	.254	-0.024	<span style="color:green">.076</span> <span style="color:green">***</span>	.002	.202	.012	<span style="color:red">-0.007</span> <span style="color:red">**</span>	-0.001	.117	.005
<i>filemover</i>	.255	.082	.476	-0.832	-0.218	.225	.020	.051	-0.003	.001	-0.086	-0.193	-0.002	.004	.302	-0.082
<i>confirmed</i>	5.62	-1.02	<span style="color:red">-4.61</span> <span style="color:red">*</span>	4.75	-6.27	.598	.187	-0.414	-0.041	<span style="color:red">-0.077</span> <span style="color:red">*</span>	<span style="color:green">1.34*</span> <span style="color:green">**</span>	<span style="color:red">-0.218</span> <span style="color:red">***</span>	<span style="color:red">-0.004</span> <span style="color:red">*</span>	-0.003	<span style="color:green">.130</span> <span style="color:green">**</span>	<span style="color:green">.022</span> <span style="color:green">**</span>
<i>abusefilter</i>	-0.789	-0.299	-0.322	-1.03	.823	.919	.216	-0.012	.092	.044	.217	.062	<span style="color:green">.012</span> <span style="color:green">***</span>	<span style="color:green">.004</span> <span style="color:green">**</span>	.069	<span style="color:red">-0.008</span> <span style="color:red">*</span>
<i>sysop</i>	<span style="color:red">-3.52</span> <span style="color:red">***</span>	-0.275	<span style="color:green">.406</span> <span style="color:green">*</span>	<span style="color:green">2.43</span> <span style="color:green">***</span>	-0.201	.351	<span style="color:green">.172</span> <span style="color:green">*</span>	<span style="color:green">.025</span> <span style="color:green">***</span>	<span style="color:green">.078</span> <span style="color:green">***</span>	<span style="color:green">.038</span> <span style="color:green">***</span>	<span style="color:green">.480</span> <span style="color:green">***</span>	.046	.002	<span style="color:green">.003</span> <span style="color:green">***</span>	-0.054	.017
<i>bureaucrat</i>	-0.689	<span style="color:red">-1.29</span> <span style="color:red">*</span>	5.48	1.79	1.21	-0.196	-0.201	-0.046	.089	.038	.007	.214	-0.016	-0.007	-0.087	-0.019
<i>checkuser</i>	-0.014	-0.300	.100	.042	-0.065	.139	<span style="color:red">-0.278</span> <span style="color:red">*</span>	<span style="color:red">-0.011</span> <span style="color:red">*</span>	.094	<span style="color:green">.010</span> <span style="color:green">*</span>	.254	.159	<span style="color:red">-0.010</span> <span style="color:red">*</span>	<span style="color:red">-0.004</span> <span style="color:red">**</span>	-0.084	<span style="color:red">-0.032</span> <span style="color:red">*</span>
<i>oversight</i>	-1.71	-0.436	.443	2.77	-1.34	.333	<span style="color:red">-0.284</span> <span style="color:red">*</span>	<span style="color:red">-0.037</span> <span style="color:red">*</span>	.031	-0.018	.301	.099	<span style="color:red">-0.009</span> <span style="color:red">*</span>	<span style="color:red">-0.003</span> <span style="color:red">*</span>	-0.111	<span style="color:red">-0.027</span> <span style="color:red">*</span>
<i>ipblock-exempt</i>	.726	<span style="color:green">.371</span> <span style="color:green">*</span>	<span style="color:red">-0.054</span> <span style="color:red">***</span>	<span style="color:red">-2.66</span> <span style="color:red">***</span>	<span style="color:green">.450</span> <span style="color:green">*</span>	<span style="color:green">.297</span> <span style="color:green">*</span>	<span style="color:green">.417</span> <span style="color:green">***</span>	<span style="color:red">.008</span> <span style="color:red">***</span>	<span style="color:red">-0.006</span> <span style="color:red">***</span>	.001	.419	<span style="color:green">.040</span> <span style="color:green">***</span>	<span style="color:green">.004</span> <span style="color:green">*</span>	<span style="color:red">-0.001</span> <span style="color:red">*</span>	<span style="color:red">-0.039</span> <span style="color:red">***</span>	<span style="color:green">.038</span> <span style="color:green">***</span>
<i>templateeditor</i>	1.61	-1.15	-1.09	-1.63	-0.969	-1.19	.043	<span style="color:red">-0.075</span> <span style="color:red">*</span>	-0.049	.162	.933	.336	-0.023	.022	.878	.140
<i>massmessage-sender</i>	3.84	-1.70	3.39	<span style="color:red">-6.94</span> <span style="color:red">*</span>	1.61	-0.673	.141	<span style="color:red">-0.080</span> <span style="color:red">*</span>	.000	-0.010	.184	.056	-0.002	-0.009	.188	.006
<i>OTRS-member</i>	-2.43	-1.37	.344	-1.33	2.38	.572	1.05	.078	-0.018	<span style="color:red">-0.036</span> <span style="color:red">**</span>	.039	<span style="color:green">.178</span> <span style="color:green">**</span>	<span style="color:red">-0.005</span> <span style="color:red">***</span>	<span style="color:red">-0.011</span> <span style="color:red">**</span>	<span style="color:green">.023</span> <span style="color:green">**</span>	.525

