What to Expect When the Unexpected Happens: Social Media Communications Across Crises

Alexandra Olteanu
Ecole Polytechnique Federale de Lausanne
alexandra.olteanu@epfl.ch

Sarah Vieweg
Qatar Computing Research Institute
svieweg@qf.org.qa

Carlos Castillo
Qatar Computing Research Institute
chato@acm.org

ABSTRACT
The use of social media to communicate timely information during crisis situations has become a common practice in recent years. In particular, the one-to-many nature of Twitter has created an opportunity for stakeholders to disseminate crisis-relevant messages, and to access vast amounts of information they may not otherwise have. Our goal is to understand what affected populations, response agencies and other stakeholders can expect—and not expect—from these data in various types of disaster situations. Anecdotal evidence suggests that different types of crises elicit different reactions from Twitter users, but we have yet to see whether this is in fact the case. In this paper, we investigate several crises—including natural hazards and human-induced disasters—in a systematic manner and with a consistent methodology. This leads to insights about the prevalence of different information types and sources across a variety of crisis situations.

Author Keywords
Social Media; Emergency Management

ACM Classification Keywords
H.3.5 Online Information Systems; K.4.2 Social Issues

General Terms
Human Factors; Measurement

INTRODUCTION
When a disaster occurs, time is limited and safety is in question, so people need to act quickly with as much knowledge of the situation as possible. It is becoming more common for affected populations and other stakeholders to turn to Twitter to gather information about a crisis when decisions need to be made, and action taken. However, the millions of Twitter messages (“tweets”) broadcast at any given time can be overwhelming and confusing, and knowing what information to look for is often difficult.

CSCW ’15, March 14 - 18 2015, Vancouver, BC, Canada

ACM 978-1-4503-2922-4/15/03$15.00
http://dx.doi.org/10.1145/2675133.2675242

One way to help those affected by a disaster to benefit from information on Twitter, is to provide an indication of *what* information they can expect to find. The capacity for affected populations to know what types of information they are likely to see on Twitter when particular kinds of mass emergencies occur, can potentially help them be more efficient in their information-seeking and decision-making processes.

To explore this idea, we collected tweets that were broadcast during 26 different crisis situations that took place in 2012 and 2013. For each crisis, we examine the types of information that were posted, and look at the sources of the information in each tweet. Our specific aim is to measure the prevalence of different types of messages under different types of crisis situations.

Our results suggest that some intrinsic characteristics of the crisis situations (e.g. being instantaneous or progressive) produce consistent effects on the types of information broadcast on Twitter. The results are of interest to members of the public, emergency managers, and formal response agencies, who are increasingly trying to understand how to effectively use social media as part of their information gathering processes.

Related Work
We know that tweets sent during crisis situations may contain information that contributes to situational awareness [49], and though disaster situations exhibit common features across various events [47], previous research has found that *information shared on Twitter varies substantially from one crisis to another* [26, 33, 34]. Indeed, some variability across disasters is expected. For instance, data from the United Nations Office for the Coordination of Humanitarian Affairs (UN OCHA) shows that disasters in high-income countries cause significantly more economic damage, but affect fewer people and have fewer fatal casualties, compared to disasters in countries with low or middle incomes [1].

Comparative research is an established discipline in communication studies [15], but to date, this method has not been extensively applied to the study of social media communications during crises. There is little overlap in the crises examined across research groups, and no attempt to date to apply the same methodology consistently to a large and diverse set of crises. The literature review by Fraustino et al. [17] indicates that research on social media during disasters “tends to examine one catastrophic event (...) and then imply that the findings are generalizable to other disasters.”

In our attempt to fill this gap, we examine tweets that were broadcast during a broad range of different crisis situations, and systematically apply the same methodology to the analysis of each event. This methodology is based on previous work that categorizes tweets by type (including [2, 7, 24, 34, 50]) or by source (including [12, 14, 27, 29, 30, 44]).

Contributions

For decision-makers and other stakeholders to be as prepared as possible, knowing what information they are likely to gain from social media can save time and help them decide where to direct their often limited resources. When stakeholders know what types of content to expect (e.g., advice, support, damage reports), and which information sources will be prevalent (e.g. news organizations, eyewitnesses, NGOs), they do not have to sift through masses of social media posts; instead, they have a reasonable expectation of what they will find, and can then make more informed decisions regarding their situational assessment process.

