Research Center
Finance & Information Management

Fraunhofer FIT
Project Group
Business & Information
Systems Engineering

Emergent User Roles of a Digital Workplace: A Network Analysis Based on Trace Data

by

Leonhard Frank, Henner Gimpel, Marco Schmidt, Manfred Schoch

to be presented at: 38th International Conference on Information Systems (ICIS), Seoul, South Korea, December 2017

University of Augsburg, D-86135 Augsburg
Visitors: Universitätsstr. 12, 86159 Augsburg
Phone: +49 821 598-4801 (Fax: -4899)

University of Bayreuth, D-95440 Bayreuth
Visitors: Wittelsbacherring 10, 95444 Bayreuth
Phone: +49 921 55-4711 (Fax: - 844710)

WI-685

www.fim-rc.de

Emergent User Roles of a Digital Workplace: A Network Analysis Based on Trace Data

Completed Research Paper

Leonhard Frank
FIM Research Center
University of Augsburg
Universitätsstraße 12
86159 Augsburg, Germany
leonhard.frank@fim-rc.de

Henner Gimpel
FIM Research Center
University of Augsburg
Universitätsstraße 12
86159 Augsburg, Germany
henner.gimpel@fim-rc.de

Marco Schmidt
FIM Research Center
University of Augsburg
Universitätsstraße 12
86159 Augsburg, Germany
marco.schmidt@fim-rc.de

Manfred Schoch
FIM Research Center
University of Augsburg
Universitätsstraße 12
86159 Augsburg, Germany
manfred.schoch@fim-rc.de

Abstract

Communication and collaboration software for knowledge workers are introduced with high expectations, especially in knowledge-intense industries. While advantages of such tools are well documented in theory, many initiatives have yet to achieve the desired outcomes in practice. Research has dealt with roles in the digital workplace and found that one-size-fits-all solutions are not suitable. However, for a lack of real-world data the matter is still not sufficiently understood. To close this gap, we conduct a sequential mixed method study. We perform an exploratory analysis based on trace data within a service organization and reconstruct its social structure. Through a cluster analysis, eight distinct emergent user roles are identified. Additionally, we analyze covariates of cluster membership, such as organizational hierarchy, through statistical testing. Lastly, semi-structured interviews help to explain our findings qualitatively. We contribute to research and practice by deepening the understanding of heterogeneous user behavior in a digital workplace.

Keywords: digital workplace, social software, digital trace data, social structure, social network analysis, emergent user roles, communication channel, collaboration platform

Introduction

The tertiary and quaternary (knowledge-intense) sectors of the economy have long been on the rise and with it, the number of knowledge-intense jobs (Kenessey 1987). Many jobs in modern organizations, especially in the western world, require extensive amounts of knowledge work (Kane et al. 2012). In recent years, digitalization has brought forward many software tools to support communication and collaboration between knowledge workers. This development has lead the digital workplace to grow continuously, particularly with new additions such as social collaboration platforms, enterprise social networks (ESN), or new communication tools like instant messaging (Drakos et al. 2015). Consequently, these market trends have prompted the development of new comprehensive software solutions (Drakos et al. 2015; Pawlowski et al. 2014). These tools have introduced many new functionalities to the digital workplace with goals such as increasing knowledge distribution beyond formal communication lines (Alavi and Leidner 2001), mediating communication and collaboration in distributed work environments (Seebach et al. 2011), helping blur organizational boundaries (Pawlowski et al. 2014), and ultimately increasing the productivity of knowledge workers (Kane et al. 2012; Köffer 2015). While companies are implementing these software solutions with great expectations, researchers and practitioners often report that adoption, usage, and impact are not yet fully understood (e.g. Berger et al. 2014; Herzog et al. 2015; Kiron et al. 2013; Kügler et al. 2012). Existing academic literature found that *one size fits* all solutions are inappropriate to address the heterogeneous job requirements and user behaviors of the digital workplace (Köffer 2015; Maruping and Magni 2015). Therefore, there is growing interest in evaluating social software initiatives in order to understand (1) why some users are adopting communication and collaboration tools and others are not, (2) which features are used by different user groups, and (3) which users create and distribute information within the organization. As a first step to better understand this heterogeneous usage behavior of knowledge workers within the digital workplace, an integrated analysis of both communication and collaboration technology is vital. While several studies exist which have brought forward first contributions regarding this issue, researchers frequently note that for privacy reasons, findings based on real-world data are scarce (e.g. Pawlowski et al. 2014; Wang and Noe 2010).

Therefore, the aim of this paper is to derive a user typology from the informal social structure of a digital communication and collaboration environment in an organization, in order to understand the heterogeneous user behavior as well as the emergent roles that knowledge workers take on, and to investigate why they do so. The latter is necessary to draw specific inferences regarding theory and practice. To approach this goal, we conduct a mixed method study (Venkatesh et al. 2013): We start by deriving the social structure of an organization that provides knowledge-intense services from a digital trace data set, i.e. data on user activity recorded by an information system (Howison et al. 2011). We do so with the tools of social network analysis (SNA) which serves as the basis of all further analyses. Subsequently, we use a cluster analysis to explore various interaction types regarding the heterogeneous behavior of users. We then evaluate explanatory variables from metadata about the users through statistical testing in order to detect covariates of cluster membership. Lastly, we conduct semi-structured interviews with a theoretical sample of users informed by our previous findings to verify and better interpret our empirical results.

This study provides the following contributions: First, we identify eight distinct user roles of the digital workplace for knowledge workers from our real-world data set and explain their characteristics. Second, we find that several of the identified user roles show a strong relationship with the organizational hierarchy. Third, we categorize multiple other user roles as task-specific and report insights about them derived from the user interviews. This suggests that knowledge-sharing can be an in-role behavior for certain types of employees (Wang and Noe 2010). Fourth, we discuss how the identified user roles relate to the existing scientific body of knowledge, such as the organizational knowledge creation theory (Nonaka et al. 2006). Fifth, we discuss practical implications for the digital workplace that have previously been derived from the literature and discuss how our approach can help with addressing them.

The remainder of this paper is structured as follows: Section 2 gives an overview of the elements of a digital workplace for knowledge workers and reviews the existing literature regarding user roles of knowledge workers. Section 3 explains our mixed method approach and its components. Section 4 contains the results of the study. We then proceed to discuss the contributions derived from these results in Section 5. Lastly, Section 6 assesses our study critically regarding its limitations and concludes the paper.

Problem Context and Literature Review

Knowledge Creation and Social Structures

According to the knowledge-based theory of the firm, knowledge is the primary resource of an organization (Grant 1996) and a superior knowledge base increases the value of an organization and its performance (Kogut 2000). Yet, despite the importance of knowledge, organizations often do not know what they know, because their body of knowledge is comprised of the knowledge of individual employees as well as shared knowledge resulting from social interactions within the organization (Alavi and Leidner 2001). The fact that knowledge is mostly owned by employees places great emphasis on knowledge application and the role of the individual (Grant 1996). For knowledge workers, it is critical to know how and from whom to obtain the valuable information required to do their jobs (Cross et al. 2002). Congruent with that, a trend towards networked organizations and an emphasis on social networks of employees is noticeable. The social interactions inherent in such networks are a manifestation of the structural dimension of social capital and are related to the extent of resource exchange within an organization (Tsai and Ghoshal 1998). It is well studied that social contacts help the members of intrafirm networks to maintain and extend their social capital within the organization (Steinfield et al. 2008). Communication and collaboration tools of the digital workplace can foster interactions, in particular between employees who are on different hierarchical levels (Behrendt et al. 2015), or who have no formal social relations between one another (Faraj et al. 2011; Kane et al. 2014). This in turn helps employees to increase their access to the network and to gain social capital. Therefore, and to study organizational networks, an investigation of the implicit social structure that emerges from those interactions between the users of the digital workplace seems promising. While this is an important step towards understanding an organization's knowledge capability, little empirical research exists in that area (Richter et al. 2010). In relation to the implicit social structure, the existence of emergent roles is a particularly interesting topic in order to improve the understanding of user behavior. Emergent roles are roles that users take on implicitly and as a result of their interactions with others. In self-organizing collaboration communities such as Wikipedia, emergent roles are a cornerstone of the knowledge-creation process (Arazy et al. 2016). However, it remains unclear whether these emergent roles can also be observed for organizational settings.