Table 4. Differences in the namespace count of activities between pre- and post-role-assignment events (in percentages). ‘\*’ signifies  $p < 0.05$ ; ‘\*\*’  $p < 0.01$ ; and ‘\*\*\*’ for  $p < 0.001$  (adjusted with FDR). Red represents a significant decrease; green a significant increase.

As observed in the namespace analysis, while some roles assignment events are associated with a significant change in activity patterns, other role transitions do not entail such activity changes. The analysis of edit count showed that activity patterns are most likely to change with the assignment of the *rollbacker*, *reviewer* and *autoreviewer* access privileges (other access

privileges show very little differences in pre/post role-assignment activity). Also, the vast majority of access privileges recorded a significant change the size of their edits as a result of role-assignment events. Turning our attention to the particular types of production edits, we see that the production edits of ‘Add Content’ and ‘Miscellaneous’ showed the most differences, whereas the edits of ‘Add Vandalism’, ‘Rephrase Text’ and ‘Add Citations’ showed the least differences.

Role	Add Citations	Add Content	Add Wiki Markup	Delete Content	Fix Typo	Reorganize Text	Rephrase Text	Add Vandalism	Delete Vandalism	Hyperlinks	Miscellaneous
<i>rollbacker</i>	-301 ***	-2.23 ***	-4.94 ***	-452 ***	-1.40 ***	-.696 ***	-.441 ***	-.043	11.39 ***	-.479 ***	-.400 ***
<i>reviewer</i>	-.414 ***	-1.11 ***	-1.96 ***	.800	.154 ***	.073 ***	-.286 ***	-.118 *	3.34	-.549 ***	.071 **
<i>autoreviewer</i>	-.248 **	-1.19 ***	-1.25	.879	1.11	-.441 ***	-.155 ***	-.158 ***	1.57	-.525 ***	.396
<i>accountcreator</i>	.405	.391	1.73	1.35	2.71	-.460	-.297	-.182	-6.69	.458	.586
<i>filemover</i>	-.014	1.17	5.38	-.007	1.87	1.66	-1.47	.441	-9.50	-.011	.581
<i>confirmed</i>	1.65	3.19	-8.59	-4.36 ***	-7.08 **	-1.30	-2.72	.489	22.22	-1.35	-2.15 *
<i>abusefilter</i>	-.767	-1.90 **	.444	1.61	.550	.119	1.10	.658	-1.75	-.458 **	.393
<i>sysop</i>	-.019	-1.53 *	-1.51	1.14 ***	-.814	-.926	-.096	-.060	3.25 **	-1.04 **	1.25 ***
<i>bureaucrat</i>	-.507	-3.38	-2.01	-.404	8.57	-.589	-.098	-.296	-1.46	-1.89	.358
<i>checkuser</i>	.754	1.18	6.23	1.11	-1.65	2.42	1.47	.127	-10.20	.395	-2.42 *
<i>oversight</i>	-.464	-.183	6.50	-3.97 *	-.743	1.13	4.34	-.083	-4.39	.472	-2.61
<i>ipblock-exempt</i>	-.074	-2.39	1.42	-.211	-1.58	1.11	-.569	-.366	1.89	-.383	1.15

Table 5. Differences in types of ‘production’ activities (in terms of activity count) between pre- and post-role-assignment events (in percentages). ‘\*’ signifies  $p < 0.05$ ; ‘\*\*’  $p < 0.01$ ; and ‘\*\*\*’ for  $p < 0.001$  (adjusted with FDR). Red represents a significant decrease; green a significant increase.

