Team Boundary Spanning through Enterprise Social Media: Exploring the Effects of Group-Level Diversity Using a Data Science Approach

Wietske van Osch
Michigan State University
VanOsch@msu.edu

Charles Steinfield
Michigan State University
steinfie@msu.edu

Yanjie Zhao
Michigan State University
zhaoyanj@msu.edu

Abstract

Effective work groups engage in team boundary spanning, that is, the use of communication ties as conduits to critical external resources. With the proliferation of social media technologies in enterprise settings and the associated increase in visibility of communication ties, understanding their impact on boundary spanning becomes imperative to improving cross-boundary knowledge creation and management inside organizations. In this paper, drawing on log data from 415 unique work groups in an enterprise social media (ESM) system, we use a machine learning approach to automatically detect three distinct team boundary-spanning activities. Using zero-inflated poisson regressions, we further show the effect of group visibility as well as three distinct sources of group structural diversity—geographic, functional, and hierarchical—on the extent to which teams engage in boundary spanning through ESM. Implications for theory and practice around the use of data science approaches as well as visibility and diversity constructs for understanding team boundary spanning are discussed.

1. Introduction

The success of today's enterprises increasingly depends on the efficiency and quality of their cross-boundary knowledge flows and processes [1]. Various information systems, specifically emerging enterprise social media (ESM) technologies, are used to increase the transparency and openness of knowledge flows with the aim of enhancing team effectiveness, collaboration, knowledge sharing, and innovation.

Today, the proliferation of ESM technologies in organizational contexts has profound implications for such boundary-spanning knowledge flows between organizational teams; i.e., with others in the company who are outside the immediate workgroup or team ([2, 3, 4] but can be the source of relevant external resources, including informational or reputational benefits.

Social media encompass a range of information and communication tools (ICTs) for supporting interaction, collaboration, and co-creation, such as blogs, content communities, and social network sites [5,6]. Studies of organizational social media use suggest that these systems have the potential to enhance boundary-spanning knowledge flows by enabling the identification of and interaction with relevant external individuals and information [7,8,9].

This paper is motivated by two goals. First, we aim to develop and test a machine-learning algorithm to facilitate the automatic detection of three distinct types of team boundary-spanning activities, using content data from an ESM platform of a large multinational corporation. The three boundary-spanning activities include representation, coordination, and general information search [3], all of which have been associated with distinct performance benefits, both for the teams performing these activities and the organization at large. Hence, our first research question concerns: *Which intra-organizational team boundary spanning activities are enacted through ESM (RQ1)?*

Second, we aim to test the effects of work group visibility and structural diversity on the extent to which teams engage in the abovementioned three team-boundary spanning activities. Hence, the focus is on understanding whether the level of visibility of the work group—i.e., the extent to which all content created by a workgroup is visible or invisible to non-members of the group—affect their representation, coordination, and information search activities. Furthermore, beyond group visibility, we also explore whether distinct sources of work group diversity—including geographic, functional, and hierarchical diversity—affect the frequency of representation, coordination, and information search activities enacted by the workgroup. Hence our second and third research questions focus on: *What are the effects of work group visibility (RQ2) and structural diversity (RQ3) on the frequency of occurrence of each type of boundary spanning activity between workgroups inside the organization?*

1530-1605/16 $31.00 © 2016 IEEE
DOI 10.1109/HICSS.2016.274

Answering these two research questions offers a number of contributions to research and practice. First, the proposed algorithm offers researchers a behavioral approach to identifying and measuring actual rather than self-reported team boundary spanning activities. Second, findings regarding the impact of work group visibility and diversity can enhance our understanding of valuable antecedents to team boundary spanning that have not hitherto been explored in the team boundary spanning literature. Third, although the emerging ESM literature has anecdotally discussed the potential impacts of using ESM for boundary spanning, this study is among the first to explore such impacts empirically. Furthermore, with respect to practice, the insights from the proposed algorithm can assist knowledge managers in evaluating and enhancing the likelihood of cross-boundary knowledge flows.

The remainder of this paper is organized as follows. We first review the literature on team boundary spanning and structural diversity and propose a set of hypotheses regarding their links. Subsequently, we describe the case organization, data collection, the development of the machine-learning algorithm, and the approach to data analyses and hypotheses testing. We then present our findings. Finally, we discuss next steps as well as important implications for research and practice.

2. Theoretical Background

2.1. Team Boundary Spanning

Team boundary spanning–the extent to which communication links work groups to external sources of information, either within or outside the organization [10]–is closely related to other popular concepts from social network theory, including bridging or weak ties [11] and structural holes and information brokerage [12]. The common denominator across these concepts is the importance of establishing and managing external linkages as conduits to critical resources and reputational benefits.