Based on our goal to ease the information overload wrought by social media during crisis situations, the question we address here is: *what are the similarities and differences in Twitter communications that take place during different crisis events, according to specific characteristics of such events?* To answer this question, we study the prevalence of different information types and sources found on Twitter during different types of crises, and correlate this with some of their intrinsic characteristics.

Methodology Overview

To perform this study, we employ the following methodology:

Step 1: We determine a set of dimensions that allow us to characterize different crises: hazard type, temporal development, and geographic spread.
Step 2: We determine a set of dimensions to characterize social media messages during a crisis: informativeness, information type, and source.
Step 3: We collect Twitter data corresponding to 26 crises that took place in 2012 and 2013, using retrospective sampling on the 1% public data stream which is publicly available in the Internet Archive.[1]
Step 4: We create, run, and evaluate a series of crowdsourcing tasks to perform content annotation on approximately 1,000 messages from each of the crises.
Step 5: We perform a statistical analysis of the dependencies between types of crises and types of messages.

STEP 1: DETERMINING CRISIS DIMENSIONS

Given that our research question connects two domains: disaster studies, and social media content analysis, the framework we use is composed of two parts. We categorize the crises according to a series of dimensions that characterize them. Next, we annotate tweets from each crisis according to dimensions that characterize different types of content.

When considering how to organize our data and approach our annotation process, we turned to dimensions used in the sociology of disaster research (p. 50 in [36]). For each crisis, we consider hazard type (natural vs. human-induced), sub-type (e.g. meteorological, hydrological, etc.), temporal development (instantaneous vs. progressive), and geographic spread (focalized vs. diffused).

C1. Hazard type

Hazard type is the first dimension we examine that may impact the types of contents disseminated through social media. The specific hazard types we consider are based on two taxonomies used in Europe[2] and the US,[3] as well as the traditional hazard categories listed by Fischer [16].

The first distinction is between those that are *natural* and those that are *human-induced*. Sub-categories and examples of each one are listed in Table 1. All sub-categories are covered by crises analyzed in this study, with the exception of the "biological" category, which we were unable to sufficiently account for regarding Twitter communications.

Table 1. Hazard categories and sub-categories.

Category	Sub-category	Examples
Natural	• Meteorological • Hydrological • Geophysical • Climatological • *Biological (N/A)*	• tornado, hurricane • flood, landslide • earthquake, volcano • wildfire, heat/cold wave • *epidemic, infestation*
Human-Induced	• Intentional • Accidental	• shooting, bombing • derailment, building collapse

C2. Temporal Development

When considering the temporal development of crises, we classify them as *instantaneous* (e.g. an earthquake or a shooting), or *progressive* (e.g. a hurricane or a heat wave) [3, 8, 35]. As we qualitatively coded the temporal aspects of the crises, we labeled a disaster *instantaneous* if it "does not allow predisaster mobilization of workers or pre-impact evacuation of those in danger," and *progressive* if it is "preceded by a warning period" [3].

C3. Geographic Spread

We look at the geographic spread of a crisis, and specify if it is *focalized* (such as a train accident) or *diffused* (such as a large earthquake) [3, 38]. A focalized crisis affects and mobilizes response in a small area, while a diffused disaster impacts a large geographic area and/or mobilizes national or international response.

We recognize that this list of crisis dimensions is not exhaustive. In particular, linguistic and cultural differences have been shown to influence message content, and the adoption of certain conventions in Twitter, e.g. [20, 37]. We also recognize that these dimensions are not independent from one another. For instance, with the exception of war and large-scale nuclear disasters, most human-induced crises tend to be focalized, while meteorological hazards are often diffused. Additionally, the interplay between these dimensions may yield complex results in terms of the types of information included

[1] https://archive.org/details/twitterstream

[2] http://www.emdat.be/classification

[3] http://www.ready.gov/be-informed

in Twitter messages, and the source of that information. For example, hazard type combined with geographic spread can affect public access to firsthand information about a crisis.