The Digital Workplace for Knowledge Workers

Many jobs in modern organizations require extensive amounts of knowledge work (Kane et al. 2012). Thus, we are particularly interested in the digital workplace of the so-called knowledge workers. Knowledge workers are characterized as employees who "think for a living" (Davenport 2005, p. 3) and turn "complex information [...] into knowledge" (Davenport 2005, p. 3). Davenport further sharpens the definition of knowledge workers, as people that "have high degrees of expertise, education or experience, and the primary purpose of their jobs involves the creation, distribution, or application of knowledge" (Davenport 2005, p.10). Köffer (2015, p. 2) introduced the digital workplace based on C. Tubb as "the collection of all digital tools provided by an organization to allow employees to do their jobs". As a first step to investigating the digital workplace for knowledge workers, it is important to understand and define the different software tools available to them. Generally speaking, there are software tools which are driven by structured and reproducible business processes rather than human interactions (van der Aalst et al. 2011), and those which foster open digital interactions between employees (Wang and Noe 2010). Examples for process-driven tools are enterprise resource planning or workflow management systems. These systems are not well-suited for the identification of an implicit social structure between employees because they follow pre-defined processes and often do not leave room for spontaneous personal interactions. Without the set perimeters of pre-defined business processes, however, an implicit social structure can emerge freely. We classify such software tools congruently with McAfee (2006) as communication channels and collaboration platforms. Communication channels include peer to peer communication tools, such as email or instant messaging, and cannot be accessed or searched by others (McAfee 2006). Collaboration platforms, such as content management systems, wikis, and blogs, by comparison, are accessible to many or all employees within the organization and the knowledge stored in them is persistent (McAfee 2006). Both of those systems foster digital interactions between employees, and therefore represent how people go about their daily business and who they interact with digitally.

Related Work on User Roles

Recently, the existence and formation of emergent roles of knowledge workers has caught the interest of researchers. Multiple current studies have identified communication and collaboration use cases including *Broadcasting, Dialog, Collaboration, Knowledge Management,* and *Sociability* (Schlagwein and Hu 2016; Schubert and Glitsch 2016). While these use cases provide a detailed outline of the functionality and capabilities of such a software environment, the authors do not attribute the use cases to specific user roles. Regarding email communication, there are a number of studies that have looked into network structures (e.g. Bird et al. 2006; Kane et al. 2012; van Alstyne and Zhang 2003), but surprisingly little research has addressed user roles. Among the notable exceptions are Alavi and Leidner (2001), who defined that in a digital environment, knowledge flows from a *Provider* to a *Seeker*, and that balancing the two is desirable. Muller et al. (2010) used real-world data to investigate the consuming behaviors of *Uploaders, Contributors*, and *Lurkers* within an enterprise file-sharing system. Reinhardt et al. (2011) created a general typology of knowledge worker roles based on a literature review. Subsequently, they verified the existence of *Controllers, Helpers, Learners, Linkers, Networkers, Organizers, Retrievers, Sharers, Solvers,* and *Trackers* through a laboratory task execution study. Their paper provides a comprehensive overview of knowledge worker roles and their behaviors, but lacks a validation based on real-world data. In contrast to that, other authors have looked at real-world data of ESN to investigate the influence of formal hierarchy on user behavior (Behrendt et al. 2015; Riemer et al. 2015). Behrendt et al. (2015) found that in ESN, the hierarchy seems to have an influence on user behavior. Riemer et al. (2015), on the other hand, found that while hierarchy has a low influence on the likelihood of responses from the network, the users' own contributions are far more important. Those findings further substantiate the relevance of informal social structures in the context of ESN. However, it remains unclear how significant the influence of formal hierarchy on emergent roles is. A study by Arazy et al. (2016) employed a SNA to identify seven emergent roles within the self-organizing collaboration platform Wikipedia. In their study, they found *All-round Contributors, Quick-and-Dirty Editors, Copy Editors, Content Shapers, Layout Shapers, Watchdogs,* and *Vandals*. A similar exploratory study by Füller et al. (2014) investigates the heterogeneous user behavior and the social structure of a collaborative open-innovation-contest community based on real-world data. In their study, they found six distinct user roles*: Socializers, (active and passive) Idea-Generators, Masters, Efficient Contributors,* and *Passive Commentators*. While their research approach is conducive to our goal of identifying user roles in a digital workplace, it is questionable whether their results can be directly transferred to the organizational context.

In summation, several researchers have previously dealt with user roles in the context of digital communication or collaboration, both within and outside of organizations. Their approaches cover a number of different software systems and reveal a number of domain-specific emergent roles. However, those studies have yet to combine both the communication and collaboration structures of a digital workplace. Additionally, to the best of our knowledge, an area that has yet to be addressed is the investigation into user behaviors in conjunction with reasons explaining why users behave the way they do or perform a certain informal role – especially in the presence of formal roles.

Empirical Study

To address the identified research gap, we use a mixed method approach (Venkatesh et al. 2013), which combines aspects of previous studies by identifying user roles in an exploratory fashion, analyzing potential influencing factors quantitatively, and interviewing users qualitatively to better understand the reasons for why employees act the way they do.

Research Setting and Data Set

Our exploratory study is based on digital trace data from a service organization that provides knowledge-intense services to corporate and individual customers. This organization is well-suited for this study for multiple reasons. First, it has two different locations with distributed teams consisting of employees from both locations. Therefore, it relies heavily on a distributed and digitally enabled work environment. Second, the organization uses the standard software Microsoft Office 365 with its social collaboration component SharePoint and the communication system Exchange. In that regard, the platform resembles a significant part of the communication and collaboration technology used in many companies today (Drakos et al.

2015). Third, the organization almost exclusively employs knowledge workers. While this organization is well-suited for our research goal, we do acknowledge that studying a single organization bears limitations on the inferences that can be drawn from our study. Further, we acknowledge the limitation of only analyzing the most dominant digital collaboration and communication system in the organization, while for example omitting interactions through phone calls or personal contact for a lack of trace data.

The organization has multiple specialized departments which are responsible for the provision of the organization's external service offerings, and support functions that provide internal shared services, such as Finance or Human Resources (HR) to all departments. Each full-time employee is a member of exactly one department and one or multiple support functions. For the purpose of our research, we were provided with digital trace data for a period of six weeks across the months of March to May 2016. At the time, the organization had a total of 146 registered employees who are users of the digital workplace. Amongst the 146 users were 6 Heads of Departments, 6 Heads of Support Functions, 8 Assistants to the Heads of Departments, 35 Full-time Employees and 91 Part-time Employees. Part-time employees have variable working hours, generally with about 10 hours per week. Almost all users can be counted towards the knowledge worker category, as they mainly have high degrees of education and work experience in professions like management, business and financial services, or computer sciences (Davenport 2005).

For our study, the digital trace data was pseudonymized by the organization's system administrator to address privacy concerns (e.g. Herzog et al. 2015; Köffer 2015; Pawlowski et al. 2014; Wang and Noe 2010). This ensures the identification of communication and collaboration patterns but prevents the researchers from knowing about the content, or from identifying individual employees (van Alstyne and Zhang 2003). Both the Exchange and SharePoint logs contain only internal communication and collaboration, but do not include recipients or users outside of the organization. To identify characteristics of users, who perform a certain role, we were provided with the user-specific binary attributes *gender*, *site* (differentiating between the company's two sites), and *length of employment* (split into "long" and "short" according to the median), as well as the *position in the organizational hierarchy* (distinguishing between five hierarchical levels). The selection of the attributes and their granularity was chosen in such a way, that each combination of attributes matched multiple (or no) employees of the organization, but never a single one.

Social Network Analysis and Interaction Patterns

We use the tools of SNA as a basis to study the heterogeneous user behaviors and derive different user roles from the resulting social structure. SNA is ideally suited to study the actors of a given social system (Wasserman and Faust 1994) and has been used in social sciences for many decades (Borgatti et al. 2009). With metrics drawn from the social structure, actors can be distinguished, potentially resulting in new insights into user roles (Arazy et al. 2016; Füller et al. 2014). The foundation of many SNA concepts, such as *centrality* and other actor-related measures, is graph theory (e.g. Füller et al. 2014; Wasserman and Faust 1994). The relational structure of a social system consists of patterns of relationships among the actors of the system. Network data is fundamentally dyadic, meaning that ties are observed for a set of two actors at a time (Borgatti and Foster 2003). The sum of those actors and the ties amongst them form a social network (Wasserman and Faust 1994). Such an approach focuses on the patterns of interconnection but tends to neglect the content of the network ties between the actors (Borgatti et al. 2009). It is based on the idea that an actor's position in a network influences their opportunities and constraints (Kane et al. 2014). This approach is conducive to our pseudonymized data set which contains communication and collaboration patterns but not their contents.