### Patterns in the Interplay between Structural and Emergent Roles

A synthesis of the results presented in Tables 4 and 5 highlights key patterns in the interplay between participants’ structural roles and the changes in their activity patterns pre/post role-assignment events. Namely, we found that (a) participants often stay close to the

“production ground floor” and persist in (some of) their editing patterns even when assigned a functional role; and (b) when changing their activity pattern as a result of role-assignment, these behavioral changes follow few common patterns.

**Staying close to the “production ground floor”.** Our findings suggest that in over half of the role-transition events, participants do not change their behavior patterns substantially (i.e. the changes are statistically insignificant, after applying the FDR correction). This is evident in both the detailed activity patterns within the ‘Main’ namespace and more broadly in the activity across Wikipedia’s various namespaces. For the detailed co-production activity, 72% of the cases (95 out of the 132 access-privilege / edit-types combinations) registered insignificant change in edit count, and 29% recoded insignificant difference in activity profile in terms of edit size. For the activity across the namespaces, 65% and 61% of the cases (out of the 240 combinations) recorded insignificant change in edit count and size, respectively. In particular, contributors seem to maintain their original levels of participation in the Main namespace (i.e. co-authoring encyclopedic entries), regardless of assuming additional structural roles. We note that even changes in activity patterns that are statistically significant sometimes represent small effects (as evident in effect size), further stressing that contributors tend to stay close to the production floor.

**Changes in activity profile associated with role-assignment events.** Our findings show that for every one of Wikipedia’s structural access privileges changes between pre- and post-role-assignment-event were statistically significant in at least one form of activity. When synthesizing the findings regarding the changes pre/post role-assignment, some key patterns emerge, namely: *functional specialization, multi-specialization, de-functionalization, changes in communication patterns, management of identity, and role definition.*

The most notable pattern associated with transitions of functional role was **functional specialization**, whereby those assigned to the role increase in activity that is directly associated with the particular role, commonly reducing content-related activity. Consider a few examples. Certain roles – namely, *rollbacker* and *sysop* – are tasked with fighting vandalism and have access to the ‘revert’ tool allowing them to roll back a wiki page to its previous version; as expected, assignment into these roles entail a significant increase in ‘Delete Vandalism’ activity (and decrease in most other forms of activity). A second example considers the role of a *reviewer*, assigned with quality assurance; the assignment into this role is associated with a significant increase in the ‘Fix Typos’ and ‘Reorganize Text’ activities, directly linked to the *reviewer*’s structural responsibilities.

A second noticeable pattern was **multi-specialization**, which was typically linked to the transition into higher-level roles. Wikipedia’s system of access privileges is organized such that some roles are nested within others (the broader, or higher-level, roles encompassing the rights and responsibilities of the lower-level access privileges) (Arazy et al. 2014). Multi-specialization is evident when the transition of a participant into a higher-level role entails assuming the access privileges that are characteristic of multiple lower-level roles. For example, those assuming the *sysop* role commonly increase both the number and scope of editing activities that are typical of *rollback*, *reviewer* and *autoreviewer*; namely, increased activity around the management of files (File and File-Talk) and interface text (MediaWiki and MediaWiki-Talk namespaces).

Third, a common pattern associated with role-transition events was **de-functionalization**. That is, in cases when participants that held a preliminary access privilege were later promoted to a broader access privilege, this transition was accompanied by a decrease in the activity typical of the lower-level privileges. For example, transitions into the structural roles of *bureaucrat* and

*checkuser* were commonly accompanied with a decrease in the scope of vandalism removal activities (typical of lower-level access privileges such as *autoreviewer* or *rollbacker*).

Fourth, we found that role-transition events were linked to **changes in communication patterns** (as reflected in activity in the User and User-Talk namespaces). The exact nature of these changes differed between functional roles, in some cases increasing in the breadth of activity and in others decreasing in the scope of communication. For example, with the transition into the structural role of *sysop* the amount of communication grew (i.e. increase in edit count in the User-Talk namespace), whereas an opposite effect was observed for the roles of *accountcreator* and *reviewer*: namely a relative decrease in the number of User-Talk edits (for *reviewer*, the role-transition event was also accompanied with an increase in the average size of edits in this namespace, reflecting a more focused communication).