STEP 2: DETERMINING CONTENT DIMENSIONS

When assessing the tweets that were broadcast during each disaster event, we turned to previous research on information broadcast via social media in disaster. We constructed a coarse-grained categorization that covers the categories of information that are highly represented in previous work (including [7, 24, 29, 49, 50] among others). Due to the large number of events and messages we consider, and the limitations of using crowdsourcing workers to perform the annotation (as opposed to experts, who would be prohibitively expensive at this scale), we formulated basic information categories broad enough to be applicable to different crisis situations. The resulting categories and the previous research represented by them, are shown in Table 2: *informativeness*, *information type*, and *source*.

M1. Informativeness

We recognize that informativeness is a subjective concept, as it depends on the person who is asking for or receiving information. In addition, as with any communication, the context in which the information exchange is taking place is critical to understanding its implications. We capture this dimension following [50], by checking whether the tweet contributes to better understanding the situation on the ground. Accordingly, we use the following annotation options:

A. Related to the crisis and informative: if it contains useful information that helps understand the crisis situation.
B. Related to the crisis, but not informative: if it refers to the crisis, but does not contain useful information that helps understand the situation.
C. Not related to the crisis.

M2. Information Type

As we closely analyzed a set of samples of messages communicated via Twitter during disasters, we found that the type of content often varies substantially across hazards; a finding corroborated by many other studies [2, 7, 24, 34, 50].

To identify a set of broad categories whose incidence (though with different degrees of occurrence) is to a large extent independent of event specificities, and to obtain a manageable coding scheme, we first identified the list of information categories used in related work studying various types of events (e.g., wildfires [50], drug wars [29], floods [6, 45], earthquake [39], nuclear power plant [46], to name a few). Then, we proceeded with merging in a bottom-up fashion those categories that overlap and/or are related. Finally, we gathered the remaining categories, typically accounting for information specific to each crisis or type of crisis (e.g., flood level, weather, wind, visibility [50]), into a "catchall" category—*other useful information*. The exact matching of information types present in the related work to each of the categories used in this paper is depicted in Table 2. The information types that we use are:

Table 2. Typologies of content used in this paper, and their relationship to some aspects mentioned in previous work

This work	Related categories from previous work
Informativeness:	
Informative	informative (direct or indirect) [24]; curating or producing content [29]; contribute to situational awareness [50]; situational information [40]; contextual information to better understand the situation [42]
Not inform.	trolling [29]; humor [28]; off-topic [34, 39, 45]; rumor [22]; humor or irrelevant/spam [42]
Information type:	
Affected individuals	medical emergency, people trapped, person news [7]; casualties (and damage), people missing, found or seen [24]; reports about self [2]; fatality, injury, missing [49]; looking for missing people [39];
Infrastruc. & utilities	(casualties and) damage [24]; reports about environment [2]; built environment [49]; damaged, closures and services [22]; collapsed structure, water shortage/sanitation, hospital/clinic services [7]; road closures and traffic conditions [48];
Donations & volunteer.	donations of money, goods or services [24]; donations or volunteering [34]; requesting help, proposing relief, relief coordination [39]; donations, relief, resources [22]; help and fundraising [6, 40]; shelter needed, food shortage/distribution [7]; volunteer information [50]; help requests [2]
Caution & advice	caution, advice [24]; warnings [2]; advice, warnings, preparation [34]; warning, advice, caution, preparation [50]; tips [28]; safety, preparation, status, protocol [22]; preparedness [51]; advice [6]; advice and instructions [40]; predicting or forecasting, instructions to handle certain situations [42];
Sympathy & emo. sup.	concerns and condolences [2]; gratitude, prayers [34]; emotion-related [39]; support [22]; thanks and gratitude, support [6, 40];
Other useful info.	fire line/emergency location, flood level, weather, wind, visibility [50]; smoke, ash [48]; adjunctive and meta-discussions [40]; other informative messages [34]; information verification, explanation of particular problems [42];
Source:	
Eyewitness	citizen reporters, members of the community [29]; eyewitnesses [6, 14, 27, 34]; local, peripheral, personally connected [45]; local individuals [43, 50]; local perspective, on the ground reports [46]; direct experience (personal narrative and eyewitness reports) [40]; direct observation, direct impact, relayed observation [48];
Government	(news organizations and) authorities [29]; government/administration [34]; police and fire services [22]; police [13]; government [6]; public institutions [46]; public service agencies, flood specific agencies [45];
NGOs	non-profit organizations [12, 46]; non-governmental organization [34]; faith-based organizations [45];
Business	commercial organizations [12]; enterprises [46]; for-profit corporation [34];
Media	news organizations (and authorities), blogs [29]; journalists, media, and bloggers [12, 14]; news organization [34]; professional news reports [28]; media [6]; traditional media (print, television, radio), alternative media, freelance journalist [46]; blogs, news-crawler bots, local, national and alternative media [45]; media sharing (news media updates, multimedia) [40];
Outsiders	sympathizers [27]; distant witness [9]; remote crowd [43]; non-locals [45, 46].