It is important to note that our use of the team boundary spanning literature is distinct from new streams of boundary research that focus on the concept of technological gatekeepers. These new streams of research analyze how technological gatekeepers use diverse media for accessing information external to the organization, and forward to intermediaries (connectors) who translate and circulate the information for others in the organization [13,14]. As such, this work is distinct in two ways. First, much of this new work does not emphasize team-level boundary spanning inside the organization, but rather focuses on the absorption of knowledge and information residing outside the organization. Second, although the concept of technological gatekeeper focuses on technology support for information search and filtering, it does not explicitly examine the role of social media and specifically ESM for changing patterns of internal or team-level boundary spanning.

In the literature on team boundary spanning, three such activities have been distinguished [15,3], namely representation, coordination, and information search. *Representation* involves the lobbying for the team up the hierarchy in order to create favorable impressions amongst senior managers, hence, is a largely vertical form of boundary spanning [3]. This process is crucial for team performance as the creation of a favorable impression among senior management is a prerequisite for obtaining access to key resources (e.g., reputation, legitimization, higher-level commitment) and financial support needed for successful product development [16]. Representation further benefits management as they stay informed of team progress that can support higher-level planning and resource allocation decisions, which in turn, can help the organization meet external client expectations (cf., [17]).

Coordination involves the facilitation of effective decision-making and design implementation through cross-boundary strategizing, planning, and evaluation; hence it is a horizontal form of boundary spanning [3,18]. This process is crucial for team performance as it involves the aligning, negotiating, and monitoring of the efforts of individuals—within and outside the team—in order to accomplish project goals (e.g., delivery deadlines). Hence, coordination is crucial for the efficiency, effectiveness, innovativeness, and flexibility of goal delivery [19].

General information search involves the general scanning of the external team environment to gain access to relevant information, knowledge, and expertise; hence, is a largely horizontal form of boundary spanning [3]. Target actors of information search activities are often loosely coupled with the focal team [1]. This boundary-spanning process is crucial for team performance as it enables them to gain project-specific expertise and an understanding of trends, opportunities, and threats in the external environment [20].

As opposed to public social media, ESM support private communication within organizations, including interactions between organizational members in teams or groups [21], thereby mitigating some of the risks of inadvertent release of proprietary information associated with public social media [22,23]. As such, ESM offers better support for team boundary spanning processes by avoiding security and confidentiality issues. This also implies that ESM restricts boundary-spanning exchanges to intra-organizational extra-team stakeholders—individuals outside the immediate workgroup, but still within the same organization. Hence, ESM-enabled boundary-spanning interactions between work groups inside an organization—rather than between organizations—form the focal point of this study.

The proliferation of enterprise social media has the potential to enhance team boundary spanning by offering group members a relatively low effort vehicle for making work activities as well as common areas of interest visible to others in the organization, which can help organizational participants locate and connect with relevant resources outside their local environment [5,7,8,9]. Hence, the visibility affordance of ESM seems to make these novel platforms particularly useful for groups to enact these three boundary-spanning activities.

2.2. Team Visibility

In addition to enhancing the potential for boundary spanning, the visibility affordance of ESM implies that individual users or groups of user can control the level of visibility of their content and conversations.

Although most of the ESM literature has focused on the positive effects of visibility for social capital formation and knowledge sharing, recently [24] argued that visibility could actually discourage employees' willingness to post information due to concerns over loss of control, job security, or accidental knowledge spill-overs.

Hence, although overall visibility appears to have a positive effect on boundary spanning in terms of rendering visible the interests and expertise of potentially unknown others residing elsewhere in the organization, in certain cases restricting visibility to only a select number of extra-team stakeholders may be beneficial.

Based on an exploratory study of ESM-based boundary spanning, [25] propose that whereas organization-wide visibility may be useful for enacting representation—i.e., for creating broad awareness for team accomplishments—coordination and information search may require a more restricted audience to create the requisite trust and safety for cross-boundary strategizing, planning, evaluation, and knowledge transfers to occur.

Hence, in this study, we propose that:

H1a: Visibility—as measured by the openness of a team space—has a positive effect on the number of representational activities enacted by the work group

H1b: Visibility—as measured by the openness of a team space—has a negative effect on the number of coordination and information search activities enacted by the workgroup

2.3. Team Diversity

Within the literature on diversity in work groups, the primary focus has been on the consequences of structural diversity for within-group communication, conflict or social integration (c.f., [26]). However, with work groups increasingly focusing on connecting members from different geographic locations and functions and the growing need for teams to engage in boundary spanning, recent research has emphasized that structural diversity plays a more significant role when it comes to exposing group members to different sources of information, knowledge, and feedback [26].