Fifth, we observed that upon assuming extra privileges, participants got more involved in the **management of their identity** within the community. Specifically, when assigned the *sysop* role, participants increased their activity in the User namespace. Finally, we noticed that after assuming structural roles, participants often get involved in the **definition of those structural roles**. For example, transitions into the structural roles of *rollbacker* and *reviewer*, were characterized with an increased activity in the particular areas of the Category-Talk namespace where the responsibilities and privileges of these roles are specified.

## Discussion

In seeking to enhance our understanding of coordination in peer-production, this study investigated the interplay between structural roles and activity patterns (i.e. emergent roles). Our investigation covered the complete history of 1000 Wikipedia articles, from their inception until

the end of 2014. Our research method relied on multiple data sources: detailed annotated data regarding contributors' 'production' activities (from (Arazy et al. 2016)), large scale activity edit, and detailed logs of role-transition events that were extracted using an external tool (WikiDAT). In order to disentangle contributors' position on the core-periphery from their activity profile, our analysis controlled for the quantity of activity, focusing solely on the nature of activity and defining a contributor's activity profile in the terms of the percentage of activity allocated to each of the activity categories. Triangulating results across the views of activity and metrics proved very powerful. Dependencies between structural roles and activity profiles that were found for the global namespace activity were often different from the dependencies observed for the detailed annotated production activity, and results recorded for using one metric were at times different than those recorded for another metric. Together, our various analyses highlighted the value of the comprehensive approach taken in the current study.

To the best of our knowledge, this is the first study to provide a comprehensive empirical evaluation of the linkage between structural roles and behavioral patterns in peer-production. Some studies have linked activity profiles to structural roles (Arazy et al. 2015; Collier et al. 2008; Collier et al. 2010; Zhu et al. 2011). By providing a broader view of structural roles and activity patterns we were able to identify meaningful patterns in the interplay between structure and agency in contributors' role-taking. Our findings reveal that the Wikipedia role system blends structure and agency in intricate ways. Despite the latitude in selecting their mode of involvement, participants' structural and emergent roles are tightly coupled, such that a contributor's decision to volunteer for additional formal responsibilities is intertwined with her decision to self-select various forms of activity.



Viewing peer-production as a highly structured form of organizing (resembling traditional command-and-control organizations) (Butler et al. 2008), we would expect role prescriptions to dictate actions and anticipate sharp behavioral changes to accompany role transitions. In contrast, when viewing online production communities as fluid and emergent organizations (Faraj et al. 2011), role transitions events are unlikely to have a substantial effect on behavior patterns. Our findings show that approximately fifty percent of the structural role transitions were associated with statistically significant changes in activity patterns, thus suggesting that neither perspective accurately captures Wikipedia's role system: the truth lies in between, and Wikipedia blends limited structure with ample opportunities for contributors to express their agency.

In particular we found that Wikipedians stay close to the "production ground floor" even when assuming additional access privileges. Prior to being assigned a structural role, contributors to Wikipedia are required to demonstrate substantial editing record. The fact that much of their profile remains unchanged with role-assignment suggests that these contributors continue with their regular activity (namely, editing the encyclopedic entries). Thus, privileged contributors continue to take part in the co-authoring of Wikipedia articles. The participation of privileged members in basic production activities serves two purposes. First, they lead others by example. Second, those holding a special access privilege serve as "on the ground traffic control", directing the activity of the masses of the non-privileged contributors.

Insights revealed from our analysis of the patterns of activity changes that are associated with structural roles assignment challenge extant conceptualizations in peer production. Some have argued that role transitions (for example, transitioning from a regular member status to serving on the board of an open-source community) are lateral movements (Dahlander and O'Mahony 2011). A competing view depicted roles in peer-productions as organized through

hierarchical power relations (Arazy et al. 2014). Interestingly, the patterns of behavioral changes identified in our study provide support for both perspectives. Namely, the patterns of *functional specialization* and *de-functionalization* suggest a lateral movement from one functional role to another; whereas the pattern of *multi-specialization* implies power relationships and a transition to a higher-level role. We, thus, infer that roles in Wikipedia are organized in a unique structure that combines hierarchical and lateral dimensions. We extend the prior conceptualizations on structural roles (e.g. the Reader-to-Leader framework).

Our study also has implications for the organizational literature conceptualization and understanding of roles. The literature currently suggests that roles could be understood through two primary lens: structural and interactionist (Turner 1986). Scholars in the area have recognized that these two perspectives complement one another: role structures provide a general framework that constrains and enables individuals' action, yet individuals have much leeway in enacting their own roles. Empirical studies of flexible and ad-hoc organizational forms provide rich descriptions of the interplay between role structure and the role enactments (Bechky 2006; Brown and Eisenhardt 1997), which inform our understanding of roles in online production communities.

