An Information Filtering and Management Model for Twitter Traffic to Assist Crises Response Coordination

Hemant Purohit\textsuperscript{a}, Andrew Hampton\textsuperscript{b}, Shreyansh Bhatt\textsuperscript{a}, Valerie L. Shalin\textsuperscript{b}, Amit Sheth\textsuperscript{a} and John Flach\textsuperscript{b}

Ohio Center of Excellence in Knowledge-enabled Computing (Kno.e.sis),
\textsuperscript{a}Department of Computer Science and Engineering,
\textsuperscript{b}Department of Psychology,
Wright State University, OH, USA

Corresponding Authors:
Hemant Purohit, Andrew Hampton
hemant@knoesis.org, andrew@knoesis.org
Kno.e.sis Center, Wright State University
3640 Colonel Glenn Highway, Dayton, OH, USA - 45324
+1-937-775-5217
ABSTRACT

Disasters such as Hurricane Sandy in 2012 result in extensive social media traffic, using networking platforms such as Twitter, as citizens report on their situations, identify needs and attempt to distribute resources. We address the challenge of finding relevant, actionable tweets from this large volume with an information filtering model. Driven primarily by concern for coordination, the initial domain independent analysis incorporates psycholinguistic theory to filter for potential messages of cooperation. The subsequent domain dependent analysis leverages a lightweight, language-driven, disaster-related domain model to extract resource references (e.g., food, shelter, etc.) in its first phase. Using a lexicon of verbs concerning the transfer of property, combined with simple syntactic frames, a second phase of domain dependent analysis assists in the identification of a particular kind of tacit cooperation, in the declarations of resource needs and availability. The results populate an annotated information repository to support the presentation of organized, actionable information nuggets regarding resource needs and availability at varying levels of abstraction. Computationally grounding the abstractions in raw data enables complex querying ability for who-what-where in coordination. Initial evaluation of the annotations relative to human judgment shows fair to good agreement. In addition to the potential benefits to the formal emergency response community of a filtered and organized corpus, the results serve as a benchmark for evaluating more computationally intensive efforts and characterizing the patterns of language behavior for coordination during a disaster.

Author Keywords

Coordination, Cooperative behavior, Distributed decision making, Emergency response, Organizational semantic analysis, Sensemaking, Targeted delivery, Twitris, Spatio-temporal-thematic analysis
[1.] INTRODUCTION

The use of social media such as Twitter during disasters, such as the 2010 earthquake in Haiti, the 2011 tsunami in Japan and Hurricane Sandy in 2012, has revolutionized the availability of information for emergency response coordination. Victims and their neighbors contribute to emergency response by sharing information (e.g., flood level, road blockages) and resources (e.g., vehicles, food, and supplies) before official help arrives. Despite the persistent citizen roles in the crisis response, command-and-control models from the traditional responder organizations do not easily accommodate the new capability (Palen and Liu 2007). One explanation is information overload; voluminous citizen postings include noise that obscures the useful content, making the information largely inaccessible to formal emergency response organizations. For example, during Hurricane Sandy, human-tagged tweets from Twitter identified a mere 4.5 thousand samples of the 2.5 million posts (0.18%) for follow up with the Red Cross (Baer 2012). Furthermore, every call of need in 2.5 million posts dwarfs the immediately available aid. While the posts also contain offers of aid, in the absence of processing these too exacerbate the overload (Fessler 2013). Social media presents the opportunity to enhance the effectiveness of resource allocation, by tapping otherwise unused resources and distribution processes in the community at large. The combination of challenge and opportunity heightens the need to study and support effective coordination between the various emergency response stakeholders and organizations leveraging social media. In particular, focusing on understanding citizen roles and needs (Palen et al. 2010) can improve information extraction to enhance coordinated emergency response.

Consistent with the work of Starbird et al. (2010), we categorize organizations based on degree of pre-established communication protocols, defined reporting hierarchies, and explicit purpose (Purohit et al. 2012a). For example, the Federal Emergency Management Agency (FEMA) in the USA is a formal organization, whose reasons for existence exclusively reflect the explicit purpose of emergency response. On the other hand, relevant informal organizations arise from agencies with independent purposes, such as manufacturing, veterinary care, restaurant services, special interest clubs, and religious congregations. Such organizations have valuable knowledge about community leaders and resources. Hybrid organizations such as the Red Cross have knowledge of the formal
system structure and defined procedures, but because they import labor from surrounding communities they lack the common knowledge and communication networks that ground the informal organizations. Multiple types of organizations combined with a severely compromised community structure create a daunting coordination problem that does not resolve with an unfiltered deluge of 2.5 million posts.

We conceptualize coordination using the dictionary^ definition -- *the harmonious functioning of parts for effective results*. Coordination can occur top-down and intentionally, in the well-defined hierarchy of a formal organization. But coordination also occurs bottom-up and fortuitously, among a loosely coupled set of independent actors constituting an emergent informal community. We argue that detecting emergent coordination in the informal citizen community is a key to filtering the relevant message traffic. Citizens contribute to coordinated emergency response in at least two ways (see Figure 1). First, they may serve as passive sensors,
Matching resource seekers to the breadth of resources providers (formal and informal) is the essence of coordinated emergency response.