SNA typically considers one or more of the following basic tie types: proximity (co-membership in groups, such as departments), relations (social relationships, such as friendship), interactions (discrete exchanges between nodes, such as a conversation), and flows (tangible or intangible material that moves from one node to another, such as information) (Borgatti et al. 2009; Kane et al. 2014). While flows are important, because "information flows drive knowledge transfer in organizations" (Alavi and Leidner 2001, p. 119), they are often difficult to measure. Consequently, and congruent with previous IS research regarding IT platforms and channels, we focus primarily on interactions (Kane et al. 2014). To understand the differences between our two IT systems, it is important to differentiate between the channel, which "pushes" information, and the platform, which requires users to "pull" information. For the push-medium email communication (i.e. Exchange), the sender initiates an interaction by sending an email. For the pullmedium content collaboration (i.e. SharePoint), however, the sender provides content to the IT system and the retriever accesses this content, resulting in an interaction.

The application of SNA in IS has long focused on single links, which contrasts multiplex approaches common in the social sciences (Howison et al. 2011). In our case, interactions can cover several distinct forms of communication or collaboration between two users. We define the following four possible dyadic interaction patterns that can be observed within the given data set, as presented in Figure 1:

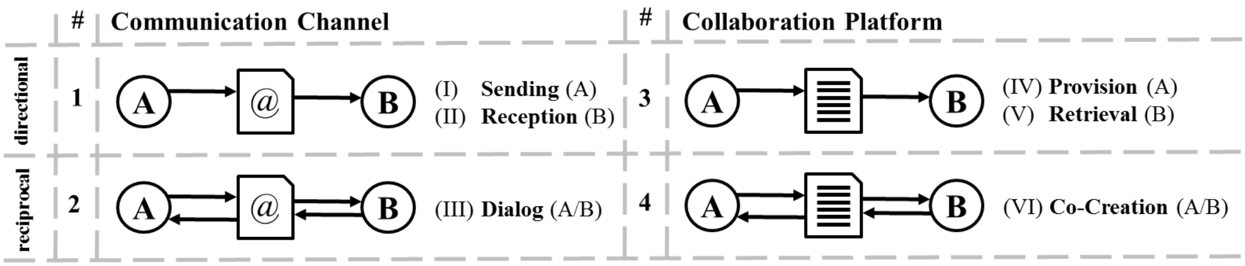

Figure 1: Interaction Patterns

Content co-creation and email dialog, as defined in this work, are by definition reciprocal and thus do not have a direction. The other two interaction types are directional, however. The strength of a tie is determined by the frequency or depth of a connection, which can be determined by interaction data (Kane et al. 2014). In our study, the strength of an interaction tie is defined by the number of different files and email subjects that two actors interact on.

In order for the observed interaction types to be transferred into input parameters for our cluster analysis, measures of contribution for the individual users need to be defined. There are several actor-based (egocentric) structural features that can be measured for a network which are commonly referred to as centrality of an actor (Füller et al. 2014; Kane et al. 2014; Wasserman and Faust 1994). Those concepts are related to the importance, prominence and visibility of an actor within a network. For the purpose of our study, we focus on degree centrality as a measure of activity (Wasserman and Faust 1994) and for greater access to network flows, such as information disseminated through interactions (Kane et al. 2014).

Analysis and Results

User Typology

To construct a social network from the log files, the defined interaction patterns were first mined from our digital trace data set. We find that the average number of colleagues a user is connected to through content collaboration is substantially lower than via email communication (10.6 and 8.9 for collaboration vs. 55.7 and 78.3 for communication). A deeper examination of the ties' intensity, which refers to the number of files or email subjects they have interacted on, reveals that users, who are connected, have on average approximately four bilateral and five unilateral communication ties (i.e. communicate on four email subjects in a discussion and on five subjects one-sidedly), but only three collaboration ties (i.e. collaborate on three files). In the social network, the overall number of interactions (weighted with their intensity) for the two directions of unilateral network ties (email sending/reception and content provision/retrieval, respectively) is identical, and therefore, the means are too. Median and standard deviation can differ depending on the directionality. For example, a single user can send emails to multiple recipients, which results in a more even distribution for email reception than for email sending. The mean number of sending and reception ties, however, stays the same. The descriptive statistics on the frequency of interactions (Table 1) show that more users are connected through communication ties (means of 271 and 297.4) than through collaboration ties (means of 33.2 and 23.2). The heterogeneous standard deviations substantiate the assumption that users behave differently from one another. A large standard deviation for the email sending measure (327.5 compared to 185.2 for email reception), for example, suggests that a limited number of users are responsible for the majority of the unilateral communication. However, due to the skewness of some of the data, the standard deviation has to be taken with a grain of salt.

	Variable	Mean	Median	SD	Skewness
I	Email Sending	271.0	170.0	327.5	3.70
II	Email Reception	271.0	212.0	185.2	1.35
III	Email Dialog	297.4	226.5	238.2	1.87
IV	Content Provision	33.2	18.5	47.3	3.41
V	Content Retrieval	33.2	22.5	43.2	4.17
VI	Content Co-Creation	23.2	11.0	29.3	2.27

Observations: n = 146 , SD = standard deviation

Table 1: Descriptive Statistics on the Frequency of Interactions

We used the interaction types to capture each user's communication and collaboration behavior as input variables for an exploratory cluster analysis aimed at identifying the distinct user types inherent in the social structure of our network. To do that, we first checked if both the measures for the unweighted graph, which records whether or not any tie exists between two users as a binary measure, and the weighted graph, which includes the strength of every tie, present a potential source of heterogeneity. We found that the spearman rank correlation coefficients between the unweighted and weighted means resides between 0.88 and 0.98, depending on the type of interaction. Therefore, we decided to only use the weighted graphs, because they contain more information and their interpretation regarding the usage patterns is more straight-forward, as it represents the extent to which the users use the interactions and not just the number of colleagues they are connected to.

For our cluster analysis, we used an agglomerative hierarchical procedure with the Ward.D2 minimum variance method and the Euclidian distance. Hierarchical clustering usually works well (Füller et al. 2014), is reproducible, and does not need the desired number of clusters, or their size, as an input parameter, which is conducive to our exploratory approach. Also, users that have been added to one cluster will remain in that cluster even if the cluster solution is changed, which helps with the process of determining the appropriate number of clusters. To eliminate outliers, we censored all values above the respective 98% quantiles.

"There is no universal definition for a good clustering size, [rather] the evaluation remains mostly in the eye of the beholder" (Rokach and Maimon 2005, p. 326, Bonner 1964). Several different stopping rules (Milligan and Cooper 1985) were employed, but yielded inconclusive results. We found that for eight clusters, the results are well interpretable. A lower cluster size joined multiple clearly distinct user groups, whereas more clusters resulted in very small cluster sizes with clusters that may be regarded as outliers rather than distinct user groups.

From our cluster analysis, we conclude the following typology: of the eight distinct user types, there are three that use both the communication channel and the collaboration platform roughly to the same extent. These clusters are labeled *All-rounders* with *low*, *mid*, and *high activity*. Four of the clusters are labeled according to a peak in one or more of six clustering dimensions. Two user types with peaks in communication interactions (Email heavy-users and broadcasters) were observed and two user types with peaks in collaboration interactions (Content co-creators and providers). Lastly, a user group that remains largely passive on both systems was identified. An overview of all clusters is provided in Table 2.

A nine cluster solution would have split Content Providers into two, creating a user group of two individuals that not only provide content, but also heavily retrieve content. As mentioned above, this group was omitted for its small size and because the characteristic attributes of Content Providers are still present in this ninth cluster. This is apparent in the data as part of the relatively high standard deviation of 0.35 in Content Retrieval of the Content Providers. A seven cluster solution, on the other hand, would have joined Content Co-Creators and All-rounders High-Activity that considerably differ in content co-creation and email dialog.