Nonetheless, without entering the debate whether or not organizations such as Wikipedia represent a novel organizational form (Puranam et al. 2014), the distinguishing characteristics of roles in online communities should be considered. First, these communities differ in that the vast majority of participants do not hold any structural role, such that the few privileged-members are responsible for coordinating the work of many who do not have formal responsibilities. Second, in online production communities such as Wikipedia, roles are organically and collectively defined. Furthermore, promotion processes and the assignment into a structural role are often

based on the decision of a group of several privileged members (rather than decided by a direct supervisor). Finally, online collaborations are often distributed and very large in scope, including massive participation, in comparison to offline organizations. These characteristics create a situation that is different from the ones investigated in role theory so far. Whereas the organizational literature suggests that role behavior is largely linked to the enactment of structural roles, our findings suggest that role-alignment in online production communities is only partial, such that much of the enactment represents emergent behavior.

We contribute to the literature by arguing that each of these perspectives – the structural and emergent conceptualizations - captures only one aspect of the intricate interplay between structure and agency around the coordination of work in peer-production. We add to the literature on emergent work (Faraj et al. 2011) and roles (Arazy et al. 2016) by showing that emergent work is bounded – at least to some extent – by participants’ structural roles. We inform the literature on governance mechanisms in peer-production (Auray 2012; Forte et al. 2009; Schroeder and Wagner 2012) by showing that whereas structural roles provide a shared understanding of work responsibilities and organizational continuity, much of the organization of work is spontaneous, whereby privileged-members continue to emergently enact activities in Wikipedia’s co-production space independent of their assigned structural roles. What, then, is the organizational function of the loose coupling between roles and enacted behavior? We believe that the permeable role boundaries help privileged members easily and transiently fill-in for one another, thus contributing to organizational flexibility and robustness.

## Conclusion

An understanding of the way in which agency and structure interact is of special importance for online production communities, given that: (a) structure is often ill-defined in such novel forms of organizing; and (b) voluntary participation provides ample opportunities for contributors to exert agency. To date, theories in the area have overlooked the ways in which coordination practices rely simultaneously on structure and enactment. Specifically, whereas both structural and emergent roles are understood to help coordinate activity, little attention has been paid to the practices by which roles and enactment are interwoven. Our empirical investigation centered on Wikipedia and findings from this study reveal some intricate patterns in the interplay between structure and agency in Wikipedia's role system. Insights regarding how Wikipedia blends structure and agency in their role system may have implications for our understanding regarding the effective design of online production communities.

Our study has also some practical implications that extend beyond the contribution to the scholarly work in the area. Many coproduction communities struggle with the decisions of how much structure to impose on the coproduction process and whether or not to formalize roles. Our findings highlight that it is possible to coordinate the activity of masses of contributors through a rather "thin" and flexible role structure. More specifically, our findings could guide community custodians in their efforts to sustain contributors' participation, engagement and commitment. An understanding of participants' activity profiles and how those profiles are linked to participants' choices to take on additional structural roles and responsibilities could community custodians in developing "career paths", such that contributors with different skill sets and interests could find suitable avenues for channeling their energy. A managerial suggestion for community custodians would be to design roles to align with contributors work routines. Our

findings may also be relevant to managers in firms who explore new collaborative production strategies. For example, many companies use wiki technology as a knowledge management tool (Majchrzak et al. 2013; Wagner and Majchrzak 2007), in particular for developing a Wikipedia-like organizational encyclopedia and have adopted (at least partially) the organic processes that characterize wiki-based collaboration over the Internet. Organizations may differ from Wikipedia in terms of both the access privileges defined and the types of activities available for contributors; still the insights gained from this study regarding the balance between structure and agency could inform those managing corporate wikis.