The state of the art for this task of identifying and matching needs and offers is highly manual, and thus resource limited, where volunteers feed the entries for available supplies and needs based on incoming reports from volunteer teams deployed on the ground. Decision makers evaluate the entered data to distribute resources. One method for making the resource seeker apparent to the formal emergency response organization hinges on imposing low-level communication templates on the informal community, as in the Tweak-the-Tweet concept (Starbird and Stamberger 2010). We have some concerns with this approach. First, compliance is a likely problem and artificial constraints are particularly prone to failure under stress or non-standard circumstances (Dietrich 2003). Second templates fail to address the information management problem that is already apparent with standard, telephone-based methods for identifying need (Flach et al., submitted for publication). Third, predefined templates and forms always risk excluding content that fails to fit (Bowker and Star 2000). Furthermore, making the formal organization the implicit recipients of the individual resource seeker’s request for assistance risks positioning formal organizations as the bottleneck in resource distribution. As such, templates for resource seekers defined by formal organizations appear to overlook the second function of citizens, as a broadcast from resource suppliers.

We believe that the identification of needs and offers of aid can be improved substantially with the help of computational approaches to filter and sort the Twitter data. To this end, we focus on detecting and abstracting patterns of coordination in communications within the informal organization to provide a computationally-grounded boundary object of use to the formal organization (Star and Griesemer 1989). Abstraction to the pattern level shifts the initial interpretive process of the Twitter stream from purely cognitive to partially perceptual, (Bennett and Flach 1992) and provides an approach to managing data overload (Henson et al. 2012). However, lower levels of abstraction (e.g., tweets with individual requests) risk overwhelming the formal organization, while high levels of abstraction risk denying a role for human interpretation. An advantage of the computational
boundary object is that it is possible to permeate the abstraction, to audit and evaluate the data sources that contribute to the organizing representation. The right dimensions of abstraction combined with computationally supported permeability are critical to supporting coordination between the formal response organization and the various hybrid, informal and emerging organizations.

We approach the problem of analyzing and summarizing of social media traffic using the coordination analysis framework shown in the Figure 2.

**Figure 2. Coordination Analysis Framework (*for extensible metadata*)**

- Our general strategy is to cull, sort, and index a tweet corpus for eventual human review, consistent with the functions of the formal organization. We rely on a small number of modular processing steps, each consisting of limited and simple heuristics. While we avoid many of the semantic challenges of more intensive Natural Language Processing (NLP), we also acknowledge the potential of missing important
but unique content. Our view is that a smaller subset of useful, actionable information trumps a more comprehensive but noisy corpus. Here we focus on the following steps in figure 2.

- Data collection in real-time
- Corpus filtering with a domain-independent classification of coordination, based on indicators of conversational behavior.
- Semantic analysis using a disaster related domain model, primarily mined from publically available sources, to identify relevant data in the tweet and any links it contains. This includes resources, actors and spatio-temporal information.
- Limited lexical and syntactic analysis of verbs identification for a particular kind of tacit cooperation in the declarations of geographical distributions of resource needs and availability.
- Repository of annotated tweets as the feeds for a queryable, spatio-temporal-thematic visualization platform for coordination augmenting our Twitris system (Purohit and Sheth, submitted for publication).

Below we provide background material for each of these steps, (to meet space constraints the elements in grey are not covered here) following a brief introduction to Twitter.

[1.1] Twitter and Tweets

Twitter is a microblogging service that provides a social network structure and medium, allowing users to distribute short messages, called tweets. A tweet spans a maximum of 140 content characters. A tweet also includes a time-stamp, author location and an author identifier preceded by the ‘@’ symbol. Twitter distributes updates about user activities, and supports conversations as well as forwarding tweets between users. Users post updates and subscribe to (or follow) tweets from other users, thus forming social networks.

The character limit influences the scope of message content and therefore constrains communication practices. Consequently, tweets may contain URL links to web-pages, blogs, etc., sometimes employing condensed URL versions shortened by external services (e.g., http://t.co/vt8fhn7). Apart from subscription, a number of platform...
features promote engagement and distribution of information. A hashtag convention e.g., #JapanEarthquake, allows users to define searchable topics to help other users find postings of interest. Platform-supported engagement features also include Reply, Retweet and Mention. Reply is a platform-provided button to communicate with an author in response to a tweet. For example, “@daddy_yankee it’s amazing to see your participation in this hard time #sandy” is a reply to daddy_yankee. Retweet forwards a tweet from users to their followers, similar to e-mail forwarding. In so doing, the writer credits the source using the built-in ‘Retweet’ feature resulting in “RT @USER_NAME ORIGINAL_TWEET_TEXT”. For example, “RT @daddy_yankee If you need a safe place to stay, shelter info avail. in #Hurricane app http://t.co/8hxpVzC7 & http://t.co/vt8fhbn7 #Sandy”. Here a new user retweeted a tweet from the user @daddy_yankee. Mention acknowledges a user with the symbolic ‘@’ sign, but without using the built-in ‘Reply’ platform feature. For example, “Thanks @daddy_yankee, we hope this message reaches the needy ones in time #sandy”.

[1.2] Domain Independent Analysis: Conversational Coordination

Information filtering begins by detecting coordinated citizen response in social media traffic, using domain-independent linguistic properties to identify coordinated verbal exchange (Honeycutt and Herring 2009). As reported by Purohit et al. (submitted for publication?) psycholinguistic theory and conversation analyses inspire our detection of coordination (Clark and Wikis-Gibbs 1986; Goodwin and Heritage 1990; Mark 2002). Accordingly, properties of an exchange, including opening and closing phrases, anaphora and deixis, reveal the existence of conversational coordination, and hence the emergence of a new informal community. By identifying a reliable set of theoretically based indicators of conversational coordination, we obtained a bootstrapped model for classifying messages as reflecting linguistic coordination, and therefore being more likely to reflect coordinated effort in the external environment. We place this step first in the processing because it is computationally cheap and substantially reduces the corpus subject to subsequent domain dependent analysis.