Specifically, the literature on structural diversity distinguishes four sources of structural diversity [26], namely:

1. Geographic locations: given that the "eyes and ears" of group members are in different environments, they can access a greater variety of information and knowledge [27]
2. Functional assignments: group members belong to different functional domains, which allows them to tap into diverse social networks (c.f., [28]).
3. Reporting managers: diversity in hierarchical reporting structures of group members also increases the opportunity accessing distinct social networks [12].
4. Business units: similar to functional assignments, group members who work in different business units are able to access diverse and unique knowledge outside of the group [29].

Each of these sources of diversity have thus been shown to make it easier for teams to access external resources and knowledge, therefore resulting in the second overarching hypothesis that:

H2: Structural diversity of a work group has a positive effect on the number of boundary spanning activities enacted by the group.

However, we propose that each of these sources of diversity—geographic, functional, or hierarchical—may offer distinct benefits or challenges for the three different types of boundary-spanning activities. Specifically, given the vertical nature of representational activities, hierarchical diversity should positively affect the extent to which groups enact such activities, whereas geographic or functional diversity would seem unimportant in affecting the frequency of representational activities enacted by a workgroup.

Furthermore, work groups with high internal levels of geographic and hierarchical diversity will have a high variety of external (to the team) reporting relations, thereby requiring a greater number of coordination attempts. Hence, we anticipate that geographic and hierarchical diversity will have a positive effect on coordination. However, as emphasized originally by [3], work groups with high internal functional diversity will need to focus on overcoming and coordinating group member in order to establish a shared understanding required for project success before engaging in external coordination. Therefore, we anticipate a negative effect of functional diversity on coordination.

Finally, given that information search activities primarily occur between individuals with the same functional background but residing elsewhere in the organization, we anticipate a negative effect of functional diversity and a positive effect of geographic diversity. No effect of hierarchical diversity is anticipated. Hence:

H2a: Geographic diversity of a work group has a positive effect on the number of coordination and information search activities enacted by a work group.

H2b: Functional diversity of a work group has a negative effect on the number of coordination activities and the number of information search activities, but no effect on representation activities.

H2c: Hierarchical diversity of a work group has a positive effect on the number of representation and coordination activities enacted by the group, and no effect on information search activities.

3. Research Design

3.1. Data Collection

Data was collected from the ESM platform of a large worldwide provider of workplace products, furnishings, and services. The company has approximately 10,000 employees around the world and is headquartered in the U.S. with offices and divisions in nearly 40 countries in North and South America, Europe, Africa, Asia, Oceania, and the Middle East.

In March 2012, the organization launched an ESM tool based on the Jive Platform. Jive[1] is a provider of corporate social technologies that support business communications and collaborations among employees. Following its global launch in March 2012, the adoption and use of the system has grown substantially, with a total user base of over 9,000 users as of 2014.

For the development of the machine-learning algorithm, we collected content data from the blogs and discussion threads of 415 groups, resulting in a total of 2029 discussions and 6500 threads. Here groups are spaces created by collections of employees with the purpose of sharing information internally and/or externally through public discussion forums, public or private blogs, and document sharing platforms. The 415 selected groups were all groups created in the ESM system that had more than 1 member. For the purpose of this study, we focused on the content within discussion forums only as these were the most external-oriented spaces of groups, hence, most likely to be used for purposes of team boundary spanning. From the 415 groups, 210 groups were characterized as highly visible and 205 groups as lowly visible meaning their content can or cannot be viewed by non-members of the group respectively.

3.2. Algorithm Development

To ensure the reliable development of the machine-learning algorithm, three graduate students were trained to perform manual coding of the content data, assigning the various posts to categories reflecting the type of boundary-spanning activity each contained (or lack thereof). Coding was preceded by an elaborate training session to familiarize the coders with the coding manual and the coding scheme.

The coding manual included five coding categories, namely three categories for each of the three boundary-spanning activities—representation, information search, and coordination. Furthermore, posts that did not fall into any of these three categories were characterized as two other types of activities, namely either work-related other or non-work related other.

An example of a coded excerpt for representation is "*Here is a great video showing the work of our team in Vodafone's new workplace*". An example of an information search post: "*Has anyone worked on an innovation centre that they could share?*" An example of a coordination post would be: "*OK, CDC folks, the cancellation of the innovation center meeting threw us a little curve ball, but here's the revised planning.*"

As for the two other activity categories, the following is an example of a work-related post that is not boundary-spanning related: "*The Steelcase interns had the opportunity to participate in Chicago yesterday. It was great to see the Steelcase show so full and have such an exciting buzz around it.*" An example of a non-work related activity is the following post: "*Have you hopped on a bike lately??? If so how long and where did you ride?*"

Examples of the five types of codes are provided in Table 2 with results below.