		Interaction Types											
		Communication Channel						Collaboration Platform					
		Reception		**Sending**		**Dialog**		**Retrieval**		**Provision**		**Co-Creation**	
User Role	#	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD	**Mean**	SD
All-rounder High-A.	9	**0.67**	0.11	**0.58**	0.14	**0.78**	0.20	**0.61**	0.23	**0.44**	0.21	**0.52**	0.13
All-rounder Mid-A.	16	**0.55**	0.16	**0.36**	0.13	**0.42**	0.13	**0.26**	0.10	**0.22**	0.10	**0.35**	0.12
All-rounder Low-A.	33	**0.30**	0.13	**0.20**	0.10	**0.28**	0.12	**0.19**	0.16	**0.13**	0.09	**0.14**	0.10
Email Heavy-User	8	**0.86**	0.13	**0.75**	0.15	**0.75**	0.15	**0.20**	0.07	**0.12**	0.07	**0.32**	0.27
Email Broadcaster	7	**0.31**	0.15	**0.89**	0.12	**0.53**	0.17	**0.11**	0.08	**0.15**	0.11	**0.07**	0.06
Content Co-Creator	11	**0.56**	0.16	**0.50**	0.15	**0.44**	0.11	**0.55**	0.14	**0.51**	0.21	**0.80**	0.20
Content Provider	8	**0.29**	0.07	**0.20**	0.07	**0.25**	0.07	**0.47**	0.35	**0.77**	0.24	**0.32**	0.13
Passive User	54	**0.17**	0.07	**0.08**	0.07	**0.13**	0.05	**0.06**	0.05	**0.03**	0.04	**0.04**	0.04

Table 2: User Typology with Corresponding Means and Standard Deviations (SD) of the Different Interaction Types

The *All-rounder High-Activity* (6.16% of 146 users) is characterized by fairly high email interactions, which suggests that this user type communicates heavily in a digital way, especially through email dialogs. A mean of 0.78 for email dialogs states that, on average, this user type has 78% of the interactions of the most active user in the network. This user type is also fairly active on the collaboration platform (1st to 3rd highest, depending on the interaction type), where they provide and retrieve content, in addition to co-creating content with their colleagues.

The *All-rounder Mid-Activity* (10.96%) is less active than its high-activity equivalent. While their number of received emails is comparable to those of an All-rounder High-Activity, they engage significantly less in reciprocal communication, as measured by the number of email dialogs.

The *All-rounder Low-Activity* (22.60%) forms the second largest cluster. This user type is considerably less active (2nd to 3rd last in all interaction types) than the formerly mentioned All-rounder types.

The *Email Heavy-User* (5.48%) engages much more heavily in email communication than in any collaborative activities. The peak in received emails is also substantial, which according to Wasserman and Faust (1994) is an indicator for a prestigious user. If this user type engages in any collaboration activity, it's mainly through co-creation of content with other users. Very rarely does this user type provide content that other users access.

The *Email Broadcaster* (4.79%) has a strong peak in outgoing email communication (most), but receives comparably little amounts of emails. However, this user type also has a fairly large amount of email dialog interactions (3rd most), likely as a result of the high number of sent emails.

The *Content Co-Creator* (7.53%) uses the collaboration platform and the communication channel fairly heavily, but has a substantial peak in content co-creation (most). This indicates that the user type collaborates heavily with other users in order to create tangible content.

The *Content Provider* (5.48%) is fairly active with regards to collaboration interactions and has a significant peak in content provision. This indicates that this user type creates tangible content that other users access frequently. The communication interactions, however, are sparse (2nd lowest) for this user type.

Finally, the *Passive User* group makes up for the majority of the users (36.99%). This user type has the lowest values across all interaction types and therefore does not participate particularly actively through digital communication or collaboration within the organization.

Covariates of Role Membership

To investigate the association between our categorical explanatory variables and the eight user types, we first examine the contingency tables illustrating the relative frequency distributions (Agresti 2007). We then apply a chi-squared-test for independence to determine whether there is a significant difference between the expected and observed frequencies. To deal with small cell values for rare user types, we simulate the associated p-values through a Monte Carlo Simulation (Adery 1968). First, we study the relationship of the identified user roles and the organizational hierarchy. Organizational hierarchy is a factor that has been mentioned frequently in literature regarding user behavior in the context of digital communication (Behrendt et al. 2015; Riemer et al. 2015). We observe a strong relation between the identified user roles and the position in the organizational hierarchy (Table 3). The association between the two variables is highly significant ($p<0.01$) with a chi-squared test statistic of $X^2 = 184.81$. We find that Heads of Departments and Heads of Support Functions tend to be heavy email-users, as observed in 50% of the cases. These users communicate heavily via email, but tend to use the collaboration platform to a substantially lesser extent. Assistants to a Head of the Department, conversely, mainly belong to the *All-rounder High-Activity* category. This user type is similarly involved in email communication than heavy email-users, but also engages heavily in collaborative activity, resulting in a more balanced usage of the collaboration platform and the communication channel. The full-time employees who don't hold a leadership role, are widely spread across the different user types, with a peak at *Content Co-Creators* and *All-rounders of Low- and Mid-Activity*. This shows that in our study regular full-time employees are generally less involved in email communication than their superiors. However, about one third of the full-time employees are heavily involved in collaborative activities, in particular content co-creation with other colleagues. This is an observation that will be subject to further qualitative investigation in the following section. Part-time employees are mostly *Passive Users*. This user type receives more emails than it sends and has a very low engagement on the collaboration platform. The rest of the part-time employees are mainly *All-rounders of Low-Activity*.

	Organizational Hierarchy					
User Role	**Head of Department**	**Head of Support F.**	**Assistant to H. of Dept.**	**Full-time Employee**	**Part-time Employee**	# of people
All-rounder High-Activity		17%	**63%**	9%		9
All-rounder Mid-Activity	17%	33%	25%	23%	3%	16
All-rounder Low-Activity	17%			17%	29%	33
Email Heavy-User	**50%**	**50%**	13%	3%		8
Email Broadcaster	17%			9%	3%	7
Content Co-Creator				**31%**		11
Content Provider				3%	8%	8
Passive User				6%	**57%**	54
# of people	6 (100%)	6 (100%)	8 (100%)	35 (100%)	91 (100%)	146

Table 3: Contingency Table for User Role and Organizational Hierarchy

In general, the organizational hierarchy does not fully explain all user types, but the different hierarchical levels show (more or less) clear tendencies towards a specific user type. To get a better picture of the factors related to the cluster membership, we proceed to analyze three additional potential covariates. First, regarding the organization's two different *sites,* we find a significant difference to the expected frequencies across all roles ($p<0.10$). According to a column-wise chi-squared test for goodness-of-fit, this is mainly due to the clusters All-rounder High and Mid-Activity, as well as due to the Email Broadcaster and Content Provider. For All-rounders High-Activity, the cause may be a higher number of Assistants to Head of Departments that are located at site A - the organization's oldest branch. Broadcasting and Content Provision activities might possibly be related to a high number of shared services, which are located at site A. Second, we examine the association between *gender* and emergent roles and do not find significant differences across our clusters ($p=0.58$). Previous studies regarding knowledge management have found significant influence of gender diversity on knowledge sharing (Wang and Noe 2010). Third, regarding the *length of employment* we find a highly significant association ($p<0.01$). We observe that Email Heavy-Users and All-rounders of High and Mid-Activity are more likely to have been with the company for a long time, while passive users have been with the company for only a short time significantly more often. However, both of those observations are correlated with the organizational hierarchy, as superiors tend to have been a part of the organization for a longer period of time than part-time employees in this organization.

User Interviews

We follow up on the quantitative results through qualitative user interviews as part of our mixed method approach to qualitatively confirm the quantitative results (Venkatesh et al. 2013). To do so, we conduct *semi-structured face-to-face interviews* with members of the organization (Myers and Newman 2007). The nine interviewees are selected based on *theoretical sampling* informed by the insights gained from our previous findings (Anderson 2010; Glaser and Strauss 2009). Because of the pseudonymized data, it is not possible to select interviewees based on their emergent role. However, due to the strong correlation between the organizational hierarchy and the identified user types, we are able to use the users' organizational positions to determine appropriate interview partners. Therefore, we select three part-time employees (A, B, C), three full-time employees (D, E, F), an Assistant to a Head of Department (G), a Head of Support Function (H), and a Head of Department (I). Similar to Behrendt et al. (2015), who used a mixed methods approach to investigate an ESN in a medical context, we defined the following two stages for the qualitative part of our study: Intended behavior and use cases of interaction types (Interview Stage 1), and addressing the findings of the quantitative section to allow for confirmation, rejection, and explanation (Interview Stage 2). All interviews were conducted, recorded, and transcribed by the authors of this paper. The transcripts were then coded iteratively to identify categories of repeated answers that address the overarching questions of the two interview stages mentioned above.