[1.3] Domain Dependent Analysis: Resource Based Classification

We extend our above work on coordination detection with the ability to identify domain content relevant to emergency response. Existing efforts to identify relevant content reflect two approaches. Using a bottom-up
approach, the data drive the extraction of an ontology from the deluge of information, using Machine Learning based techniques (Imran et al., submitted for publication²). Using a top-down approach, domain experts provide a hierarchical classification scheme to organize information while also capturing relationships between the classes. Relative to the bottom up approach, the top-down approach, e.g., the Management Of A Crisis (MOAC) vocabulary (Limbu 2012) and the Humanitarian Exchange Language (HXL) standard (Keßler et al. 2013), etc., is more comprehensive and less vulnerable to bias in the data. But, it faces the complex challenge of modeling the entire domain. Further, the resulting heavy domain model imposes a computational burden that may compromise rapid classification. Because the necessary scope to support rather than supplant human sensemaking is unknown, we favor preserving real-time performance by limiting the amount of knowledge driven processing to what we actually need. Thus, we seek a lightweight model crafted to enable the addition of new concepts without disrupting an existing ontology or requiring time-consuming synchronization. We use the lightweight disaster domain model to categorize and discriminate between disaster scenarios and nominal message traffic, as well as identify resources and actors. For the present demonstration, we focus on three resource categories: food, shelter and medical needs. Of these, the medical needs are the most specific and detailed, because they are almost certainly relevant in the disaster scenario. However, detailed and specific referents for types of food or shelter (e.g., grapes or gymnasium) have less certain relevance to the disaster scenario. Thus, we endeavor to exploit a minimum, but expandable subset that provides the maximum coverage without inducing false alarms. Our current domain model includes a “shelter” subset with the words “emergency center,” “tent,” and “shelter,” along with lexical alternatives. We also included the words “food” and “water” to identify common requests for consumable resources.

[1.4] Domain Dependent Analysis: Seeker- Supplier Identification

We combine domain-independent indications of coordination with additional linguistic analysis to determine actionable relevance. Consistent with our focus on the role of seekers and suppliers in coordination, we separate the corpus into tweets indicating resource-needs and tweets indicating available resources or help. We exploit
simple heuristics, local to the tweet, with few dependencies and limited effort to invoke clarifying context. Three sets of heuristics contribute to our analysis: lexical heuristics, syntactic heuristics and spatio-temporal analysis.

[1.4.1] **Lexical Heuristics**—We rely on a lexicon of verbs to distinguish between seekers and suppliers. While it is possible to articulate need without a verb, for example by stating the noun in question (e.g., “Water!”), such formulations are potentially ambiguous regarding the seeker-supplier distinction. We focus primarily on verbs corresponding to Schank’s P-Trans primitive (Schank 1972), reflecting the transfer of property. Levin’s analysis of verbs (Levin 1993), amply grounded in over 800 citations in the scholarly literature, provides a resource for selecting these verbs. Our lexicon of seeker-supplier verbs includes the Levin categories of: give, future having, send, slide, carry, sending/carrying, put, removing, exerting force, change of possession, hold/keep, contact, combining/attaching, creation/transformation, perception, communication. The categories of slide, exerting force, combining/attaching, creation/transformation, and perception are included to test discriminant validity.

[1.4.2] **Syntactic Heuristics**—The analysis requires syntax as well as a lexicon. For example, a subject with the main verb “have” and any noun suggests a supplier. However, the same string preceded by the auxiliary verb “do” and the pronoun “you” suggests a seeker because the combination of syntax and pronoun reverses the illocutionary force through an interrogative structure. However, the abbreviated and unconstrained Twitter medium prevents reliance on punctuation for the identification of interrogatives. Pronouns and word order assist in the seeker supplier distinction associated with interrogatives, e.g., “Can you send water? (seeker) and “I can send water” (supplier). We exploit the auxiliary verbs (‘be’, ‘do’, ‘have’ as well as the modals ‘can’, ‘could’, ‘may’, ‘might’, etc.) word order (e.g. verb-subject positions) and question words (‘wh’-words and ‘how’) and the conditional (‘if’).

An exhaustive list of such limited heuristics is still subject to error, largely due to the phenomenon of indirect speech acts, which rely on shared background knowledge to re-interpret apparently factual information (Searle 1975). Accordingly, asserting a problem is a classic approach to articulating need, e.g., “it is hot in here” means...
“I need air” and/or “open the window”. Similarly, “The Red Cross can provide housing” provides a supplier fact. However, the “The Governor can provide housing” could imply a disgruntled seeker in the form of an indirect speech act, because unlike the Red Cross, the Governor does not directly supply housing. Moreover, we are unable to identify the implicit interrogative in “Sam thought that Beth had water”, which calls into question whether Beth in fact had water (Higgenbotham 1997). The factual statement could also imply that Beth is seeking water, Sam is seeking water, or the speaker is seeking water, none of which is actually asserted. However, politeness often motivates indirect speech acts (Clark 1979). Urgency, combined with the space limitations of twitter, may reduce the prevalence of indirect speech acts in the disaster scenario.

[1.4.3] Spatio-Temporal analysis—A disaster prompts a global response, much of which is irrelevant to the near-term response of the formal organization. Hence a subsequent phase of processing must address the spatio-temporal dimensions of the tweet. Proximal seekers and suppliers have more immediate utility to a particular response organization than remote seekers and suppliers. On the other hand, remote suppliers could influence longer term planning. For these reasons, we separate and organize the data on the dimensions of space and time.

