The Impact of Key Words on Knowledge Reuse in Emergency Management

Social Media

Full paper

Dave Yates
University of Denver
dave.yates@du.edu

Scott Paquette
McMaster University
paquets@mcmaster.ca

Abstract

This paper describes analysis of 645 blog posts from an Emergency Management blog used by the United States Department of Defense. Information technology, particularly social media such as blogs permits knowledge sharing and reuse among a diverse set of responders and staff agencies who are loosely organized around different contingency events. Knowledge seekers in these situations are faced with shared knowledge from many sources and contexts and they apply cognitive strategies for quickly evaluating that knowledge. Building on Selection theory, we hypothesize the effects key words in blog post titles may have on knowledge reuse when traditional mechanisms such as source credibility and trust are not available. Results indicate key words have a strong association with knowledge reuse; further, we classify key words and compare the types of key words typically used against the types of knowledge shared in emergency management. We discuss the implications of our results for theory and practice.

Keywords

Knowledge reuse, knowledge management, emergency management, social media, selection theory.

Introduction

Organizations, including governmental, non-governmental, and corporate, are increasingly using social media technologies to share knowledge during contingencies, emergencies, and disasters (von Krogh 2012, Leidner et al. 2009, Yates and Paquette 2011). These technologies, which include social media such as wikis and weblogs (or blogs) permit open communication, knowledge sharing, and knowledge re-use within an organization and between organizations (Majchrzak et al. 2013). Loose-formed (formal or informal) collections of organizations connected by social media who share knowledge during emergencies, contingencies, and disasters (hereafter collectively referred to as emergencies) are not uncommon (Cumbie and Sankar 2012, Pan et al. 2012). Knowledge shared between organizations in these networks informs decisions about resource allocation, response actions, coordination requirements, and future planning considerations (Shan et al. 2012, Yang et al. 2012).

Knowledge sharing can be particularly difficult during emergencies (Bharosa et al. 2010, Turoff 2002) as, emergencies are unpredictable, unfold quickly, involve a large number of people (both the affected public and agents of multiple response organizations), have short decision and reaction times for responders, much uncertainty, often a degraded physical and informational environment, and are times of high pressure and stress (Day et al. 2009, Xie et al. 2011). As a result, users often have little time and attention for sharing and reusing important knowledge, giving rise to greater confusion, misunderstanding, inefficiencies, and poor decision making. Social media employed by organizations during these responses must compensate for these limitations while simultaneously promoting effective knowledge sharing and decision making (Yates and Paquette 2011, Palen et al. 2011). Knowledge shared in emergencies is dynamic and socially constructed (Yang et al. 2012) and focused on the “critical problem at the moment”
Impact of Key Words in Emergency Management

(Turoff 2002). Yet in the dynamic and uncertain response to emergencies it is not always clear what should be shared and who needs access to it (Pan et al. 2012, Turoff 2002).

Knowledge sharers often explicitly use keywords to not only classify and describe shared knowledge, but also in an emergency to help other users identify how knowledge will may useful to their organizations. In emergencies keywords may indicate the shared knowledge is of interest to certain other responders or decision makers, that the knowledge pertains to a specific event or region, which helps busy users strategize whether they need to access and reuse that knowledge or not. Keywords are ubiquitous and have been well documented in the information retrieval literature, yet despite their popularity little is known about how keyword selection leads to reuse. Are there types of keywords that are more appropriate for describing certain types of shared knowledge? How prevalent are keywords, and do they help users select which knowledge to reuse? Our goal is to answer these questions. We hope to offer strategies for sharers and re-users to help foster increased value from shared knowledge; and, we hope to contribute to the information system and emergency management literature by offering a theoretically grounded approach to knowledge sharing that complements decision makers’ needs during emergencies.