Intended Behavior and Use Cases

In the first stage, we intend to learn more about why the interviewees use the communication channel and collaboration platform respectively, and why they engage in the respective identified interaction types. In general, email communication is used for coordination, information sharing, or to document decisions in written form particularly with other employees who are not physically available. Email dialog is mainly used for coordination and status updates, while unanswered emails are for announcements, triggers or simply to inform somebody about something – for example through a copy of an email.

The collaboration platform, on the other hand, is used to co-create and archive knowledge, to make content accessible to a larger audience, and to look for and find information. For content co-creation, people frequently mentioned use-cases, which require intensive teamwork. In addition to co-creating content, they also mentioned receiving input or detailed in-text feedback through that kind of interaction. It was frequently mentioned that content stored on the platform is persistent, durable, and save. Additionally, administrative tasks such as shared lists, instructions and tutorials were mentioned. Content retrieval is used to access (or provide) input for knowledge creation, informational lists, meeting minutes, and other protocols. Overall, this shows that users are making conscious decisions about when they use which software. It also confirms that our defined interaction types are indeed recording heterogeneous behavior and that the patterns capture distinct information.

When asked about the most important influencing factors for why somebody would use communication channels or collaboration platforms more or less intensely, the interviewees almost unanimously confirmed the position in the hierarchy to be of relevance, and also mentioned the nature of the individual tasks. Interviewee H stated: "You have to view it in the context of the task. [A part-time employee] has vastly different communication requirements than an Assistant to the Head of Department, who has to coordinate important strategic issues with multiple stakeholders". Experience with the software systems, as well as personal preference and IT skills were also mentioned in this context.

Addressing the Quantitative Findings

In the second stage, we asked the interviewees to address our quantitative findings and to provide explanations as to why the observed patterns may exist. For that, they were shown versions of Figure 1, and Table 2 and 3 before being asked questions such as: "We observed that Assistants to a Head of Department are more heavily involved in content collaboration than other employees. Judging from your experience and interaction with them, is this a plausible observation and if so, why do you think they are?"

All but two *Passive Users* are part-time employees. Per our interviewees, part-time employees communicate and collaborate significantly less because they work less hours and have fewer tasks: "They have fewer duties that they need to communicate and collaborate on. Things like delegating, controlling, and guiding are mainly done through communication – and that's not typically part of a part-time employee's job description", Interviewee H.

We identified three levels of *All-rounders*, who use the two systems with rather similar intensity. Thus, we conclude that Mid-Activity All-rounders represent the average usage amongst employees who work full hours, while Low-Activity All-rounders use both systems to a lesser degree. High-Activity All-rounders are occupied by middle managers who depend on documenting decisions in a structured way: "Depending on the size of their department, they have to maintain a lot of lists to keep an overview of all the topics that they deal with. They also gather a lot of information from the entire organization and transform or condense it for their bosses", Interviewee G. They also often organize meetings and bring decisions made by the participants into practice, which requires extensive amounts of communication: "It has got to do with our responsibilities. Management assistants are the binding element between their superiors and the other employees. They have to gather a lot of information, condense it, and pass it on. That happens mainly via email, as many employees are working on external projects during the week", Interviewee H.

According to our interviewees, *Email Broadcasters* are (1) organizers of certain expert group meetings and other regular events, who ask for input from the participants, send agendas, and schedule meetings, or (2) the main secretary's office, which often sends emails to multiple recipients to inform them about changes regarding meetings, updates about decisions, or forward emails that they receive centrally but for which they are not responsible, or (3) single-point-of-contacts: "I receive emails with some brief information from my boss, based on which I write a proper email and communicate the matter to everybody else in the department", Interviewee B.

Email Heavy-Users communicate more than they collaborate with others. The high number of incoming emails indicates that these users are particularly prestigious (Wasserman and Faust 1994). First, managers "have exponentially more tasks" than employees on lower hierarchy levels. "It's a cascading effect. For every task you receive status updates which accumulate accordingly", Interviewee E. They give input, set goals, and monitor progress, but do not necessarily get involved operationally. Secondly, the reason why this communication is done via email, was explained by a lack of in-person availability. "That's why they depend heavily on emails. Usually, they answer a bulk of emails in the evening", Interviewee G. Interviewee I added that he uses emails frequently because he "travels a lot and the integration of the email client works flawlessly on the smartphone".

Content providers are all located at site A where most shared services are situated. We therefore suggest that this user behavior is task-specific. According to our interviewees, there are employees who are responsible for creating and updating tutorials, descriptions, FAQs, or templates. Frequently mentioned were the IT, Public Relations, and Finance departments. Given the fact that most Content Providers are part-time employees, and that the information stored in the mentioned documents is rather broad, we conclude that Content Providers are employees who gather and document information, rather than necessarily creating it themselves in the first place. Another interesting finding from the self-assessment was that content provision was rated low across the board, which suggests that providers of content are often unaware of others using their work.

For *Content Co-Creators***,** extensive team work is an important factor. Interviewee F said: "that's again task-related. More time for projects, proposals, or evaluation reports means more collaboration with others." Some interviewees, mentioned that teams which work in distributed environments, such as different internal locations or external projects, might engage more in content co-creation.

Meta-findings

To sum up our insights from the three parts of this study, we provide the following meta-inferences from integrating the qualitative and quantitative findings (Venkatesh et al. 2013). The results of the different parts of our study are presented in Table 4.

We found that part-time employees use the communication channel and the collaboration platform less frequently than full-time employees. However, task-specific exceptions, such as Content Providers, or Email Broadcasters are possible. In the user role Content Provider, part-time employees do not necessarily create new knowledge, but document existing tacit knowledge or merge dispersed knowledge to make it tangible. Full-time employees occupy many different user roles. The majority of them use both systems with relatively equal intensity and tend to be All-rounders of Low- or Mid-Activity. However, for task-specific reasons, about one third of them are engaged in tacit knowledge creation with their co-workers and are therefore Content Co-Creators. All of the user roles observed for full-time employees communicate significantly less than the roles most frequently observed for top managers (Head of Support Function, Head of Department) and middle managers (Assistant to Head of Department). Assistants to the Heads of Departments are highly active on both systems, and are thus High-Activity All-rounders. They have a broad portfolio of tasks where they are required to obtain information from employees and restructure or condense them to suit the needs of their superiors. In addition to that, they frequently organize meetings and take minutes to document decisions made by their superiors. Heads of Departments, just like Heads of Support Functions, are mainly using the communication channel, and not the collaboration platform. Their job profile requires extensive amounts of coordination and communication, because they are ultimately responsible for all tasks within their departments and are required to keep up with all developments, as well as to give high level input or feedback where necessary. Due to their limited in-person availability the communication is often asynchronous and therefore digital.

Several outliers that do not follow the observed correlations between user roles and organizational positions, are also apparent. For users who communicate or collaborate less than the rest of their co-workers on the same hierarchical level, this could be for personal factors such as vacation time, which we did not include into the quantitative part of our study for privacy reasons. Particularly interesting, however, are users who communicate and collaborate more than their peers. For example, part-time employees who are Mid-Activity All-rounders, or full-time employees who are High-Activity All-rounders. We suggest and our interviews support, that these users might be so called *hidden leaders*. Such employees use relationships and interactions with others to manifest their leadership, and do not rely on a hierarchical position to influence others (Edinger and Sain 2015).