[1.5] Information Visualization platform: Twitris

The resulting database of annotated tweets is potentially accessible to human analysis and query. However, the combination of volume and detail requires a combination of search and abstraction capabilities, what we are calling a permeable boundary object. We provide a prototype visualization platform for coordination, primarily to illustrate the functionality that is both likely required and feasible given a computationally accessible repository. This enables a view of information filtered by various dimension of relevance to response, e.g., by geography and time, by need category, by seeking vs. supplying nature etc. For this purpose, we use the Twitris platform (http://twitris.knoesis.org), currently in version 3 (Purohit and Sheth, submitted for publication³), which provides social media analysis along three dimensions: spatio-temporal-thematic (Nagarajan et al. 2009), people-content-network (Purohit et al. 2012b) and sentiment-emotion-subjectivity (Smith et al. 2012). Twitris presents important nuggets (weighted key-phrases) extracted from the tweet data for a chosen time and location, thus, providing
sense of community activity from a spatio-temporal perspective. Twitris also presents a cumulative version of all such information nuggets (http://twitris.knoesis.org/hurricane/topics/).

[2.] METHOD

Below we first describe the data collection method and the tweet annotation framework (see Figure 3). Then we describe processing associated with the major attributes in the annotation framework: Domain model implementation and use, the resource seeker-resource supplier analysis, the annotated repository and visualization.

[2.1] Data collection

To reflect language behavior in response to a disaster, we examined the Haitian and Japanese earthquakes and Hurricane Sandy. The Twitter Streaming API provides real-time tweet collection, including tweet text and metadata, such as timestamp, location and author information. Our crawling modules are based on the Twitter Storm distributed real-time computation system (Marz 2011), that utilizes Streaming API to provide real-time data compatible with additional real-time computation. Prior to Streaming API, Twitter provided Search API for data collection based on keyword-based search. We had used the old method for data collection during Haiti 2010 and Japan 2011 disasters (Nagarajan et al. 2009). To study specific events, our crawler in the Twitris system constantly collected the filtered stream of English language tweets from the Twitter Streaming API for event-related keywords (e.g., “hurricane sandy” for the Hurricane Sandy storm 2012) for the duration of the event period as noted in the Table 3. The initial set of keywords and hand-selected hashtags served as seed words. We then expanded the initial set by extracting the top terms in the collected data.

The goal of subsequent analysis is to filter out the irrelevant tweets, and retain a repository of annotated tweets, according to the attributes shown in Figure 3. Each tweet has an author, associated text, keywords, a date and a
location. Further analysis of the text indicates whether it corresponds to a seeker or a supplier, an organization, a location or a resource. Below we focus on resource analysis and the seeker-supplier distinction.

![Diagram of extensible metadata for an annotated tweet](image)

**Figure 3: Sample of extensible metadata for an annotated tweet**

### 2.2 Resource-based Tweet Classification

**Disaster Domain Model**

Drawing on previous interactions with domain experts (Flach et al. submitted for publication⁴), our Community Emergency Response Team (CERT) training, Rural Domestic Preparedness Consortium training and publically available references, e.g., (U.S. Department of Homeland Security 2010; FEMA 2012; OCHA 2011), we crafted a lightweight domain model that classifies various resource-needs and captures relationships between the concepts.

The full model covers disasters including earthquake, flood, political, terrorist, toxic/radioactive, weather and technological, and resources [e.g., edible, infrastructure, law enforcement]. Using a first aid handbook (Swienton and Subbarao 2012), we created an extensive “medical” subset of emergency indicators. We identified words which pertained specifically to first aid or injuries and included those words along with variations in tense (i.e., breath, breathing, breathes) and common abbreviations (i.e. mouth to mouth, mouth 2 mouth, CPR). A local expert with FEMA experience augmented the model with additional indicators and provided anecdotal context.
The current model for food, medical and shelter contains 43 concepts and 45 relationships, and shall be available at http://twitris.knoesis.org/images/datasets-and-models/.

We created the domain model in OWL using the Protégé (Stanford 2013) ontology editor. Each type of disaster is listed as an entity type with indicators for that disaster listed as individuals under a corresponding entity. Therefore a relationship is declared stating that a particular disaster, say flood, has positive indicator, flood_i, which includes all words under that heading. Each disaster has a declared negative relationship with the negative indicator list. Finally resources are declared as individuals under the appropriate entity in the same way, but relationships are not explicitly stated with any disaster in order to provide flexibility.

Using the disaster domain model to identify resources, we use a part-of-speech tagger for entities (nouns) and actionable verbs, and then identify the class of relevant resource-need based on the presence of the entities in the lexicon of the disaster domain model. To express detail tweets may include URL references to an external medium, e.g., “RT @Chillie_Mo: Overnight shelter for Hurricane victims in NY http://t.co/feLdI7hj.” The frequency of these links suggested the need to examine the relevance of the content of these external sources, using the resource-need lexicon in the disaster domain model. If the content contained such words, we considered it to be potentially useful message. But we note that URL extraction, resolution and its content analysis are computationally expensive and can be very slow. During high-volume processing this may entail a million number of URL resolutions (resolving short URLs to Long URLs and then fetching the content from URL hosts).

Mention of resources does not distinguish between seeking behavior and supplying behavior, which has utility for emergency response. Therefore, in the next section we present approaches to classify information in the resource-class related corpus for seeking and supplying behavior.

[2.3] Resource Seeking-Supplying Behavior Identification
Bottom-up analysis of the top-500 frequent tweets suggested some simple key-word heuristics for distinguishing between resource seekers and resource suppliers, as shown in Table 1.