A. Affected individuals: deaths, injuries, missing, found, or displaced people, and/or personal updates.
B. Infrastructure and utilities: buildings, roads, utilities/services that are damaged, interrupted, restored or operational.
C. Donations and volunteering: needs, requests, or offers of money, blood, shelter, supplies, and/or services by volunteers or professionals.
D. Caution and advice: warnings issued or lifted, guidance and tips.
E. Sympathy and emotional support: thoughts, prayers, gratitude, sadness, etc.
F. Other useful information not covered by any of the above categories.

M3. Source

When people turn to Twitter to learn about a disaster, they are often concerned with the source of information. Hence, we focused on *content source*, which may be different from tweet author; e.g. if the Twitter account of a large media organization quotes a government official, the "source" is the government official. Sources are categorized as: *primary sources* (eyewitness accounts) or *secondary or tertiary sources* (typically mainstream media or others engaged in journalistic acts) [12, 14, 27, 29, 30, 44].

For the former, we chose to broaden the definition of an eyewitness account as originating from "a person who has seen something happen and can give a first-hand description of it"[4] to also accommodate those cases when the account does not include a direct observation, yet the user is personally impacted by the event, or it "is about a direct observation or impact of a person who is not the micro-blogger" [48]—typically relaying the observations of friends or family.

In the latter case, we can find several organizations who often aggregate information about a crisis, including business, governmental, and non-governmental sources:

A. Eyewitness: information originating from eyewitnesses of the event or of response/recovery operations, or from their family, friends, neighbors, etc.
B. Government: information originating from the local or national administration.
C. Non-governmental organization: information originating from NGOs.
D. Business: information originating from for-profit business (except news organizations).
E. Traditional and/or Internet media: information coming from sources such as TV, radio, news organizations, web blogs, or journalists.
F. Outsiders: information originating from individuals that are not personally involved/affected by the event.

STEP 3: DATA COLLECTION

List of Events

Table 3 shows our datasets, which are available for research purposes at http://crisislex.org/. They correspond to a set of 26 events during 2012 and 2013, and which spawned significant activity on Twitter. Table 3 also includes crisis dimensions of hazard type, development, and spread (we consider the Singapore haze to be partially human-induced due to intentional fires to clear land). We note that in our dataset, all human-induced crises are focalized and instantaneous, while all natural hazards are diffused, but may be instantaneous or progressive.

To obtain our list of events, we started with a set of disasters compiled mainly from Wikipedia.[5] We then filtered it by choosing events that had at least 100,000 tweets associated with them—which is reflected by at least 1,000 tweets in the 1% public data stream we used.

Floods are the most frequent type of natural hazard in our data, and also the natural hazard that affects the most people in the world. According to data from the United Nations for the 2002–2011 period, an average of 116 million people were affected by a flood every year, followed by 72 million people a year affected by drought, 40 million by storms, 9 million by extreme temperatures, and 8 million by earthquakes [1].