User Role	Profile	Most Common Hierarchical Position	Other important Attributes	Qualitative Insights
All-rounder High-Activity	Frequent email communication, especially dialog. Frequent content collaboration	Assistant to Head of Dept.	Long employment & Site A	Middle management; broad portfolio of tasks; structured documentation; efficiency of coordinative tasks.
All-rounder Mid-Activity	Moderate email communication. Moderate to low content collaboration	All levels	Long employment	Average usage of channel and platform.
All-rounder Low-Activity	Moderate to low email communication. Low content collaboration.	Part- & Full-time Employee	-	Below average usage of channel and platform.
Email Heavy-User	Frequent email communication, especially reception. Low content collaboration.	Head of Support Function & Head of Department	Long employment	Limited in-person availability; lots of coordination, input, and feedback through cascading effects of responsibilities.
Email Broadcaster	Moderate email communication, but very frequent email sending. Low content collaboration.	Part- & Full-time Employee	Site A	Task-specific: scheduling of meetings; newsletters; single-point-of-contact in certain shared services, e.g. IT department, secretary's office.
Content Co-Creator	Moderate email communication. Frequent content collaboration, especially content co-creation.	Full-time Employee	-	Task-specific: when extensive team work is required and in distributed teams: e.g. research, written proposals, internal and external projects.
Content Provider	Low email communication. Frequent content collaboration, especially content provision.	Part-time Employee	Site A	Shared services and administrative tasks: e.g. instructions, tutorials, and templates in Finance, IT, HR departments.
Passive User	Very low email communication. Very low content collaboration.	Part-time Employee	Short employment	Fewer tasks & work hours; mainly operational tasks; more in-person contact through open-plan office, less meetings.

Table 4: Meta-Findings - User Roles with Quantitative and Qualitative Factors

Discussion

Theoretical Implications

Several researchers have previously dealt with roles of knowledge workers, different use cases of communication and collaboration software, and hierarchical differences in social software usage. However, the previous findings leave room for further contributions. This is due to several reasons: First, little research relies on real-world data. Second, the rare exceptions do not combine both collaboration and communication systems in an integrated way. Third, the mentioned studies rarely investigate exogenous covariates for a specific user behavior. Our paper identifies and analyzes eight heterogeneous user roles to address this gap.

Previous research regarding ESN has found relationships between the organizational hierarchy, on the one hand, and communication and knowledge sharing, on the other hand (Behrendt et al. 2015). Others, however, call for deemphasizing the role of hierarchy in knowledge sharing (Wang and Noe 2010). In our study, we find strong associations to the organizational structures for many user roles. However, for other roles, specific tasks that the users perform seem to be the distinguishing factor. For example, the user group identified as Content Providers has frequently been described in the literature as Providers or Sharers (Alavi and Leidner 2001; Reinhardt et al. 2011). According to several statements of the software environment's users in the qualitative part of our study, Content Providers are people whose jobs require them to gather information and create content that is frequently accessed by other users. This is congruent with Wang and Noe (2010) who state that knowledge sharing can be an in-role behavior for certain employees. The same applies to Email Broadcasters. Schlagwein and Hu (2016) observed broadcasting behavior in the context of ESN, and directly compare it to email broadcasting. According to the authors, broadcast in general is primarily aimed at reaching many users with a preconceived message. Such messages usually contain formal rather than informal information, when transmitted via email (Schlagwein and Hu 2016). Based on our user interviews, the respective user group is indeed tasked with broadcasting of information, e.g. in the form of internal newsletters. In addition to that, we learn from our interviews that the group might also be involved in the planning and scheduling of meetings, which according to Reinhardt et al. (2011) is the task of an Organizer. Due to the pseudonymized data set, we cannot conclusively say whether organizing is a relevant factor for the emergence of Email Broadcasters. For instance, according to our interviews Assistants to the Heads of Departments are also frequently involved in such activities, but in addition to that they also heavily participate in other interactions. Therefore, while we find users who perform tasks attributed to an Organizer, we cannot say with certainty whether some of them would form their own user group if the content of their interactions were considered.

A large part of the users in our study are all-rounders, which is congruent with a study by Arazy et al. (2016), who investigated emergent user roles in the open collaboration platform Wikipedia. For example, in our study, the majority of Assistants to the Heads of Departments – who are middle managers – are High-Activity All-rounders characterized by high levels of communication and collaboration activities. The organizational knowledge creation theory (Nonaka et al. 2006) can provide an explanation for this observation. It has, amongst other things, dealt with the role of leadership in knowledge management. According to Nonaka et al. (2006), top level managers communicate and coordinate visions about knowledge throughout the organization. Congruent with that, we find that Heads of Departments and Support Functions – who are top managers – are heavily involved in email communication and not so much in collaborative activities such as content provision or co-creation. For reasons of cost and time, not all knowledge can be shared (Nonaka et al. 2006). This is particularly the case for people high up in the hierarchy whose time is particularly precious. According to our interviews, this might be a reason for why Heads of Departments and Support Functions tend to create less tangible content through the collaboration platform and use asynchronous and verbal communication more frequently. Middle managers, on the other hand, bring the visions of top managers into concepts and facilitate organizational knowledge creation by synthesizing knowledge of front line employees as well as of their top managers and help make it explicit (Nonaka et al. 2006). These users are described in our user interviews as employees who gather information and reshape it to suit the needs of their superiors. In that sense, their behavior also resembles that of Linkers who "mash up information from different sources to generate new information", as found in a study by Reinhardt et al. (2011).

Contrary to previous studies which hypothesized and found Retrievers, Learners or Seekers (Alavi and Leidner 2001; Reinhardt et al. 2011), we do not find a user group that has peaks in content retrieval in our real-world data set. While many of the identified user types rely heavily on content retrieval, they also convert that information into tangible content to a similar extent. Because our study is based on social network data, we only consider content that was modified within the six-week observation period. It remains unclear whether the absence of Retrievers might be influenced by that restriction. However, it seems reasonable that employees do not look for information simply for the sake of knowing it, but that they do something with the obtained information. This then results in more balanced user types, which according to Alavi and Leidner (2001) is desirable, at least on an aggregated organizational level.

Several previous studies regarding digital social structures report about a dense network core and a large periphery of rather passive users (e.g. Füller et al. 2014; Muller et al. 2010). We, too found a passive user type, however, we are uncertain whether this is due to the uncommon organizational structure with many part-time employees or if it is a phenomenon that can generally be observed for employees with operative tasks. Congruent with our observation, and within a different organization, Behrendt et al. (2015) found that lower hierarchical levels are less active in ESN. In their study, the lowest hierarchical levels barely participate in ESN at all, average hierarchical levels have the most social relationships, middle managers communicate actively, and top managers reach many users at once. In our study, some part-time employees pointed out, that their lack of digital communication and collaboration might be due to a higher level of personal interactions in their open-plan offices. However, the effect of such personal interactions on digital interactions are not considered in our quantitative analysis.

Lastly, we find several employees who do not fall into task specific roles, but also are not in the same cluster as their colleagues on the same hierarchical level. We consider these to be outliers that communicate and collaborate more than their peers. According to social capital theory, users can gain social capital on an individual and relationship level from such informational exchanges with their colleagues (Steinfield et al. 2008). Our interviewees state that being well-connected in the digital workplace can be one aspect of several important aspects for a promotion. Congruent with that, they also state that there are a number of colleagues who are particularly involved in communication and collaboration, for example because they are experts in a particular field. Therefore, it might be possible that some of these users are hidden leaders or experts of some sort.

Managerial Implications

Our contributions, can be used to help practitioners with addressing *six* of the practical challenges for collaborative work in the digital workplace, which Köffer (2015) extracted through a literature review. First and most generally, we show a way to *monitor general work behaviors (1)* through digital trace data with our study. While privacy issues might limit the usefulness of such an analysis in an organizational context, our approach does provide a way to investigate how communication and collaboration systems are being used on an organizational level. This might help organizations to assess the overall adoption rates and identify areas for improvement. It could also be interesting for platform owners, who can study which features – if defined as interaction types – are being used by which user groups. Second, Maruping and Magni (2015) report that with the diversity of work practices, no *one size fits all* strategy regarding the incorporation of collaboration technology can be pursued. With our typology of user roles, we provide guidance for practitioners to *segment employees (2)*, not only regarding their collaboration behavior, but also regarding their communication requirements (Cameron and Webster 2013). Third, through identifying different user types in our study, we also help organizations to better understand user needs based on which they can *provide support and training (3)*, tailored to the individual needs of their employees. As mentioned in Section 3, for data privacy reasons it would be challenging for organizations to recreate this analysis in order to identify individual employees, however, in our analysis of covariates of cluster membership, as well as our qualitative interviews, we described the user types and their characteristics in depth. This might help organizations to target entire homogeneous groups of knowledge workers with their support or training efforts, rather than individual users. Fourth, and connected to the previous point, through the identification of Passive Users, employees with a small number of ties can be encouraged to interact with others (Zhang and Venkatesh 2013), which in turn helps to *enable social interactions (4)*. Fifth, by getting a better idea of the communication and collaboration requirements of each hierarchical level, practitioners are also supported to more adequately *consider individual characteristics (5)*, such as digital skills and experience in their hiring or promotion decisions. For example, the 9% of full-time employees that reside in the High-activity All-rounder cluster and the Email Heavy-Users cluster might be candidates for a more communication-heavy job in management. Last, top management support is often cited as a critical success factor for the adoption of new software tools and for a positive knowledge sharing culture (e.g. Wang and Noe 2010). We found that middle managers are particularly engaged in communication and collaboration as per their job requirements, which might make them better advocates to *demonstrate leadership* (6) on novel (social) collaboration platforms or ESN.