<table>
<thead>
<tr>
<th>Behavior Type</th>
<th>Heuristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeker rules</td>
<td>looking for</td>
</tr>
<tr>
<td></td>
<td>in need of</td>
</tr>
<tr>
<td></td>
<td>“no” followed by &lt;a resource name&gt;</td>
</tr>
<tr>
<td>Supplier rules</td>
<td>providing &lt;a resource name&gt;</td>
</tr>
<tr>
<td></td>
<td>&lt;a resource name&gt; available</td>
</tr>
</tbody>
</table>

Table 1. Initial heuristic patterns for identifying seeking-supplying behaviors

However, these failed to capture the range of linguistic mechanisms for expressing need and availability. Moreover, the absence of syntax made the classification ambiguous. For example, preceding “looking for” with “If you are” changes the illocutionary force to a supplier. Thus, we developed a more systematic approach to the identification of seeker and supplier, based on a lexicon of verbs concerning the transfer of property, question words and syntax. We developed a number of simple templates that combine syntax and verbs associated with the transfer of property to suggest illocutionary force. Some of these appear in Table 2.

<table>
<thead>
<tr>
<th>Behavior Type</th>
<th>Linguistic Rules</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seeker rules</td>
<td>(Pronoun except you = yes) ∧ (need/want = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Seeker</td>
</tr>
<tr>
<td></td>
<td>(need = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Seeker</td>
</tr>
<tr>
<td></td>
<td>(please = yes) ∧ (Levin Verb = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Seeker</td>
</tr>
<tr>
<td></td>
<td>(who = yes) ∧ (has/had = yes) ∧ (Determiner = yes/no) ∧ (Adjective = yes/no) ∧ (Things = yes) → Seeker</td>
</tr>
<tr>
<td></td>
<td>(what = yes) ∧ (can/could = yes) ∧ (Pronoun/Proper Noun = yes) ∧ (do = yes) → Seeker (also a Weak Supplier)</td>
</tr>
<tr>
<td></td>
<td>(shall/will = yes) ∧ (Pronoun/Proper Noun = yes) ∧ (Levin Verb = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Seeker</td>
</tr>
<tr>
<td></td>
<td>(do/does/did) ∧ (Pronoun/Proper Noun = yes) ∧ (have/has/had = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Seeker</td>
</tr>
<tr>
<td>Supplier rules</td>
<td>(Pronoun/Proper Noun = yes) ∧ (have/has/had = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Supplier</td>
</tr>
<tr>
<td></td>
<td>(you = yes) ∧ (need/want = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Supplier</td>
</tr>
<tr>
<td></td>
<td>(if you = yes) ∧ (Pronoun except you/Proper Noun = yes) ∧ (can/could/would/should = yes) ∧ (Levin Verb = yes) ∧ (Determiner = yes/no) ∧ (Adjective = yes/no) ∧ (Things = yes) → Supplier</td>
</tr>
<tr>
<td></td>
<td>(Pronoun/Proper Noun = yes) ∧ (may/might/must/can/could = yes) ∧ (help/assist/aid/lend a hand = yes) → Supplier</td>
</tr>
<tr>
<td></td>
<td>(do = yes) ∧ (you = yes) ∧ (need = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Supplier</td>
</tr>
<tr>
<td></td>
<td>(Pronoun/Proper Noun = yes) ∧ (shall/will = yes) ∧ (Levin Verb = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Supplier</td>
</tr>
<tr>
<td></td>
<td>(shall/will = yes) ∧ (Pronoun/Proper Noun = yes) ∧ (Levin Verb = yes) ∧ (Adjective = yes/no) ∧ (Things = yes) → Supplier</td>
</tr>
</tbody>
</table>
Table 2: Linguistics based rules to identify seeking-supplying behavior. \((x = \text{yes})\) holds when the tweet contains the feature \(x\). The lowercase word \(x\) implies literal usage, e.g., ‘need/want’ implies presence of either of need or want word in the tweet. A capitalized word implies presence of any of the class of word types, e.g., ‘Adjective’ for adjectives and ‘Things’ for resources.

[2.4] Tweet Annotation and Augmented Twitris system

As shown in Figure 3, we created the annotated corpus for a disaster event by associating metadata with each tweet, encoded in the semantic web technology format RDF-- Resource description framework (W3C- RDF Working group 2004). Our metadata includes resource class, author name, time of posting, geo-location (from both tweet metadata as well as author profile location) via Geonames dataset on Linked Open Data (LOD) cloud as well as Google maps APIs, and hashtags (user classified topics), resource seeker-provider behavior as explained in the following, etc. We extend the annotation mechanism for further enrichment of the metadata such as the DBPedia entities (a Wikipedia’s semantic web version available on LOD) that helps in fetching associated information about people, places and organizational affiliation of the entities that can be leveraged for complex queries.

We created a visualization interface to present the information to coordinators in a more user-friendly way. The annotated corpus and this interface supports search/querying for complex questions, such as ‘give me a list of all the users looking for shelter in the New Jersey area’ or ‘Where is the pet-friendly shelter information?’ We are obligated by the Twitter policy not to share data in the raw form and therefore cannot currently make the query interface of the annotated information repository public. We are, however, planning to provide the research community with the tweet-id sets for the past disaster events within the confines of the Twitter policy. With this information, our research colleagues can re-crawl the data using Twitter APIs with a simple crawling script for these tweet-id sets.

[3.] EXPERIMENTAL RESULTS
We experimented with this two-step analytical model (domain independent followed by domain dependent) on a diverse set of disaster events in terms of the affected population, the degree of devastation, the nature of casualties, etc. Table 3 summarizes the data sets.