Data Sampling

Our data collection method is shaped by limitations to data access through Twitter, and is based on first collecting a base data sample and then retrospectively sub-sampling it. The base data sample was obtained by constantly monitoring Twitter's public stream via Twitter's Sample API, which consists of a sample of approximately 1% of all tweets[6] and it is accessible via Internet Archive[7], allowing full reproducibility of this work. In the 2012-2013 period, this collection contains on average about 132 million tweets (amounting to 38 GB of compressed data) per month. The quality of Twitter data samples acquired via the publicly available APIs that offer limited access to the full Twitter stream has been studied extensively, to understand the nature of the biases of such data samples [18, 19, 25, 31, 32]. Yet, while [32] have shown biases with respect to hashtag and topic prevalence in the Streaming API (which we do not use in this study), [31] shows that the data obtained via the Sample API closely resemble the random samples over the full Twitter stream, which corroborates the specifications of this API. Additionally, given the daily volume of tweets "the 1% endpoint would provide a representative and high resolution sample with a maximum margin of error of 0.06 at a confidence level of 99%, making the study of even relatively small subpopulations within that sample a realistic option" [18].

The sub-samples are obtained by running keyword searches over the base data—keyword searches that mimic the way in which Twitter does keyword tracking to obtain a sample of the data that one can obtain in real time.[8] An advantage of this retrospective sampling method is that one can capture the entire period of the event, which is not the case for other

[4] http://www.oxforddictionaries.com/definition/english/eyewitness

[5] From the list of significant events per month, e.g. for January 2013 we consulted http://en.wikipedia.org/wiki/January_2013

[6] https://dev.twitter.com/docs/api/1.1/get/statuses/sample

[7] https://archive.org/details/twitterstream

[8] https://dev.twitter.com/docs/streaming-apis/parameters#track

Table 3. List of crises studied, sorted by date, including the duration of the collection period for each dataset, the number of tweets collected, and several dimensions of the crises

Year	Country	Crisis Name	Days	Tweets	Hazard category	Hazard subcategory	Hazard type	Development	Spread
2012	Italy	Italy earthquakes	32	7.4K	Natural	Geophysical	Earthquake	Instantaneous	Diffused
2012	US	Colorado wildfires	31	4.2K	Natural	Climatological	Wildfire	Progressive	Diffused
2012	Philipinnes	Philipinnes floods	13	3.0K	Natural	Hydrological	Floods	Progressive	Diffused
2012	Venezuela	Venezuela refinery explosion	12	2.7K	Human-induced	Accidental	Explosion	Instantaneous	Focalized
2012	Costa Rica	Costa Rica earthquake	13	2.2K	Natural	Geophysical	Earthquake	Instantaneous	Diffused
2012	Guatemala	Guatemala earthquake	20	3.3K	Natural	Geophysical	Earthquake	Instantaneous	Diffused
2012	Phillipines	Typhoon Pablo	21	1.9K	Natural	Meteorological	Typhoon	Progressive	Diffused
2013	Brazil	Brazil nightclub fire	16	4.8K	Human-induced	Accidental	Fire	Instantaneous	Focalized
2013	Australia	Queensland floods	19	1.2K	Natural	Hydrological	Floods	Progressive	Diffused
2013	Russia	Russian meteor	19	8.4K	Natural	Others	Meteorite	Instantaneous	Focalized
2013	US	Boston bombings	60	157.5K	Human-induced	Intentional	Bombings	Instantaneous	Focalized
2013	Bangladesh	Savar building collapse	36	4.1K	Human-induced	Accidental	Collapse	Instantaneous	Focalized
2013	US	West Texas explosion	29	14.5K	Human-induced	Accidental	Explosion	Instantaneous	Focalized
2013	Canada	Alberta floods	25	5.9K	Natural	Hydrological	Floods	Progressive	Diffused
2013	Singapore	Singapore haze	19	3.6K	Mixed	Others	Haze	Progressive	Diffused
2013	Canada	Lac-Megantic train crash	14	2.3K	Human-induced	Accidental	Derailment	Instantaneous	Focalized
2013	Spain	Spain train crash	15	3.7K	Human-induced	Accidental	Derailment	Instantaneous	Focalized
2013	Phillipines	Manila floods	11	2.0K	Natural	Hydrological	Floods	Progressive	Diffused
2013	US	Colorado floods	21	1.8K	Natural	Hydrological	Floods	Progressive	Diffused
2013	Australia	Australia wildfires	21	2.0K	Natural	Climatological	Wildfire	Progressive	Diffused
2013	Phillipines	Bohol earthquake	12	2.2K	Natural	Geophysical	Earthquake	Instantaneous	Diffused
2013	UK	Glasgow helicopter crash	30	2.6K	Human-induced	Accidental	Crash	Instantaneous	Focalized
2013	US	LA Airport shootings	12	2.7K	Human-induced	Intentional	Shootings	Instantaneous	Focalized
2013	US	NYC train crash	8	1.1K	Human-induced	Accidental	Derailment	Instantaneous	Focalized
2013	Italy	Sardinia floods	13	1.1K	Natural	Hydrological	Floods	Progressive	Diffused
2013	Phillipines	Typhoon Yolanda	58	39.0K	Natural	Meteorological	Typhoon	Progressive	Diffused

collections built during the disasters, which generally lack the first minutes or hours of the event.