Theoretical Background

In this research we focus on organizational knowledge sharing (Leidner et al. 2009, Majchrzak et al. 2007, Yates and Paquette 2011), since organizations own and maintain the social media systems used during disasters and organizational users are more often focused on allocating resources and making strategic decisions. By contrast, public information sharing during disasters focuses more on technological and social mechanisms for sharing (Palen et al. 2011). According to Majchrzak et al. (2007) organizational users form loosely formed networks to fulfill knowledge needs, and social media in particular is a useful platform enabling knowledge exchange between these users (Leidner et al. 2009, Pan et al. 2012, Yates and Paquette 2011). Disasters create highly dynamic sharing environments (Turoff 2002), and organizations have differing levels of pre-emergency involvement Pan et al. (2012). Thus knowledge needs emerge during the emergency and organizational contexts are fluid as organizations join the response. Where social media is flexible, more formal information structures become too slow and inflexible to support emergent knowledge reuse (Day et al. 2009). Social media enables both horizontal and vertical knowledge exchange, allowing decision makers instant access to ‘front line’ knowledge as it is shared, reducing response times (Yang et al. 2012).

However, the flexibility enabled by social media comes with a cost since knowledge sharers have little insight into how the knowledge will be used by others. Von Krogh (2012) points to quality concerns, as well as difficulty “codifying” knowledge for reuse and difficulty “representing” the right knowledge needed at the right time during the disaster. The large amounts of data that can be placed on social media must be sifted through and processed by decision makers (Yates and Paquette 2011) suggesting not all knowledge shared is necessarily useful. Knowledge sharing may even create a bottleneck in the emergency response process (Bharosa et al. 2010) if users struggle to find an effective way to evaluate shared knowledge for use (Day et al. 2015, Pan et al. 2012). Sharers need a mechanism to quickly convey the qualities of their shared knowledge that users may reliably evaluate in order to avoid these problems.

The functionalist theory of mass media (Lasswell 1948, Wright 1960) proposes a concept called focus which describes the actual content of the media. Lockwood and Dennis (2008) in a study of corporate blogs identified a number of pertinent dimensions of focus for how blog posts may be categorized. Three types of focus are applicable to emergency blogs: events, non-events, and time-based focus. Lockwood and Dennis explain events as posts which something the organization is responding to at the moment. Others looking to guide their own organizations’ emergency response may examine event posts for knowledge (Majchrzak and More 2011). The second type of focus is non-events. In contrast to events, non-events are informational posts not tied to a specific and discrete event. Non-events may describe plans, general efforts, or information related to but not directly resulting from response actions (Shan et al. 2012). As such, they will be of interest for reuse, but their direct value will be harder to realize than Events. The third type of focus is a time-based focus, a post that provides a summary of actions to date, or a regularly schedule update on capability or function or status. Timed-based posts are routine in nature, often
supporting organizational routines such as shift changes and daily briefings (Pan et al. 2012, Yates and Paquette 2011). Thus their value may be limited to narrow situations or a subset of users only.

Keywords may help indicate the value contained in the different focus types of posts by offering a user insight into the suitability of the post for reuse. The utility of keywords has mostly been explored from the context of information retrieval (Salton and Harman 2003, Tran et al. 2007), where keywords provide classification, indexing, and search terms to represent a knowledge contribution. More recent work has focused on keywords and automatic search and retrieval models (c.f. Li et al. 2008), examining the semantics of keywords, how different users represent the same meaning with keywords, and how machines can leverage keywords to retrieve the ‘right’ user-specified knowledge in a web search. However, less is known about how keywords enable knowledge management processes. For example, little work has focused on knowledge sharers deliberate attempts to direct attention to particular knowledge contributions, although keywords are commonly used for that purpose. Keywords associated with conference papers and journal articles, for instance, may be evaluated by users who are uncertain whether the main body of content is appropriate for their knowledge needs.