<table>
<thead>
<tr>
<th>Event</th>
<th>Period</th>
<th>#Tweets</th>
<th>#Authors</th>
<th>#Reply</th>
<th>#RT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hurricane Sandy</td>
<td>Oct 27 - Nov 7, 2012</td>
<td>4904815</td>
<td>1813730</td>
<td>7762</td>
<td>2828190</td>
</tr>
<tr>
<td>Japan Earthquake</td>
<td>Mar 11 - Mar 30, 2011</td>
<td>609853</td>
<td>26916</td>
<td>60223</td>
<td>240090</td>
</tr>
<tr>
<td>Haiti Earthquake</td>
<td>Jan 13 - Mar 10, 2010</td>
<td>583747</td>
<td>26460</td>
<td>56896</td>
<td>200955</td>
</tr>
</tbody>
</table>

Table 3: Dataset Overview

[3.1] Results

Table 4 illustrates some characteristics of the domain independent, analysis-based filtering. Table 5 presents the distribution of resource-based need classification for our data sets and Table 6 shows statistics about the seeking-supplying behavior identification for different cases of need-types for the output of domain independent analysis for a disaster event data stream.

<table>
<thead>
<tr>
<th>Event</th>
<th>#Tweets</th>
<th># D.I. Filtered Tweets</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hurricane Sandy</td>
<td>4904815</td>
<td>2834098</td>
</tr>
<tr>
<td>Japan Earthquake</td>
<td>609853</td>
<td>45057</td>
</tr>
<tr>
<td>Haiti Earthquake</td>
<td>583747</td>
<td>24367</td>
</tr>
</tbody>
</table>

Table 4: Domain Independent Analysis (D.I.) based filtering

<table>
<thead>
<tr>
<th>#D.I. Filtered Tweets</th>
<th>Sandy 2012</th>
<th>Japan 2011</th>
<th>Haiti 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>2834098</td>
<td>45057</td>
<td>24367</td>
</tr>
<tr>
<td>Shelter related</td>
<td>2330</td>
<td>64</td>
<td>130</td>
</tr>
<tr>
<td>Medical related</td>
<td>7048</td>
<td>1155</td>
<td>1323</td>
</tr>
<tr>
<td>Food related</td>
<td>10002</td>
<td>1529</td>
<td>530</td>
</tr>
</tbody>
</table>

Table 5: Domain Dependent Analysis: Resource classification results over D.I. filtered tweets
Table 6: Domain Dependent Analysis: Seeker-Supplier Identification over Domain Independent analysis

(D.I.) filtered tweets

[3.2] Visualization of Spatio-Temporal resource distribution

The visualization prototype provides several functions that hinge on the prior analyses and resulting annotated data repository. The prototype in Figures 4.1 – 4.4 employs our Twitris platform for social media event analysis (see also section 1.5). Figure 4.1 illustrates the basic capability to index tweets by geographic location.

We augment our Twitris platform functionality by presenting information in the filtered form of “slicing and dicing” from multiple dimensions. First, we use color to layer a classification of resources and needs on the data (Figure 4.2). This feature also provides viewers with the capability to detect implicit need in the statement of conditions, e.g., headaches, fainting etc. The tabs in figures 4.2 and 4.3 illustrate the specific ability to sort on resource seekers and suppliers. Figure 4.3 illustrates the potential to permeate the abstraction to reach the tag-cloud of information nuggets (shown in Figure 4.4). Thus, one can ask the question, ‘in a particular location X, users Y are asking for what kinds of resources Z’, etc. Consistent with the practice of having television sets in the emergency operations center, we also include access to the richer contextual information from Web of data—related news, blogs/articles and Wikipedia.
We make no claims concerning the optimal usability of this prototype interface. Our point is the real-time abstraction, pattern recognition, sorting and querying capability that we can support given the corresponding annotation of the underlying tweet corpus.

Figure 4.1. Visualization interface showing location points (pushpins) on the map with substantial data clusters on a chosen day
Figure 4.2. Visualization interface showing important key-phrases (tags) from the Supplier corpus for a chosen pushpin on the map.
Figure 4.3. Visualization interface showing important key-phrases (tags) from the Seeker corpus for a chosen pushpin on the map (Resources tab: for content driven influential topics in the form of Seeker-Supplier key-phrases)

Figure 4.4. Content in the bottom section for a chosen tag “hurricane sandy victims” in the tag-cloud

[3.3] Evaluation

The methods covered in this paper include two major steps: a resource-based needs classification using a domain model and seeker-supplier identification. The first step of resource-based classification performs the task of entity spotting in the tweet text based on the given sets of lexicon in the domain model for respective resource-based needs. We skip this simple keyword matching task in favor of the second step.

For the second step, we developed two sets of 1000 randomly selected tweets from the unprocessed Hurricane Sandy corpus. Using native speaker language skills, a research assistant unfamiliar with the project technology defined the illocutionary force (serving as ground truth) for each tweet as: seeker, supplier, both or none. Below we present the agreement between these ground truth assignments for the combined corpus.
Table 7: Evaluation of combined Seeker-Supplier Identification using human labeled datasets

<table>
<thead>
<tr>
<th>URL distribution for sample set A</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total tweets with URL</td>
<td>518</td>
</tr>
<tr>
<td>Total seeker tweets with URL</td>
<td>85</td>
</tr>
<tr>
<td>Total supplier tweets with URL</td>
<td>26</td>
</tr>
<tr>
<td>Total seeker tweets</td>
<td>222</td>
</tr>
<tr>
<td>Total supplier tweets</td>
<td>31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>URL distribution for sample set B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total tweets with URL</td>
<td>543</td>
</tr>
<tr>
<td>Total seeker tweets with URL</td>
<td>75</td>
</tr>
<tr>
<td>Total supplier tweets with URL</td>
<td>37</td>
</tr>
<tr>
<td>Total seeker tweets</td>
<td>290</td>
</tr>
<tr>
<td>Total supplier tweets</td>
<td>56</td>
</tr>
</tbody>
</table>

Table 8: URL distribution in the evaluation sets

Table 7 contains two types of analyses. The first type, which we call “relevance”, examines the agreement between ground truth and the algorithm, collapsing across the seeker/supplier distinction. Any mismatched rating between “none” and “seeker or supplier” constitutes an error. Similarly any combination of “seeker” and “supplier” ratings constitutes a hit. The second type specifically examines the seeker/supplier decision, by separating out only those tweets that received either a seeker or supplier label for both ground truth and the algorithm. For all analyses we present D’, which combines False Alarm and Hit data, along with Cohen’s Kappa metrics, which take into account biases in the distribution of positive and negative cases relative to ground truth.