Following the training, the coders were supervised in the independent coding of 14% of the content to compute interrater agreement. An initial interrater agreement of 89.6% with a corresponding .71 Cohen's

[1] http://www.jivesoftware.com

kappa (i.e., substantial agreement; c.f. [30]) provided confirmation of coding scheme validity and coding process reliability. Following the reconciliation of differences, the remainder of the content data for manual coding was divided across the three coders.

Within the next stage, the manually coded data was used to create an algorithm for automated text classification. The problem of text data classification belongs to the area of natural language processing, which is one of the most popular applications of machine learning. Compared to machine-learning problems that deal with numerical data, text data mining and classification is more tedious.

In this study, we used a support vector machine (SVM) learning algorithm to develop the prediction model. A support vector machine is a supervised learning algorithm. The idea of this algorithm is to construct a (set of) hyperplane(s) in a high dimensional space, which can be used to separate data samples belonging to different classes. The hyperplane chosen should have the largest distance to the nearest training data points from different classes as a larger margin will lead to lower generalization error.

SVM is particularly well-suited for text categorization for a number of reasons [31]. For example, text data has many features with each unique word as a feature, and SVM deals well with high-dimensional data since SVM offers overfitting protection. Also text data, after being processed, will generate a sparse matrix and SVM is well suited for sparsity.

Although the original SVM algorithm was invented in 1963 as a linear model, a kernel trick was later introduced to be applied to nonlinear classifiers [32]. In this project, we applied a sigmoid kernel function in the nonlinear classification model, which was determined using cross validation by comparing the performance to other kernel functions including linear and RBF functions.

Algorithm accuracy was calculated by accepting the classification result as correct if one in two labels matches the true category. The overall accuracy of tenfold cross validation of training samples for the boundary spanning algorithm is 86.2%.

3.3. Operationalization of Variables

Table 1 provides definitions and operationalizations for all variables in the regression models.

Variable	Type	Definition	Measurement/ Operationalization
Visibility	IV	The level of visibility of content created by a group in ESM to all employees and teams outside the group	0 = visible 1 = invisible
Geographic Diversity	IV	The extent to which different group members reside in different geographic locations	Entropy (0 – 1; where 0 is no diversity and 1 is maximal diversity) (see Harrison & Klein, 2007 [33])
Functional Diversity	IV	The extent to which different group members belong to different functional departments or units	Entropy (0 – 1; where 0 is no diversity and 1 is maximal diversity) (see Harrison & Klein, 2007 [33])
Hierarchical Diversity	IV	The extent to which different group members belong to different hierarchical levels	Entropy (0 – 1; where 0 is no diversity and 1 is maximal diversity) (see Harrison & Klein, 2007 [33])
Represen-tation	DV	The lobbying for the team up the hierarchy to create favorable impressions amongst senior managers	Count variable representing the number of representation instances within a particular work group
Coordination	DV	The facilitation of effective decision-making and design implementation through cross-boundary strategizing, planning, and evaluation	Count variable representing the number of coordination instances within a particular work group
Information Search	DV	The general scanning of the external team environment to gain access to relevant information, knowledge, and expertise	Count variable representing the number of information search instances within a particular work group

Table 1. Operationalization of Variables

3.4. Analysis and Hypothesis Testing

To test our hypotheses regarding the effects of group visibility and group structural diversity on each of the three boundary-spanning activities, we use Zero-inflated poisson (ZIP) regression using the PSCL package in R. Poisson regressions are best used for regression analyses in which the dependent variables are heteroscedastic and bounded by zero (i.e., count variables), such is the case with the three boundary spanning dependent variables. [34] and [35] indicated that the probability of zero in a mixed Poisson distribution is greater than the probability of zero in an ordinary Poisson distribution with the same mean, thus we applied the ZIP model which contains two parts, a poisson count model and the logit model for predicting excess zeros. The data was cleaned by excluding the groups without any interactive activity and the count was normalized and mapped to 0 to 100 based on the total discussion number generated by each group.

4. Findings

In what follows, we first discuss the findings from our machine-learning approach to detecting the distribution and frequency of the three team boundary-spanning activities, representation, coordination, and information search. Then we present the findings from the Poisson regressions for testing the hypotheses regarding the effects of group visibility and group structural diversity on the frequency of team boundary spanning activities.