Two main theories may explain why this type of keyword evaluation is valuable in knowledge management. The first centers around the concept of elaboration and models of persuasion, such as the elaboration likelihood model (ELM) (Petty and Cacioppo 1986). For instance, Meservy et al. (2014) describe how users evaluate answers in online knowledge repositories and in part rely on elaborating information to filter useful from non-useful responses. Elaboration models however describe situations where the elaborating information provides an additional lens by which the main content is evaluated, providing heuristical methods to evaluate source credibility, for instance, by qualities besides the knowledge itself. In the case of users evaluating titles of blog posts in an emergency when the users may not even read the post if it does not seem applicable to them, elaboration in the manner described above is not likely to happen. Instead, users may apply selection mechanisms using the keywords in the title as cues or signals that the shared knowledge is valuable for reuse.

Selection theory (Broadbent 1958, Smith and Sewell 2013) is the second main theory which explains why keywords are valuable, and in the context of this research more accurately describes the cognitive processes at work. Whereas elaboration models focus on persuasion, selection models focus on attention. Do users have and evaluate the appropriate cues to make decisions on the appropriateness of shared knowledge? IS work building on selection theory has largely examined how users evaluate online reviews (Mudambi and Schuff 2010) and content ratings (Poston and Speier 2005). However, Butler et al. (2014) applied a selection model towards online community membership. And Lin et al. (2005) examined the signaling between sender and receiver that enable fast and efficient knowledge transfer. They particularly noted that in the knowledge transfer they studied, ELM models could not explain the phenomenon since the central route of elaboration was infeasible, as is the case when users in emergency management evaluate blog post titles.

Thus in summary, selection models that rely on keywords to cue or signal the value of the shared knowledge for reuse may help explain why keywords are valuable in the context of emergency management. In the next section we develop our model, creating a framework for keyword types and relating those types to the different knowledge foci (event, non-event, time-based) found in emergency management blog posts.

Model Development

In their study of first responders’ knowledge needs, Yang et al. (2009) described a framework of four general categories of information requirements for ongoing operations: 1) environmental conditions; 2) information on the response participants themselves; 3) status of the ongoing event such as casualties; 4) resources applied toward the response. Yang et al. (2012) noted that key information included environmental conditions, information about the responders, information on the response, and information about safety equipment. Yang et al. (2012) expanded on this classification, noting that decision makers have particular information needs beyond the current response activities such as risk
assessment and planning, thus shared knowledge must address their concerns as well as those of first responders.

These frameworks suggest that keywords whose purpose is to provide signals for selection of knowledge for reuse should follow this categorization. Namely, keywords should indicate that the shared knowledge is of interest for one of the following reasons:

1) Decision Maker Interest: the shared knowledge is primarily geared towards decision makers; it focuses on a critical change in the status of resources or capabilities, a decision that must be made, or an item flagged for high-level interest (such as a political concern).

2) Internal Interest: the shared knowledge provides information on response participants themselves, including capabilities and limitations, response actions, and span of operations.

3) External Interest: the shared knowledge provides information on the emergency itself, such as impacts of the emergency, actions of other actors besides responders (such as looting) or other external activities.

4) Environmental Interest: the shared knowledge concerns environmental or broad-reaching factors which impact the emergency response as a whole, or might be peripheral to the emergency response.

Selection theory suggests that keywords in general which serve as a signal may help users evaluate shared knowledge for reuse, thus their mere presence would drive more users to reuse knowledge because the keywords cue the selection process to occur. As long as the keywords were visible to other users, such as in the post titles of blog posts, and had applicability to different knowledge areas or contexts, they have the potential to lead to reuse. Without keywords, and in the absence of other elaborating information, users may not even apply selection criteria and thus avoid shared knowledge in general, focusing attention on other sources of information. Thus, our first hypothesis is that keywords in general, when present, lead to increased reuse.

\[ H1: \text{Posts with key words in the title will be reused to a significantly greater extent than posts without key words in the title.} \]

Our second hypothesis delves deeper into the four types of keywords specified above. Theory and prior research suggests each of these types of keywords would help users evaluate shared knowledge. Therefore, we hypothesize distinct and significant impacts of each type on knowledge reuse.

\[ H2: \text{The four types of keywords will each have a unique and significant impact on reuse.} \]