Although yielding some interesting results, this study is not without limitations. First, few dependencies exist between structural access privileges and contributors' activity profiles some are irrelevant for this study (for example, non-registered members cannot create a new article, but those are excluded from this study), but other dependencies have a direct effect on the findings in this study<sup>vi</sup>. Second, our analysis provided a high-level view of each role, looking at the pre- and post-role-assignment event activity profiles. A more detailed analysis could “zoom-in” to consider the other access privileges held at that time. That is, we propose that future research would distinguish between, for example, progression into *sysop* when holding at that time the *filemover* role from promotions to *sysop* when not holding the additional role. Such an analysis has the potential to yield more nuanced understanding of the relationship between structural and emergent roles in Wikipedia. Third, our findings were based on a large-scale quantitative analysis, relying on Wikipedia's activity logs. While such an analysis is useful in capturing the breadth of the phenomenon and in providing a vivid portrait of the intricacies of roles-activities relationships within Wikipedia, other research methods – namely, qualitative research - are particularly suitable for addressing questions of “How?” and “Why?”. Hence, we

suggest that future research in the area will employ a mixed method approach, combining qualitative analysis of activity logs with qualitative methods to yield more refined explanations for the relationship between structural and emergent roles. Lastly, an additional limitation of this study concerns the generalizations of our results. Our analysis focused on the English Wikipedia, whereas other Wikipedias employ a somewhat different roles system (de Laat 2012). Furthermore, we recommend that future research moves beyond Wikipedia to other online production communities, in particular those that differ from Wikipedia in terms of their governance structure (and especially, the structural role system), and investigate roles and activities in those settings, paying attention to the effects of contextual factors.

In conclusion, the tension between structure and agency is paramount for understanding collaboration patterns within online production communities. The literature has often depicted community-based projects as “bazaars” (stressing participants’ agency), juxtaposing them with top-down “cathedral-style” structures that characterize traditional organizations (Raymond 1999). This paper presents a more nuanced view, focusing on the structure/agency tension with relation to roles. Demonstrating the way in which contributors’ choices to enact particular actions are intertwined with their formal responsibilities is of key importance, given that roles – both emergent and structural – represent one of the chief coordination mechanisms in organizations. Our study provides some preliminary results and we offer a first step towards a theory of permeable role boundaries; clearly, this area warrants further research.

## References

- Aaltonen, A., and Lanzara, G. F. 2015. "Building Governance Capability in Online Social Production: Insights from Wikipedia," *Organization Studies* (36:12), pp. 1649-1673.
- Amrit, C., and Van Hillegersberg, J. 2010. "Exploring the Impact of Socio-Technical Core-Periphery Structures in Open Source Software Development," *journal of information technology* (25:2), pp. 216-229.