4.1. Algorithm Results: The Distribution of Team Boundary Spanning Activities

Table 2 presents the distribution of the three boundary-spanning activities as enacted through ESM as well as examples of each activity. Information search is by far the dominant activity conducted through ESM, accounting for approximately 35% of all classified activities. Furthermore, representation and coordination account for ~19% and ~20% of the classified activities. Finally, 32% of the activities communicated and enacted through ESM were unrelated to boundary spanning[2].

4.2. Variable Descriptive Statistics

Table 3 presents the descriptive statistics of all variables included in the final Poisson regression model.

Variable	Type	Descriptives
Visibility	IV	Mean: 0.506 Std. Deviation: 0.500 Min: 0 / Max: 1
Geographic Diversity	IV	Mean: 0.215 Std. Deviation: 0.242 Min: 0/ Max: 1
Functional Diversity	IV	Mean: 0.481 Std. Deviation: 0.312 Min: 0/ Max: 1
Hierarchical Diversity	IV	Mean: 0.336 Std. Deviation: 0.202 Min: 0/ Max: 1
Representation	DV	Mean: 3.905 Std. Deviation: 6.931 Min: 0/ Max:22
Coordination	DV	Mean: 4.110 Std. Deviation: 5.772 Min: 0/ Max: 39
Information Search	DV	Mean: 7.193 Std. Deviation: 27.11 Min: 0/ Max: 129

Table 3. Variable Descriptive Statistics

4.3. Group-Level Antecedents: The Effects of Group Visibility and Structural Diversity

The findings of the hypotheses testing for the effects of group visibility and group structural diversity on the frequency of boundary spanning activities are summarized in Table 4 and discussed as follows.

With respect to the effect of group visibility, we found that high-visibility groups enact significantly higher numbers of representational activities (B = -.577, p = .000; where B is the Poisson regression coefficient, and p is p-values for this coefficient) confirming hypothesis 1a. Furthermore, we found that low visibility groups enact significantly higher numbers of coordination (B = .236; p = .000) and information search activities (B = .363; p = .000), providing support for hypothesis 1b.

With respect to the overall effect of structural diversity—i.e., the combined scores of geographic, functional, and hierarchical diversity—on the sum of boundary spanning activities—representation, coordination, and information search—enacted by

Categories	Representation	Coordination	Information Search	Work Related Activity	Non-Work Activity
%	19.01%	19.88%	35.05%	11.95%	21.05%
Example Coded Excerpts	Here is a great video showing the work of our team in Vodafone's new workplace.	OK, CDC folks, the cancellation of the innovation center meeting threw us a little curve ball, but here's the revised planning.	Has anyone worked on an innovation centre that they could share?	The Steelcase interns had the opportunity to participate in Chicago yesterday. It was great to see the Steelcase show so full and have such an exciting buzz around it.	Have you hopped on a bike lately??? If so how long and where did you ride?

Table 2. Boundary Spanning Frequencies and Examples of Coded Excerpts in ESM

2 Please note that the percentages exceed 100% when summed because of the use of an assembled classifier that treats activities as non-exclusive. Using an assembled classifier was preferred given the significant improvement in the reliability of performance (86.2%) compared to single classifier algorithms (67%).

Hypothesis	Poisson Regr. Coefficient/p-value/Std. Err.	Supported	Hypothesis Change
H1a	B = 0.577; p = 0.000; Std. Err. 0.116	Supported	N/a
H1b (coordination)	B = 0.-236; p = 0.000; Std. Err. 0.083	Supported	N/a
H1b (information search)	and B = -0.363; p = 0.000; Std. Err. 0.051		N/a
H2	B = -0.166; p = 0.000; Std. Err. 0.017	Not supported	We found a negative, instead of positive effect of structural diversity on number of boundary spanning activities
H2a (representation)	B = -.691; p = .000; Std. Err. = 0.117	Partially Supported	Although no effect was expected for representation, the effect was found to be negative.
H2a (coordination)	B = .012; p = .891; Std. Err. = 0.012		Positive effect for coordination was not significant
H2a (info search)	; ; B = .122; p = .036; Std. Err. = 0.058		Positive effect for information search is confirmed.
H2b (representation)	B = -1.709; p = .000; Std. Err. = 0.115	Partially supported	Although no effect was expected for representation, the effect was found to be negative.
H2b (coordination)	B = -1.540; p = .000; Std. Err. = 0.091		Negative effect for coordination was confirmed.
H2b (info search)	B = -1.350; p = .000); Std. Err. = 0.057		Negative effect for information search was confirmed.
H2c (representation)	B = 1.052; p = .000; Std. Err. = 0.197	Supported	N/a
H2c (coordination)	B = 1.223; p = .000; Std. Err. = 0.16584		N/a
H2c (info search)	B = .027; p = .794; Std. Err. = 0.10382		N/a

Table 4. Hypotheses Testing Results

work groups (H2), our findings showed that there was a negative effect on the number of boundary spanning activities enacted by diverse groups (B = -.166; p = .000). Therefore, we need to reject H2 that structural diversity has a positive effect on the occurrence of boundary spanning activities. However, looking at each of the sources of structural diversity and each of the types of boundary spanning separately generates a much more nuanced view of the relation between diversity and boundary spanning.