Finally, we are interested in how the types of knowledge focus interact with the types of keywords, and if certain keywords are more likely to be associated with focus types. Decision Maker keywords are those intended to draw high-level interest to critical response actions. These actions are most likely to be conveyed as Events, since Events describe the ongoing response actions. Non-events which primarily convey supporting information, and Time-based events which provide regular updates, are less likely to draw Decision Maker attention. Similarly, External keywords are most likely associated with Event focused posts. Any ongoing activities will have External interest, whereas Non-Events and Time-based events may not. Environmental keywords are also more likely to be associated with Events. Environmental interest items, which not specifically tied to response activities, are likely to be described in the context of their effect on those actions. For example, heavy rain may increase the incidence of traffic accidents or slow down responders. Finally, and by contrast, Internal keywords should be more highly associated with Non-event focused posts. Non-event posts provide supporting information about the response which more likely describes response agencies and capabilities themselves, rather than the response actions. Time-based posts, since they provide regular and expect updates, are not anticipated to be more highly associated with any particular type of keyword. Therefore, in summary, we hypothesize:

\[ H3: \text{Decision Keywords will be most highly associated with Event type posts.} \]

\[ H4: \text{Internal Keywords will be most highly associated with Non-Event type posts.} \]

\[ H5: \text{External Keywords will be most highly associated with Event type posts.} \]

\[ H6: \text{Environmental Keywords will be most highly associated with Event type posts.} \]
Analysis

We sought an emergency management blog used for a variety of contingencies and emergencies rather than a single emergency (Shan et al. 2012) to test the hypotheses. An emergency management blog operated by the U.S. Defense Information Systems Agency called the Strategic Knowledge Integration Web, or SKIWEB, was an ideal platform for the study. SKIWEB is used constantly and regularly rather than for a single activity (Yang et al. 2012). Users (military and civilian) across the Department of Defense share knowledge in real time about ongoing contingencies, disaster response, and emergency activity. This knowledge includes operations status, news items, intelligence, and anything that might be of interest to others subscribed to the blog. SKIWEB connects users worldwide in tactical (field) units, operations centers and staff agencies alike.

Users see a chronological (most recent on top) list of posts. Each listing shows the post title (which may include keywords), and the sharer’s organization and optionally their name (sharers can share as individuals or as representatives of an organization). Users then select specific entries of interest to view the entire post and access embedded knowledge or provide comments or updates. Figure 1 shows a screenshot of the SKIWEB welcome page with a link to the event log view.