Cohen’s Kappa suggests “fair” agreement for the relevance judgment for the full evaluation tweet set. We also split the full evaluation set into separate subsets of tweets with and without URL links (see Table 8), with the expectation that material in the URLs unavailable to the ground truth judgment was potentially increasing false alarms. However, splitting the randomly selected set into subsets that includes tweets with URLs and those without provides an unexpected improvement in both D’ and Kappa.

We examined seeker/supplier agreement using just those tweets receiving seeker/supplier judgments from both ground truth and algorithm. Because both the algorithm and ground truth had the potential to judge a tweet as both a seeker and a supplier, we required a policy to assess agreement for a “both” rating with both “seeker” and
“supplier”. We examined a liberal policy, in which the ambiguity was resolved in favor of an agreement, and a conservative policy, in which the tweets in question were deleted from the analysis. D’ and Cohen’s Kappa values are both “good”. The example tweets in the introduction to this paper illustrate this classification capability.

[5.] DISCUSSION

We have leveraged a combination of modest conversation analysis, NLP, a lightweight domain model and database technology to annotate tweets for potential relevance to assist in emergency response. Below we identify conclusions and implications regarding the computational approach, language behavior and the promise of such techniques for emergency response.

[5.1] Computational Approach

Our general strategy is to annotate incoming Tweets with metadata, calculated in near real time and stored in a database for later query. The existence of this annotated database avoids additional complex, time-consuming processing at query, thereby supporting real-time human interaction. The metadata are also critical to the ability to present abstract, permeable summaries of the twitter data, accessible for further human analysis, while sensitive to the problem of data overload.

Modular, relatively modest processes and knowledge bases generated promising results for the ability to separate relevant, actionable tweets from the noise. Our work establishes a preliminary benchmark for subsequent improvements, either employing more computationally intensive approaches, or, as we favor, a more complete set of heuristics and enriched knowledge base with an additive impact on processing time. Nevertheless, a key concern is creating false alarms with a knowledge base that is too generic. This was one motivation for developing our own knowledge base on this project, targeting domain-specific resources, and the effort to transfer property between seeker-supplier.
We have demonstrated the ability to evaluate embedded links, which is almost certainly beyond human capability in near-real time. We were surprised by a potential correlation between URL existences with a human judge. The source of this improvement is unclear. Perhaps the URL is a proxy for complex content that we have not yet captured in our rules. Or perhaps the mere presence of a URL suggests substance (as opposed to irrelevant rants). Therefore URLs affect the ground truth judgment by creating a general impression of the tweet trustworthiness, assuming restraint in the creation of URLs. If a URL simply serves as a proxy and/or creates a general impression, we may be able to use URL existence rather than expensive URL content analysis to make this content accessible for emergency response.

All of these features are consistent with the goal of developing fast, cheap, domain specific heuristics for decisions that are as good as, if not better than, thorough and slow analysis (Gigerenzer and Goldstein 1996).

[5.2] Language Behavior

We exploit the language behavior that appears in the medium, rather than impose templates suited to the recipient. Our modest incorporation of Natural Language Processing focused on the syntax pertaining to the contributions of seekers and suppliers. This is admittedly limited to factually-oriented declarations and potentially misses indirect speech acts, in which speakers indicate a problem that implies a request for assistance. We simply do not know the prevalence of indirect speech acts in these corpuses. While politeness is often a motivation for the indirect speech act, politeness seems unlikely in the disaster context. An alternative motivation is the potential absence of domain specific knowledge on the part of the seeker, who does not know what to request specifically (e.g., a HAZMAT crew, an ambulance or both). The ability to sort and search on concepts in the domain model (e.g., medical properties, shelter and food) provides some leverage for the problem of indirect speech acts. For example, we support identifying the prevalence of population complaints like a headache in a region affected by a chemical spill, or gastrointestinal problems in the case of an epidemic or contamination.
The seeker/supplier agreement was good, despite a number of limitations in the analysis. For example, our algorithms do not account for the ambiguity of seeking donations, which may function as both an articulation of need as well as a potential source of help. We are also missing rules for handling negation, although like indirect speech acts, we do not know the prevalence of negation in seeking and supplying behavior. We suspect that potential seekers are not likely to articulate the absence of need using twitter, e.g., “I don’t need shelter”. Here negation doesn’t flip seeker to supplier, but rather seeker to noise. The absence of rules for negation could be increasing our false alarms on the relevance decision.

[5.3] Domain Considerations

Part of the motivation for working with the natural data is to discover how the informal community is exploiting the medium, both for better coordinating with the informal community, and for discerning evolutions in the delivery of emergency services. In this regard, we make several observations. First, we note a substantial effort to supply resources, from both the local and global informal community. This constitutes a potentially dramatic change to the conduct of emergency response operations, which otherwise must rely on either pre-arranged delivery contracts or idiosyncratic intuition about the location of resources (Flach et al., submitted for publication¹). Currently there is no mechanism in place for someone outside of the formal emergency response organizations, for example an auto supply store, to communicate a stockpile of car chargers for cell phones. Geographic-tagging is critical to making this function useful.