First, with respect to geographic diversity—the extent to which different group members reside in different geographic locations—we anticipated no effect on representation, however, there our results reveal a negative significant effect (B = -.691; p = .000). With respect to coordination, we expected a positive effect of group geographic diversity, but this effect was not significant (B = .012; p = .891). Finally, with respect to information search, we anticipated a positive effect of group geographic diversity, which was indeed found to be significant (B = .122; p = .036), thereby providing partial support for H2a.

Second, with respect to functional diversity; we again expected no significant effect on representation activities, but found the effect to be significant and negative (B = -1.709; p = .000). Furthermore, we expected a negative effect of functional diversity on coordination, which was indeed found to be significant (B = -1.540; p = .000). Similarly, for information search we expected a negative effect, which was also confirmed significant (B = -1.350; p = .000). Therefore, we found partial support for H2b.

Third, with respect to the effects of hierarchical diversity on the three boundary spanning activities, we expected a positive effect on representation activities, which was indeed found to be highly significant (B = 1.052; p = .000). Furthermore, we expected a positive effect of hierarchical diversity on coordination, which was also significant (B = 1.223; p = .000). Finally, we anticipated no effect for hierarchical diversity on information search, which was also confirmed (B =

.027; p = .794). Hence, hypothesis 2c was fully confirmed.

5. Discussion

In this paper, we first developed and tested a machine-learning algorithm to facilitate the automatic detection of three distinct types of team boundary-spanning activities by 415 unique work groups, using content data from an ESM platform of a large multinational corporation. With a reliability of 86.2%, our algorithm revealed that interactions between different teams inside the Company were primarily focused on information search (56.5%), i.e., the general scanning of the external team environment to gain access to relevant information, knowledge, and expertise. Additionally, we found that another 37% of interactions were unrelated to boundary spanning. Furthermore, we found that coordination made up an additional 19.8% of all inter-team interactions, focusing on the facilitation of effective decision-making and design implementation through cross-boundary strategizing, planning, and evaluation. Finally, the remaining 8.1% of all interactions between work groups focuses on the lobbying for the team up the hierarchy to create favorable impressions amongst senior managers. These findings are interesting as they are very different from the results from [25] who—based on a manual content analysis of two highly visible organizational work units—found that the vast majority of content was representational. It also conflicts findings in previous boundary spanning literature that has suggested that "talking up" was one of the most dominant strategies [36, 37, 14].

Following the measurement of dependent variables—representation, coordination, information search—through the machine learning approach, we next measured the effects of group visibility and group structural diversity on the frequency with which each of these boundary-spanning activities was enacted by groups. Several interested findings emerged. First, we found that high group visibility—i.e., all group content is visible to any extra-team member of the organization—is positively related to the number of representational activities; whereas it is negatively related to coordination and information search activities. This is in line with the theorizing of [25] who suggest that the visibility affordance of ESM may effectively support representational activities, but that more private communications are required for effective information search and coordination to occur.

Furthermore, we found that—although the literature on structural diversity (c.f., [26])—using self-reported data suggests that high structural diversity has a positive effect on inter-team knowledge transfer, we found that a negative relation exists between structural diversity and boundary spanning. Nonetheless, it appears that some sources of diversity may have positive effects on specific boundary spanning activities. Specifically, we find that geographic diversity has a positive effect on information search and hierarchical diversity has a positive effect on representation and coordination activities. Nonetheless, we found that functional diversity was negatively related to all three boundary-spanning activities, representation, coordination, and information search. These findings imply that dispersed work groups enact higher rates of information search and that work groups with greater hierarchical differences enact higher rates of representation and coordination.

5.1. Implications for Research and Practice

The above findings offer several contributions to team boundary-spanning research and practice. First, the proposed machine-learning algorithm offers researchers a behavioral approach to identifying and measuring actual rather than self-reported boundary spanning activities, which have so far dominated all of the boundary spanning literature to date.

Second, findings regarding the impact of work group visibility help to challenge the "ideology of openness" [24] that has dominated much of the ESM literature to date arguing that performance benefits of ESM are a direct consequent of its visibility affordance. Yet, our results show that visibility is only beneficial for one of the three boundary-spanning activities, namely representation, which only accounted for 8% of all interactions. Rather, effective coordination and information search—jointly constituting 76.3% of all work group interactions—require restricted visibility.