Of the 645 posts, 224 or 35% had a keyword in the post title. Keywords, as previously specified, were terms that indicated what the post was about and why or how it might be of interest to others in different contexts. The other posts had titles that indicated the post purpose but not in a way that provided any specific selection criteria or additional insight as to why another user would be interested in reusing that knowledge. When keywords were present, our coders recorded them on a separate list, then afterwards examined the entire list of keywords looking for natural similarities and differences representative of unique keyword types. The easiest type of keyword to identify was those identified as Decision Maker Interest. These posts had special keywords such as Operational Report (OPREP) or Commander’s Critical Information Review (CCIR). Similarly, differentiation between Internal Interest vs. External Interest keywords was straightforward and unambiguous. Internal keywords identified news or events related to response agencies, such as Accident, or Deployment. External keywords indicated response activity...
focused outside of the responders themselves – an affected area of group, or in the case of military operations, an adversary. These keywords included Attack, Smuggling, Hostile, or a specific Country or Area of Responsibility. The most difficult keywords to differentiate (although generally the easiest to identify as keywords) were Environmental Interest. These keywords included both environmental events such as Earthquake and Storms, and also environmental conditions impactful for both internally and externally-focused activities (such as Outbreak, Flooding). Once classified, keywords were q-sorted by four other SKIWEB users following procedures in Moore and Benbasat (1991). These users were trained on the keyword types and then presented the keywords on separate cards in random order. Each sorter placed the 224 keywords in the exact same type categories, indicating 100% agreement. A full list of keywords is available but was not included due to space limitations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coding Scheme</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reuse</td>
<td>Number of times a post was read (from log)</td>
<td>19.47</td>
<td>32.30</td>
</tr>
<tr>
<td>Rank of Poster</td>
<td>1-15; 1 for military E-1 (Private or Airman Basic); 15 for military O-6 (Colonel)</td>
<td>7.82</td>
<td>3.75</td>
</tr>
<tr>
<td>Title word count</td>
<td>Number of words in post title</td>
<td>12.95</td>
<td>5.85</td>
</tr>
<tr>
<td>Event Post</td>
<td>1 if post focus was Event; 0 otherwise</td>
<td>301</td>
<td>.47</td>
</tr>
<tr>
<td>Non-event Post</td>
<td>1 if post focus was Non-event; 0 otherwise</td>
<td>229</td>
<td>.36</td>
</tr>
<tr>
<td>Time-based Post</td>
<td>1 if post focus was Time-based; 0 otherwise</td>
<td>115</td>
<td>.18</td>
</tr>
<tr>
<td>Keyword</td>
<td>1 if post title contained a keyword; 0 otherwise (examples: ‘hostile’, ‘earthquake’, ‘damage’)</td>
<td>224</td>
<td>.35</td>
</tr>
<tr>
<td>Organization Poster</td>
<td>1 if posted by an organization account; 0 if an individual</td>
<td>179</td>
<td>.28</td>
</tr>
<tr>
<td>Civilian Poster</td>
<td>1 if posted by a civilian (no military rank); 0 otherwise</td>
<td>71</td>
<td>.11</td>
</tr>
</tbody>
</table>

Table 1: Summary of variables

Results

Hypothesis testing was conducted in SPSS with Reuse as the dependent variable, following procedures in Hair et al. (2009). Hypothesis 1 and Hypotheses 2 were evaluated using multiple regression, controlling for organizational vs. individual poster, rank of the poster, and post title word length. For Hypothesis 1, we found a significant impact of keywords on reuse (β = .581, p < 0.001) as well as a significant (negative) effect of title word count (β = -0.143, p<0.01). Thus H1 was supported. For Hypothesis 2, we modeled the four distinct keyword types in the regression, and ran the regression step-wise with each of the types added first, second, third, and fourth to the model. Results were consistent across all runs. For each keyword type there were positive and significant effects on the DV after accounting for the control variables, as shown in Table 2. Thus, H2 was supported. The variables explained 53.2% of the variance in post reuse.

Hypotheses 3-6 concerned the association of the four different keyword types with the three types of focus: events, non-events, and time-based. Chi-squared tests were conducted in SPSS (Agresti and Kateri 2011) using crosstabs to examine the incidence of each type of keyword with each focus type for significance. Because the 0-0 condition (no keyword-not the selected focus) was numerically prevalent
compared to other combinations of focus and keyword we removed the posts without keywords, running
the analysis for the remaining posts and for each of the four keyword types. Table 3 shows the significant
findings.

<table>
<thead>
<tr>
<th>Keyword Type</th>
<th>β (standardized)</th>
<th>Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decision Making</td>
<td>0.398</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Internal</td>
<td>0.537</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>External</td>
<td>0.198</td>
<td>p &lt; 0.001</td>
</tr>
<tr>
<td>Environmental</td>
<td>0.108</td>
<td>p &lt; 0.01</td>
</tr>
</tbody>
</table>