We note concern for the alignment of any technology with tacit, socio-cultural assumptions about disaster response. For example, consider the apparent ratio of suppliers to seekers in twitter traffic. The informal community may not view twitter as a suitable medium for articulating individual need. More generally, the imposition of any abstraction on individual data presupposes a value system (Bowker and Star 2000). In our case, the focus on patterns of need as opposed to individual need implies a culture that values distributing resources according to where they best serve the community. This may conflict with an assumption of the informal community, particularly wealthy communities who ordinarily enjoy immediate response to a request for help, that
resources exist to address individual need on demand. However, in the crisis scenario, resources are necessarily limited, and a cry for help using either 911 or Twitter does not guarantee a timely response. Worse, persisting emphasis on the individual contact may in fact distract from the effort to maximize benefit from those resources that are available. Abstraction also reflects values in more explicit fashion, potentially interfering with the discovery of unanticipated patterns. For example, we highlight medical, shelter and food. We could include religious resources, or pet care. While we are optimistic about expanding the annotation scheme, our only response to the certain risk of an incomplete analysis is permeability of the abstraction to the raw, albeit massive, data set.

[5.] FUTURE WORK

Our future work addresses the three themes just discussed. Regarding the computational approach, we are currently working to extend our domain model with other disaster ontologies such as HXL (Keßler et al. 2013) and MOAC (Limbu 2012), specifically the inclusion of technological disasters, where computer systems in control of public utilities are compromised. We must also extend the scope of seeking behavior to include information (e.g., missing persons) in addition to the concrete resources that we currently consider. Finally, we plan to address the event evolution problem using the Continuous Semantics framework (Sheth et al. 2010), to model evolving knowledge for improving coverage in the Twitter streaming data collection process. Nevertheless, reducing false alarms despite a growing knowledge base is a key to making the medium useful. As in our previous work (Purohit et al., submitted for publication²), machine learning techniques can help us sort through our heuristics to identify the most diagnostic rules.

Work remains on matching seekers with suppliers. Presently, we rely on viewers to discern the appearance of proximity that appears at the spatio-temporal visualization interface. Mapping the location information and citizen sensed signals to create corresponding transportation paths (and delivery feasibility estimates for remote
suppliers) must play a role in this matching process. We also note that recovering location information from tweet data suffers from missing geo-location metadata of the tweets and author profiles.

We will leverage our existing ability to gather and annotate tweets to continue a characterization of the language behavior in this environment. The prevalence of indirect speech acts and the function of negation bear on the general influence of the medium on language behavior and the influence of disaster on talk. Our use of a single source of ground truth is unrealistic, if only because any human observer brings an idiosyncratic, culturally laden perspective to the problem of evaluating the pragmatics of any utterance. Furthermore, the proper perspective is not clear, as both the informal and formal response communities play a role in the interpretation.

As we improve our technical capability, empirical evaluation with users takes on greater urgency. Sufficient capability exists to support performance studies with human users. We face the usual measurement challenge of identifying suitable process and performance quality metrics in complex domains, along with the potential modifications to the task resulting from the introduction of technology. For example, abstraction and summarization masks the role of individual need and could alter decision-making in the distribution of resources. In addition, credit assignment in a complex, distributed task is not straightforward, as subtle variations apart from the function of the technology can influence process and outcome. For this reason, we will also continue to emphasize verification that the technology is performing as claimed.

[6.] CONTRIBUTIONS

We propose a linguistically inspired information filtering model for social media posts (tweets), to support the coordination of needs and available resources, including both the formal and informal communities formed during disaster. The first step uses domain independent analysis to filter out messages with potential coordination relevance. The second step uses a lightweight, domain model for message classification along two dimensions: resources and seeker-supplier. The output of these analyses resides in a semantic repository of annotated tweets.
Initial evaluation demonstrated fair to good agreement with a human reviewer. We plan to release datasets for the three major disaster events we have examined, with 6 million tweets to support benchmarking in the research community.

The domain models, processing and the resulting repository scale additively with increases in the scope of the analysis, to enable near real-time processing in the creation of the repository, and real-time response to queries anticipated in prior processing. To illustrate the engagement this database supports, we present a visualization prototype that makes activity in the informal community accessible to the formal response organization. Summaries of the data, in the form of spatio-temporal aggregations and high frequency terms categorized according to the seeker-supplier distinction, counter the potential for data overload. But, the computational foundation of these summaries allows reviewers to permeate below the summaries, to access and interpret the raw data. This technology has potential impact on the conduct of emergency response, by allowing the formal response community to focus on patterns of need in addition to individual need and by making available resources directly apparent to the formal response organization. Future work will expand the domain analysis, assess the contributions of existing processes, characterize the usage of social media and increase and evaluate capability to support the formal organization.

ACKNOWLEDGEMENT

This work is supported by the NSF (IIS-1111182, 09/01/2011 - 08/31/2014) SoCS program. We thank the NSF for their generous support, our colleagues for the valuable comments on the draft and preliminary discussion about designing a coordination analysis framework, James Gruenberg from Wright State’s Calamityville for valuable insight into disaster management, and research assistants, especially Dylan Clericus (undergraduate), Meagan Aldridge and Harry Abramovitz.

1 From Merriam-Webster, http://www.merriamwebster.com/dictionary/coordination
REFERENCES

Swienton, Raymond and Italo Subbarao (eds.) (2012). Basic Disaster Life Support, Course Manual (3rd Ed.). American Medical Association. USA.


Fessler, Pam (2013). Thanks, But No Thanks: When Post-Disaster Donations Overwhelm.


Starbird, Kate and Jeannie Stamberger (2010). Tweak the Tweet: Leveraging Microblogging Proliferation with a Prescriptive Grammar to Support Citizen Reporting. Presented at the 2010 Information Systems for Crisis Response and Management Conference (ISCRAM 2010), Seattle, WA