Limitations and Future Research

Our study has a number of limitations and leaves room for further research. While our data set is taken from an organization that is well-suited to study knowledge workers in the digital workplace, it only represents a small sample of knowledge workers. Additionally, we only capture white-collar knowledge workers with our study, therefore our results cannot necessarily be generalized to other knowledge workers, such as healthcare practitioners or engineers. Also, while many of the user types found in this study overlap with those identified in previous studies in other settings, we cannot say with certainty that these user types are also inherent in the social structure of other organizations. Therefore, further research based on different data sets is necessary to validate the generalizability of our findings. Likewise, we follow an "eye of the beholder" clustering approach, which leans heavily on the interpretation of the identified clusters. While we provided extensive qualitative details to support our selected clustering solution, this remains an explorative approach which, again, needs to be validated in future research contributions. The maturity of the software usage within the organizations and personal IT skills could be considered to draw comparisons between organizations. A problem that is frequently mentioned in the context of SNA based on digital trace data is that by definition it only considers social interactions within the software environment. For example, it neglects undocumented face-to-face interactions and interactions through other software tools (Wang and Noe 2010). Howison et al. (2011) caution not to over-interpret the number of digital events between employees, because the intensity and content of the interactions is unknown. Yet, researchers could define more distinct interaction patterns for future work, to distinguish further between user types. For example, Gleave et al. (2009) present different ego-networks and hypothesize that their shapes can give hints about the roles of actors. Additionally, for privacy reasons our analysis neglects the content of the interactions and the actual information flows transmitted through them. *Hashing* and *speech acts* have been used in the past to allow for an automatic analysis while maintaining the anonymity of the data (Carvalho and Cohen 2005; van Alstyne and Zhang 2003) and could be applied to this context as well. Another interesting question for further research is whether the employees keep or change their user roles over time. And if they change, what external factors cause those role changes. Researchers in the context of Wikipedia have found a turbulent stability of emergent roles, which describes the phenomenon that individual user roles may change, but the overall composition remains the same (Arazy et al. 2016).

Conclusion

In this study, we addressed the need to gain a better understanding of the heterogeneous behaviors of knowledge workers within their digital workplace in an organization. The importance of this question is rooted in the understanding that one size fits all solutions regarding the incorporation of such software into the diverse work practices are not adequate. Therefore, and to improve our knowledge of how these work practices differ, we set out to identify emergent user roles of a communication and collaboration environment. This endeavor is rooted in the knowledge-based theory of the firm and social capital theory, as well as in a fragmented body of research on the digital workplace and user roles in digital communication and collaboration environments. As a result of a cluster analysis, we found eight distinct user roles. In contrast to other studies in different contexts, we found that the presence of organizational roles can help explain many behavioral differences through factors such as the organizational hierarchy and the individual job requirements of the users. Those findings are routed in a quantitative analysis of influencing factors and qualitative user interviews. We observe that, congruent with the organizational knowledge creation theory, top managers are heavily involved in communication, while middle managers bridge the gap between top managers and employees by turning visions into tangible content. For user types that distribute information and provide content, we observed usage patterns that can be explained through an in-role understanding of knowledge sharing. Similarly, for employees who are heavily involved in tasks that require team work, a tendency towards co-creation of content with colleagues was observed. Lastly, and congruent with the positive effects of social connections on social capital, we argue that outliers can potentially be hidden leaders and candidates for promotions. With our approach, we contribute to the scientific progress in the field and support practical implications of communication and collaboration in the digital workplace. Future research should refine our interaction types and validate our findings with different data sets, particularly through but not limited to longitudinal designs.

References

Adery, C. A. H. 1968. “A Simplified Monte Carlo Significance Test Procedure,” *Journal of the Royal Statistical Society. Series B (Methodological)* (30:3), pp. 582–598.

Agresti, A. 2007. *An Introduction to Categorical Data Analysis*, Hoboken, NJ: John Wiley & Sons, Inc.

Alavi, M., and Leidner, D. E. 2001. “Review: Knowledge Management and Knowledge Management Systems: Conceptual Foundations and Research Issues,” *Management Information Systems Quarterly* (25:1), pp. 107–136.

Anderson, C. 2010. “Presenting and Evaluating Qualitative Research,” *American Journal of Pharmaceutical Education* (74:8).

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 Coproduction,” *Information Systems Research* (27:4), pp. 792–812.

Behrendt, S., Klier, J., Klier, M., and Richter, A. 2015. “The Impact of Formal Hierarchies on Enterprise Social Networking Behavior,” in *Proceedings of the 36th International Conference on Information Systems,* Fort Worth, Texas, USA. December 13-16, 2015, Association for Information Systems.

Berger, K., Klier, J., Klier, M., and Richter, A. 2014. “”Who is Key…?” - Characterizing Value Adding Users in Enterprise Social Networks,” in *Proceedings of the European Conference on Information Systems (ECIS),* Tel Aviv, Israel. June 9-11, Association for Information Systems.

Bird, C., Gourley, A., Devanbu, P., Gertz, M., and Swaminathan, A. 2006. “Mining Email Social Networks,” in *Proceedings of the 2006 International Workshop on Mining Software Repositories*, S. Diehl (ed.), Shanghai, China. May 22-23, ACM, pp. 137–143.

Bonner, R. E. 1964. “On Some Clustering Techniques,” *IBM Journal of Research and Development* (8:1), pp. 22–32.

Borgatti, S. P., and Foster, P. C. 2003. “The Network Paradigm in Organizational Research: A Review and Typology,” *Journal of Management* (29:6), pp. 991–1013.

Borgatti, S. P., Mehra, A., Brass, D. J., and Labianca, G. 2009. “Network Analysis in the Social Sciences,” *Science* (323:5916), pp. 892–895.

Cameron, A.-F., and Webster, J. 2013. “Multicommunicating: Juggling Multiple Conversations in the Workplace,” *Information Systems Research* (24:2), pp. 352–371.

Carvalho, V. R., and Cohen, W. W. 2005. “On the Collective Classification of Email ‘Speech Acts’,” in *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval,* Salvador, Brazil. August 15-19, pp. 345–352.

Cross, R., Borgatti, S. P., and Parker, A. 2002. “Making Invisible Work Visible: Using Social Network Analysis to Support Strategic Collaboration,” *California Management Review* (44:2), pp. 25–46.

Davenport, T. H. 2005. *Thinking for a Living: How to Get Better Performances and Results from Knowledge Workers*, Boston, MA: Harvard Business School Press.

Drakos, N., Gotta, M., and Mann, J. 2015. “Magic Quadrant for Social Software in the Workplace,” Gartner Research.

Edinger, S. K., and Sain, L. 2015. *The Hidden Leader: Discover and Develop Greatness within Your Company*, New York: AMACOM.

Faraj, S., Jarvenpaa, S. L., and Majchrzak, A. 2011. “Knowledge Collaboration in Online Communities,” *Organization Science* (22:5), pp. 1224–1239.

Füller, J., Hutter, K., Hautz, J., and Matzler, K. 2014. “User Roles and Contributions in Innovation-Contest Communities,” *Journal of Management Information Systems* (31:1), pp. 273–308.

Glaser, B. G., and Strauss, A. L. 2009. *The Discovery of Grounded Theory: Strategies for Qualitative Research*, New Brunswick: Aldine Transaction.

Gleave, E., Welser, H. T., Lento, T. M., and Smith, M. A. 2009. “A Conceptual and Operational Definition of Social Role in Online Community,” in *42nd Hawaii International Conference on System Sciences,* Waikoloa, Hawaii, USA. January 05-08.

Grant, R. 1996. "Toward a Knowledge-Based Theory of the Firm," *Strategic Management Journal* (17:S2), pp. 109–122.