Third, our findings regarding group diversity can enhance our understanding of valuable antecedents to team boundary spanning. In contrast to the literature on group diversity, which suggests all sources of structural diversity to have a positive effect on information transfer, we find only limited benefit for boundary spanning, with functional diversity having no benefit at all. Furthermore, geographic and hierarchical diversity offer only limited benefit to information search and representation as well as coordination respectively.

Furthermore, not only do these findings help to advance theories of boundary spanning by providing behavioral insights into group-level antecedents of boundary-spanning activities, they further offer two practical contributions. First, the machine-learning algorithm can be used or modified by knowledge managers interested in assessing the extent to which boundary spanning occurs through ESM or other content-based and interaction-oriented platforms.

Second, our findings can inform managers of two important group-level antecedents, visibility and structural diversity, that may or may not be conducive to the occurrence of each of the boundary spanning activities. Understanding and testing these boundary-spanning antecedents helps to improve the creation and organization of work groups with the aim of enhancing the effectiveness of intra-organizational representation, information search, and coordination. Specifically, our results are useful in highlighting that different boundary-spanning activities require vastly different levels of visibility and distinct sources of diversity.

5.2. Challenges and Future Research

There are a few challenges of the existing study that provides insights into important avenues for future research. First, our dependent variables focus on te occurrence frequencies of each of the boundary spanning activities rather than measuring the success thereof. Whether or not a group enacts many representational activities does not tell us much about how successful each of these activities are. Hence, future research should (i) find algorithmic and behavioral methods to assessing the success of these three team boundary-spanning activities and (ii) explore the effects of group visibility and group structural diversity on the effectiveness rather than mere frequency of boundary spanning.

Furthermore, our machine-learning algorithm reveals that beyond the three boundary-spanning activities, inter-team interactions also encompass many activities (~30%) that do not fall into the categories of representation, coordination, and/or information search. Hence, future research should explore whether additional boundary-spanning activities exist in ESM that have not been previously identified in the knowledge management literature.

Finally, although this study aimed to address two important group-level antecedents to boundary spanning frequency, additional antecedents may exist that reside either at the work group, individual, or even organizational level. For instance, organizational culture, work group cohesion, or individual job satisfaction may all impact the extent to which groups engage in boundary spanning. However, the challenge with these variables is that they are not readily operationalized and assessed through log data, hence, would require a mix of behavioral and self-reported approaches.

6. Acknowledgement

This material is based in part upon work supported by the National Science Foundation under Grant Number IIS-1422316. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

7. References

[1] J.A. Marrone, "Team boundary spanning: A multilevel review of past research and proposals for the future," Journal of Management, 2010, 36(4), pp. 911-940.

[2] D.G. Ancona, "Outward bound: Strategies for team survival in an organization", Academy of Management Journal, 1990, 33, pp. 334-336.

[3] D.G. Ancona, and D.F. Caldwell, "Bridging the boundary" Administrative science quarterly, 1992, 37, pp. 634-665.

[4] J.A. Marrone, P.E. Tesluk, and J.B. Carson, "A multi-level investigation of antecedents and consequences of team member boundary spanning behavior," Academy of Management Journal, 2007, 50, pp. 1423-1439.

[5] P.M. Leonardi, M., Huysman, and C. Steinfield, "Enterprise social media" Journal of Computer Mediated Communication, 2013, 19(1).

[6] J.W. Treem, and P.M. Leonardi, "Social media use in organizations: Exploring the affordances of visibility, editablity, persistance, and association," Communication Yearbook, 2012, 36.

[7] J.M. DiMicco, D.R. Millen, W. Geyer, W., and C. Dugan, "Research on the Use of Social Software in the Workplace," Computer Supported Collaborative Work. San Diego, CA, USA, 2008.

[8] J.M. DiMicco, W. Geyer, C. Dugan, B. Brownholtz, and D.R. Millen. "People Sensemaking and Relationship Building on an Enterprise Social Networking Site," Proceedings of HICSS, 2009.

[9] C. Steinfield, J.M. DiMicco, N.B. Ellison, and C. Lampe, "Bowling online: social networking and social capital within the organization," Proceedings of C&T '09, ACM, New York, NY, USA, 2009, pp. 245-254.

[10] M.L. Tushman, and T.J. Scanlon "Boundary spanning individuals," Academy of Management Journal, 1981, 24(2), pp. 289-305.

[11] M.S. Granovetter, M. S., "The strength of weak ties," American Journal of Sociology, 1973, 78(6), pp. 1360-1380.