Table 2. Regression Coefficients for Keyword Types on Reuse

<table>
<thead>
<tr>
<th>Focus Type</th>
<th>Event Focus</th>
<th>Non-Event Focus</th>
<th>Time-Based Focus</th>
</tr>
</thead>
<tbody>
<tr>
<td>Keyword Type</td>
<td>χ²(1) = 1.021</td>
<td>χ²(1) = 0.433</td>
<td>χ²(1) = 0.917</td>
</tr>
<tr>
<td>Decision Making</td>
<td>n/s</td>
<td>n/s</td>
<td>n/s</td>
</tr>
<tr>
<td>Internal</td>
<td>χ²(1) = 10.743*</td>
<td>χ²(1) = 13.832</td>
<td>χ²(1) = 2.256</td>
</tr>
<tr>
<td>p = 0.001</td>
<td>p &lt; 0.001</td>
<td>n/s</td>
<td></td>
</tr>
<tr>
<td>External</td>
<td>χ²(1) = 1.126</td>
<td>χ²(1) = 3.332</td>
<td>χ²(1) = 4.883*</td>
</tr>
<tr>
<td>n/s</td>
<td>p &lt; 0.001</td>
<td>p &lt; 0.05</td>
<td></td>
</tr>
<tr>
<td>Environmental</td>
<td>χ²(1) = 2.385</td>
<td>χ²(1) = 1.540</td>
<td>χ²(1) = 0.604</td>
</tr>
<tr>
<td>n/s</td>
<td>n/s</td>
<td>n/s</td>
<td></td>
</tr>
</tbody>
</table>

Table 3. Regression Coefficients for Keyword Types on Reuse

* indicates significant finding of the absence of association; that is, keyword is NOT associated with focus

The χ² test indicates the strength of the association between keyword and focus type; thus Internal
keywords are strongly and positively associated with Non-Event focus posts, meaning that they are
significantly more likely to be found in these posts than in other posts; conversely, there is a significant
negative association between Internal keywords and Event posts, indicating these keywords are NOT
found in Event posts. The only other significant association was between External keywords and Time-
Based posts, and again the External keywords were NOT associated with these posts with a high degree of
significance. The rest of the associations tested were not significant, meaning there was no strong
association between categories; i.e., a keyword was just as likely to be associated with the focus type as not
associated with the focus type. Thus, Hypotheses 3, 5, and 6 were not supported; Hypothesis 4 was
supported since Internal keywords were strongly associated with Non-Event posts (and NOT Event posts),
as expected.
Discussion and Conclusion

Results indicate strong support for the importance of keywords as selection criteria for reuse of shared knowledge. Whereas keywords are often seen as ubiquitous and universal, we have attempted to examine their utility more deeply to hopefully provide useful insights to both research and practitioners of both social media and emergency management. Our research contributes to theory by integrating selection theory, social media, and emergency management research, extending prior work and validating the four information requirements suggested in prior research through the context of keywords. We show how keywords listed in blog post titles provide signals to users for the suitability of the post for reuse, particularly important when users are time and attention constrained as they are in emergencies. Internal and Decision Making Interest keywords seem to have the strongest signal value for users, suggesting that these and not External issues are either less well known by others or of more universal interest. This suggests emergency response systems should have mechanisms encouraging knowledge sharing of this type, as well as External and Environmental information. This also suggests that knowledge management theories and research that examines the impact of mechanisms such as keywords must also consider the dynamics of the situation and the type of knowledge shared (e.g., events) when hypothesizing effects.

For practitioners, these results indicate how shared knowledge might be tagged to make it more useful for others. They also demonstrate both the practicality and necessity of keywords in blog post titles to signal the reuse value of the shared knowledge to others. There is a danger of using ‘too many’ keywords and diluting the effect. Although we did not detect this effect, future work may examine if there is such as thing as too many or too frequent keywords. If keywords are no longer distinguishable as a signal, they likely will loose their utility. At the utilization rate in this study (35%), keywords seem to provide a useful way to sift through dynamic and constantly updating emergency posts.

Despite the many limitations of this study (generalizability from this type of emergency management blog, lack of robust coding mechanisms and prior standards, and non-proportional representation of the types of keywords in the sample are but a few), we hope this provides useful analysis for future research on social media use for emergencies. As responders are critically busy during emergency response and have limited time to share knowledge, mechanisms which reliably improve the effectiveness of their efforts will strengthen the overall response effort, increasing knowledge reuse, efficiency, and decision quality.

REFERENCES