Herzog, C., Richter, A., and Steinhueser, M. 2015. "Towards a Framework for the Evaluation Design of Enterprise Social Software," in *Proceedings of the 36th International Conference on Information Systems,* Fort Worth, Texas, USA. December 13-16, 2015, Association for Information Systems.

Howison, J., Wiggins, A., and Crowston, K. 2011. "Validity Issues in the Use of Social Network Analysis with Digital Trace Data," *Journal of the Association for Information Systems* (12:12), pp. 767–797.

Kane, G., Ransbotham, S., and Boynton, A. 2012. "Is High Performance Contagious among Knowledge Workers?" in *Proceedings of the 33rd International Conference on Information Systems (ICIS),* Orlando, Florida, USA. December 16-19, Association for Information Systems.

Kane, G. C., Alavi, M., Labianca, G., and Borgatti, S. P. 2014. "What's Different about Social Media Networks? A Framework and Research Agenda," *Management Information Systems Quarterly* (38:1), pp. 275–304.

Kenessey, Z. 1987. "The Primary, Secondary, Tertiary and Quaternary Sectors of the Economy," *Review of Income and Wealth* (33:4), pp. 359–385.

Kiron, D., Palmer, D., Phillips, A. N., and Berkman, R. 2013. "Social Business: Shifting Out of First Gear," *MIT Sloan Management Review* (55:1), pp. 1–32.

Köffer, S. 2015. "Designing the Digital Workplace of the Future – What Scholars Recommend to Practitioners," in *Proceedings of the 36th International Conference on Information Systems,* Fort Worth, Texas, USA. December 13-16, 2015, Association for Information Systems.

Kogut, B. 2000. "The Network as Knowledge: Generative Rules and the Emergence of Structure," *Strategic Management Journal* (21:3), pp. 405–425.

Kügler, M., Smolnik, S., and Raeth, P. 2012. "Why Don't You Use It? Assessing the Determinants of Enterprise Social Software Usage: A Conceptual Model Integrating Innovation Diffusion and Social Capital Theories," in *Proceedings of the 33rd International Conference on Information Systems (ICIS),* Orlando, Florida, USA. December 16-19, Association for Information Systems.

Maruping, L., and Magni, M. 2015. "Motivating Employees to Explore Collaboration Technology in Team Contexts," *Management Information Systems Quarterly* (39:1), pp. 1–16.

McAfee, A. 2006. "Enterprise 2.0: The Dawn of Emergent Collaboration," *MIT Sloan Management Review* (47:3), pp. 21–28.

Milligan, G. W., and Cooper, M. C. 1985. "An Examination of Procedures for Determining the Number of Clusters in a Data Set," *Psychometrika* (50:2), pp. 159–179.

Muller, M. J., Shami, N. S., Millen, D. R., and Feinberg, J. 2010. "We are All Lurkers: Consuming Behaviors Among Authors and Readers in an Enterprise File-Sharing Service," in *Proceedings of the 16th ACM International Conference on Supporting Group Work,* Sanibel Island, Florida, USA. November 07-10, ACM, pp. 201–210.

Myers, M. D., and Newman, M. 2007. "The Qualitative Interview in IS Research: Examining the Craft," *Information and Organization* (17:1), pp. 2–26.

Nonaka, I., Krogh, G. von, and Voelpel, S. 2006. "Organizational Knowledge Creation Theory: Evolutionary Paths and Future Advances," *Organization Studies* (27:8), pp. 1179–1208.

Pawlowski, J. M., Bick, M., Peinl, R., Thalmann, S., Maier, R., Hetmank, L., Kruse, P., Martensen, M., and Pirkkalainen, H. 2014. "Social Knowledge Environments," *Business & Information Systems Engineering* (6:2), pp. 81–88.

Reinhardt, W., Schmidt, B., Sloep, P., and Drachsler, H. 2011. "Knowledge Worker Roles and Actions-Results of Two Empirical Studies," *Knowledge and Process Management* (18:3), pp. 150–174.

Richter, D., Riemer, K., and Vom Brocke, J. 2010. "Social Transactions on Social Network Sites: Can Transaction Cost Theory Contribute to a Better Understanding of Internet Social Networking?" in *Proceedings of the 23rd Bled eConference: eTrust: Implications for the Individual, Enterprises and Society,* Bled, Slovenia. June 20-23.

Riemer, K., Stieglitz, S., and Meske, C. 2015. "From Top to Bottom," *Business & Information Systems Engineering* (57:3), pp. 197–212.

Rokach, L., and Maimon, O. 2005. "Clustering Methods," in *Data Mining and Knowledge Discovery Handbook*, New York, NY: Springer, pp. 321–352.

Schlagwein, D., and Hu, M. 2016. "How and Why Organisations Use Social Media: Five Use Types and Their Relation to Absorptive Capacity," *Journal of Information Technology*.

Schubert, P., and Glitsch, J. H. 2016. "Use Cases and Collaboration Scenarios: How Employees Use Socially-Enabled Enterprise Collaboration Systems (ECS)," *International Journal of Information Systems and Project Management* (4:2), pp. 41–62.

Seebach, C., Beck, R., and Pahlke, I. 2011. "Situation Awareness through Social Collaboration Platforms in Distributed Work Environments," in *Proceedings of the 32rd International Conference on Information Systems (ICIS),* Shanghai, China. December 04-07, Association for Information Systems.

Steinfield, C., Ellison, N. B., and Lampe, C. 2008. "Social Capital, Self-Esteem, and Use of Online Social Network Sites: A Longitudinal Analysis," *Social Networking on the Internet Developmental Implications* (29:6), pp. 434–445.

Tsai, W., and Ghoshal, S. 1998. "Social Capital and Value Creation: The Role of Intrafirm Networks," *Academy of Management Journal* (41:4), pp. 464–476.

van Alstyne, M., and Zhang, J. 2003. "EmailNet: A System for Automatically Mining Social Networks from Organizational Email Communication," NAACSOS, Pittsburgh: Carnegie Mellon.

van der Aalst, W., Adriansyah, A., Medeiros, Ana Karla Alves de, Arcieri, F., Baier, T., Blickle, T., Bose, J. C., van den Brand, P., Brandtjen, R., Buijs, J., Burattin, A., Carmona, J., Castellanos, M., Claes, J., Cook, J., Costantini, N., Curbera, F., Damiani, E., Leoni, M. d., Delias, P., van Dongen, B. F., Dumas, M., Dustdar, S., Fahland, D., Ferreira, D. R., Gaaloul, W., van Geffen, F., Goel, S., Günther, C., Guzzo, A., Harmon, P., Hofstede, A. t., Hoogland, J., Ingvaldsen, J. E., Kato, K., Kuhn, R., Kumar, A., La Rosa, M., Maggi, F., Malerba, D., Mans, R. S., Manuel, A., McCreesh, M., Mello, P., Mendling, J., Montali, M., Motahari-Nezhad, H. R., Zur Muehlen, M., Munoz-Gama, J., Pontieri, L., Ribeiro, J., Rozinat, A., Pérez, H. S., Pérez, R. S., Sepúlveda, M., Sinur, J., Soffer, P., Song, M., Sperduti, A., Stilo, G., Stoel, C., Swenson, K., Talamo, M., Tan, W., Turner, C., Vanthienen, J., Varvaressos, G., Verbeek, E., Verdonk, M., Vigo, R., Wang, J., Weber, B., Weidlich, M., Weijters, T., Wen, L., Westergaard, M., and Wynn, M. 2011. "Process Mining Manifesto," Springer, Berlin, Heidelberg, pp. 169–194.

Venkatesh, V., Brown, S., and Bala, H. 2013. "Bridging the Qualitative–Quantitative Divide: Guidelines for Conducting Mixed Methods Research in Information Systems," *Management Information Systems Quarterly* (37:1), pp. 21–54.

Wang, S., and Noe, R. 2010. "Knowledge Sharing: A Review and Directions for Future Research." *Human Resource Management Review* (20:2), pp. 115-131.

Wasserman, S., and Faust, K. 1994. *Social Network Analysis*: *Methods and Applications*, Cambridge: Cambridge University Press.

Zhang, X., and Venkatesh, V. 2013. "Explaining Employee Job Performance: The Role of Online and Offline Workplace Communication Networks," *Management Information Systems Quarterly* (37:3), pp. 695–722.