[12] R.S. Burt. Structural holes: The social structure of competition. Cambridge, MA: Harvard University Press, 1992.

[13] E. Whelan, R. Teigland, B. Donnellan, and W. Golden, “How Internet Technologies Impact Information Flows in R&D,” R&D Management, 2010, 40(4), pp. 400.

[14] E. Whelan, and R. Teigland, R., “Transactive Memory Systems as a Collective Filter for Mitigating Information Overload in Digitally Enabled Organizational Groups,” Information and Organization, 2013, 23(3), pp. 177-197.

[15] D.G. ncona, and D.F. Caldwell, “Beyond boundary spanning: Managing external dependence in product development teams,” The Journal of High Technology Management Research, 1990, 1(2), pp. 119-135.

[16] G. Grabher, “Temporary architectures of learning” Organization studies, 2004, 25(9), pp. 1491-1514.

[17] L.A. Bettencourt, S.W. Brown, and S.B. MacKenzie, “Customer-oriented boundary-spanning behaviors,” Journal of retailing, 2005, 81(2), pp. 141-157.

[18] M.A. Marks, J.E. Mathieu, and S.J. Zaccaro, “A temporally based framework and taxonomy of team processes,” Academy of Management Review, 2001, 26(3).

[19] S.A. Mohrman, S.G. Cohen, and A.M. Mohrman. Designing team-based organizations: New forms for knowledge work. San Francisco: Jossey-Bass, 1995.

[20] A.B. Hargadon, “Firms as knowledge brokers” California Management Review, 1998, 40, pp. 209-227.

[21] M.H. Jarrahi, and S. Sawyer, S. “Social Technologies, Informal Knowledge Practices, and the Enterprise,” Journal of Organizational Computing and Electronic Commerce, 2013, 23(1-2), pp. 110-137.

[22] M.M. Skeels, and J. Grudin, J. “When social networks cross boundaries: a case study of workplace use of Facebook and Linkedin,” Proceedings of GROUP 2009. ACM Press, 2009.

[23] G. Von Krogh, “How does social software change knowledge management?” The Journal of Strategic Information Systems, 2012, 21(2), pp. 154-164.

[24] J.L. Gibbs, N.A. Rozaidi, and J. Eisenberg, J., “Overcoming the “Ideology of Openness” Journal of Computer-Mediated Communication, 2013, 19(1), pp. 102–120.

[25] W. Van Osch, and C.W. Steinfield, “Boundary Spanning through Enterprise Social Software” Proceedings of the 2013 International Conference on Information Systems, Milan, Italy, 2013.

[26] J.N. Cummings, “Work groups, structural diversity, and knowledge sharing in a global organization,” Management Science, 2004, 50(3), pp. 352-364.

[27] P.R. Monge, L.W. Rothman, E.M. Eisenberg, K.I. Miller, and K.K. Kirste, “The dynamics of organizational proximity,” Management Science, 1985, 31(9), pp. 1129-1141.

[28] W.M. Cohen, and D.A. Levinthal, “Absorptive capacity: a new perspective on learning and innovation,” Administrative science quarterly, 1990, pp. 128-152.

[29] M.T. Hansen, “Knowledge networks: Explaining effective knowledge sharing in multiunit companies,” Organization science, 2002, 13(3), pp. 232-248.

[30] J.R. Landis and G.G. Koch, “The measurement of observer agreement for categorical data,” Biometrics, 1977, 33(1), pp. 159–174.

[31] T. Joachims, Text categorization with support vector machines: Learning with many relevant features. Springer Berlin Heidelberg, 1998, pp. 137-142.

[32] B.E. Boser, I.M. Guyon, and V.N. Vapnik, "A training algorithm for optimal margin classifiers", Proceedings of COLT '92, 1992, pp. 144.

[33] D.A. Harrison, and K.J. Klein, “What's the difference? Diversity constructs as separation, variety, or disparity in organizations,” Academy of management Review, 32(4), 2007, 1199-1228.

[34] W. Feller, “On a general class of “contagious" distributions,” Annals of Mathematical Statistics, 1943, 16, 319-329.

[35] J. Mullahy, “Heterogeneity, excess zeros, and the structure of count data models,” Journal of Applied Econometrics, 1997, 12, 337-350.

[36] N. Levina, and E. Vaast, “The emergence of boundary spanning competence in practice: Implication for implementation and use of information systems,” MIS Quarterly, 2005, 29(2), pp. 335-363.

[37] E. Whelan, S. Parise, J. de Valk, and R. Aalbers, “Creating Employee Networks That Deliver Open Innovation,” MIT Sloan Management Review, 2011, 53(1), pp. 37.