- Arazy, O., Daxenberger, J., Lifshitz-Assaf, H., Nov, O., and Gurevych, I. 2016. "Turbulent Stability of Emergent Roles: The Dualistic Nature of Self-Organizing Knowledge Co-Production," *Information Systems Research* (27:4), pp. 792-812.
- Arazy, O., Lifshitz-Assaf, H., Nov, O., Daxenberger, J., Balestra, M., and Cheshire, C. 2017. "On the How and Why of Temporal Dynamics of Emergent Roles in Wikipedia," *20th Conference on Computer-Supported Cooperative Work and Social Computing (CSCW 2017)*, Portland, OR, USA.
- Arazy, O., Nov, O., and Ortega, F. 2014. "The [Wikipedia] World Is Not Flat: On the Organizational Structure of Online Production Communities," *The 22nd European Conference on Information Systems (ECIS'2014)*, Tel-Aviv, Israel.
- Arazy, O., Nov, O., Patterson, R., and Yeo, L. 2011. "Information Quality in Wikipedia: The Effects of Group Composition and Task Conflict," *Journal of Management Information Systems* (27:4), pp. 71-98.
- Arazy, O., Ortega, F., Nov, O., Yeo, L., and Balila, A. 2015. "Functional Roles and Career Paths in Wikipedia," *Computer Supported Cooperative Work (CSCW)*, Vancouver, BC, Canada: ACM.
- Arazy, O., Yeo, L., and Nov, O. 2013. "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)* (64:8), pp. 1634-1648.
- Auray, N. 2012. "Online Communities and Governance Mechanisms," in *Governance, Regulation and Powers on the Internet*, E. Brousseau, M. Marzouki and C. Méadel (eds.). Cambridge University Press.
- Bechky, B. A. 2006. "Gaffers, Gofers, and Grips: Role-Based Coordination in Temporary Organizations," *Organ. Sci.* (17:1), pp. 3-21.
- Benjamini, Y., and Hochberg, Y. 1995. "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing," *Journal of the royal statistical society. Series B (Methodological)*, pp. 289-300.
- Benkler, Y. 2006. *The Wealth of Networks: How Social Production Transforms Markets and Freedom*. Yale University Press.
- Bigley, G. A., and Roberts, K. H. 2001. "The Incident Command System: High-Reliability Organizing for Complex and Volatile Task Environments," *Academy of Management Journal* (44:6), pp. 1281-1299.
- Blau, J. R., Coser, R. L., and Goodman, N. 1995. *Social Roles & Social Institutions: Essays in Honor of Rose Laub Coser*. Transaction Publishers.
- Blumer, H. 1986. *Symbolic Interactionism: Perspective and Method*. Univ of California Press.
- Brown, S. L., and Eisenhardt, K. M. 1997. "The Art of Continuous Change: Linking Complexity Theory and Time-Paced Evolution in Relentlessly Shifting Organizations," *Administrative science quarterly*, pp. 1-34.
- Burke, M., and Kraut, R. 2008. "Mopping Up: Modeling Wikipedia Promotion Decisions," *Proceedings of the 2008 ACM conference on Computer supported cooperative work (CSCW)*, San Diego, California, USA: ACM, pp. 27-36.
- Butler, B., Joyce, E., and Pike, J. 2008. "Don't Look Now, but We've Created a Bureaucracy: The Nature and Roles of Policies and Rules in Wikipedia," *Proceedings of the CHI conference on human factors in computing systems*, Florence, Italy: ACM, pp. 1101-1110.
- Butler, B., Sproull, S., Kiesler, S., and Kraut, R. E. 2007. "Community Effort in Online Communities: Who Does the Work and Why," in *Leadership at a Distance Research in Technologicallysupported Work*, S. Weisband (ed.). London UK: Lawrence Erlbaum Associates, pp. 171-194.
- Collier, B., Burke, M., Kittur, N., and Kraut, R. 2008. "Retrospective Versus Prospective Evidence for Promotion: The Case of Wikipedia," *2008 annual meeting of the Academy of Management*.
- Collier, B., Burke, M., Kittur, N., and Kraut, R. 2010. "Promoting Good Management: Governance, Promotion, and Leadership in Open Collaboration Communities," in: *International Conference on Information Systems*. St. Louis, USA.
- Dahlander, L., and O'Mahony, S. 2011. "Progressing to the Center: Coordinating Project Work," *Organization science* (22:4), pp. 961-979.
- de Laat, P. B. 2012. "Coercion or Empowerment? Moderation of Content in Wikipedia as 'Essentially Contested' bureaucratic Rules," *Ethics and information technology* (14:2), pp. 123-135.
- Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. "Knowledge Collaboration in Online Communities," *Organization Science* (22:5), pp. 1224-1239.
- Forte, A., Larco, V., and Bruckman, A. 2009. "Decentralization in Wikipedia Governance," *Journal of Management Information Systems* (26:1), pp. 49-72.
- Fuchs, S. 2001. "Beyond Agency," *Sociological Theory* (19:1), pp. 24-40.
- Halavais, A., and Lackaff, D. 2008. "An Analysis of Topical Coverage of Wikipedia," *Journal of Computer Mediated Communication* (13:2), pp. 429-440.

- Hughes, E. C. 1958. *Men and Their Work*. Free Press.
- Kane, G. C. 2011. "A Multimethod Study of Information Quality in Wiki Collaboration," *ACM Transactions on Management Information Systems (TMIS)* (2:1), p. 4.
- Kane, G. C., Johnson, J., and Majchrzak, A. 2014. "Emergent Life Cycle: The Tension between Knowledge Change and Knowledge Retention in Open Online Coproduction Communities," *Management Science* (60:12), pp. 3026-3048.
- Kittur, A., and Kraut, R. E. 2010. "Beyond Wikipedia: Coordination and Conflict in Online Production Groups," *Proceedings of the 2010 ACM conference on Computer Supported Cooperative Work (CSCW'2010)*: ACM, pp. 215-224.
- Lakhani, K. R., and Panetta, J. A. 2007. "The Principles of Distributed Innovation," *Innovations* (2:3), pp. 97-112.
- Liu, J., and Ram, S. 2011. "Who Does What: Collaboration Patterns in the Wikipedia and Their Impact on Article Quality," *ACM Transactions on Management Information Systems* (2:2), pp. 1-23.
- Majchrzak, A., Wagner, C., and Yates, D. 2013. "The Impact of Shaping on Knowledge Reuse for Organizational Improvement with Wikis," *MIS Quarterly* (37:2), pp. 455-469.
- Mead, G. H. 1934. *Mind, Self and Society*. Chicago University of Chicago Press.
- Niederer, S., and Van Dijck, J. 2010. "Wisdom of the Crowd or Technicity of Content? Wikipedia as a Sociotechnical System," *New Media & Society* (12:8), pp. 1368-1387.
- Preece, J., and Shneiderman, B. 2009. "The Reader-to-Leader Framework: Motivating Technology-Mediated Social Participation," *AIS Transactions on Human-Computer Interaction* (1:1), pp. 13-32.
- Puranam, P., Alexy, O., and Reitzig, M. 2014. "What's ``New'' About New Forms of Organizing?," *Academy of Management Review* (39:2), pp. 162-180.
- Puranam, P., and Hakonsson, D. D. 2015. "Valve's Way," *Journal of Organization Design* (4:2), pp. 2-4.
- Ransbotham, S., and Kane, G. C. 2011. "Membership Turnover and Collaboration Success in Online Communities: Explaining Rises and Falls from Grace in Wikipedia," *MIS Quarterly* (35:3), pp. 613-627.
- Raymond, E. 1999. "The Cathedral and the Bazaar," *Knowledge, Technology & Policy* (12:3), pp. 23-49.
- Schroeder, A., and Wagner, C. 2012. "Governance of Open Content Creation: A Conceptualization and Analysis of Control and Guiding Mechanisms in the Open Content Domain," *Journal of the American Society for Information Science and Technology* (63:10), pp. 1947-1959.
- Shirky, C. 2009. *Here Comes Everybody: The Power of Organizing without Organizations*. New York: Penguin Group USA.
- Stvilia, B., Twidale, M. B., Smith, L. C., and Gasser, L. 2008. "Information Quality Work Organization in Wikipedia," *Journal of the American Society for Information Science and Technology* (59:6), pp. 983-1001.
- Thomas, E. J., and Biddle, B. J. 1966. *Role Theory: Concepts and Research*. Wiley.
- Turner, J. H. 1986. *The Structure of Sociological Theory*. Chicago, IL: University of Chicago Press.
- Turner, R. H. 1978. "The Role and the Person," *American Journal of Sociology* (84:1), pp. 1-23.
- von Krogh, G., Spaeth, S., and Lakhani, K. R. 2003. "Community, Joining, and Specialization in Open Source Software Innovation: A Case Study," *Research Policy* (32:7), pp. 1217-1241.
- Wagner, C., and Majchrzak, A. 2007. "Enabling Customer-Centricity Using Wikis and the Wiki Way," *Journal of Management Information Systems* (23:3), pp. 17-43.
- Welser, H. T., Cosley, D., Kossinets, G., Lin, A., Dokshin, F., Gay, G., and Smith, M. 2011. "Finding Social Roles in Wikipedia." iConference '11 Proceedings of the 2011 iConference
- Yang, D., Halfaker, A., Kraut, R., and Hovy, E. 2016. "Who Did What: Editor Role Identification in Wikipedia," *Tenth International AAAI Conference on Web and Social Media*.
- Zhu, H., Kraut, R. E., Wang, Y.-C., and Kittur, A. 2011. "Identifying Shared Leadership in Wikipedia," *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: ACM, pp. 3431-3434.



---

<sup>i</sup> Please see <http://www.mediawiki.org/wiki/Help:Namespaces>.

<sup>ii</sup> [https://en.wikipedia.org/wiki/Wikipedia:Your\\_first\\_article](https://en.wikipedia.org/wiki/Wikipedia:Your_first_article)

<sup>iii</sup> Contributors not holding these access privileges are allowed to submit a revert and wait for approval (by a contributor with these specific access privileges) <https://en.wikipedia.org/wiki/Help:Reverting>

<sup>iv</sup> See [https://en.wikipedia.org/wiki/Help:MediaWiki\\_namespace](https://en.wikipedia.org/wiki/Help:MediaWiki_namespace). Nonetheless, a contributor without a *sysop* access privilege can submit a contribution and wait for approval

<sup>v</sup> We also performed a sensitivity analysis, restricting the duration that activity was recorded before and after role assignment events. Our analysis of a 6-month and 12-month restriction revealed that the most significant differences between the ‘pre’ and ‘post’ lists were obtained for the unrestricted activity periods.

<sup>vi</sup> [https://en.wikipedia.org/wiki/Wikipedia:User\\_access\\_levels](https://en.wikipedia.org/wiki/Wikipedia:User_access_levels)