Keywords were selected following standard practices commonly used for this type of data collection [6, 23, 34, 46], and typically include hashtags or terms that pair the canonical name of the disaster with proper names of the affected locations (e.g., Manila floods, #newyork derailment), the proper names of the meteorological phenomena (e.g., Hurricane Sandy), or, at times, hashtags promoted by governments, response agencies, or news media. Previous work has shown that this method produces a sample of messages whose distribution of information categories closely resembles the sampling by other methods e.g. *geofencing*, which samples all tweets from users in the affected area [34].

To identify the keywords/hashtags used during each event, one of the authors used a search engine to lookup for "Hashtags ⟨Event Name⟩." The search results often included news articles discussing the social media use during the searched event,[9] resources from NGOs using social media for crisis management,[10] Internet media platforms,[11] governmental resources on social media use[12] or research papers [6, 34]. Using these resources, we built an initial list of hashtags/keywords, which we further validated and iteratively improved by manually searching for them on Twitter. In those cases in which the hashtag/keyword had been used for other purposes, we also looked for the combination of the hashtag/keyword, and the event name. When other keywords frequently appear with those already on our list, we also searched for them in Twitter. If there were at least few instances in which they appeared in relevant tweets without the other keywords, we added them to the list. The size of the resulting keywords lists vary, yet [34] suggests that keywords lists of various sizes retrieve collections which exhibit comparable representativeness with respect to a reference sample.

For the *instantaneous* hazards we start the collection from the moment when the event happen, while for the *progressive* hazards we start from the moment the hazard was detected (e.g., when a storm formed for a hurricane). The volume of tweets in each collection decreases after onset, but we continue collecting data until that volume stabilizes to a low value (specifically, when the standard deviation of the daily number of tweets becomes less than 5).

As a post-processing step, we remove very short tweets (i.e. those made up of 3 tokens or less), as they are in general hard to classify and rarely contain any useful information. We do not remove near-duplicates or re-tweets (RTs) because we are interested in the extent to which people repeat and pass along existing messages.

STEP 4: CROWDSOURCED DATA ANNOTATION

We employed crowdsource workers to perform manual annotation of our datasets in April and May 2014.[13] The workers

[9] http://mashable.com/2012/06/29/colorado-wildfire-social-media/, http://www.techinasia.com/singapore-haze-infographic/ and others.

[10] http://wiki.crisiscommons.eu/wiki/Crises, http://crisiswiki.org/

[11] http://twitchy.com/2014/07/07/earthquake-hits-southern-mexico-and-guatemala-fatalities-damage-reported-pics/, https://storify.com/ABC13Houston/plant-explosion-in-west-texas

[12] http://www.gov.ph/2013/11/09/online-efforts-for-typhoon-yolanda/

[13] We employed workers through the crowdsourcing platform *CrowdFlower: http://crowdflower.com/*

were provided with detailed instructions and examples of correctly labeled tweets, so they could successfully complete the annotation task.

Task Description

Below are the instructions given during the annotation phase to crowdsource workers. "You," in the task description refers to the crowdsourcing worker. The underlined parts, and the examples, changed for each crisis.

M1. Informativeness. The instructions used for this annotation task are shown below, and include examples for each class.

> Categorize tweets posted during the 2013 Colorado floods. Please read them carefully, following links as necessary, and categorize them as:
> A. Related to the floods and informative: if it contains useful information that helps you understand the situation:
> – "RT @NWSBoulder Significant flooding at the Justice Center in #boulderflood"
> – "Flash floods wash away homes, kill at least one near Boulder via @NBCnews"
> B. Related to the floods, but not informative: if it refers to the crisis, but does not contain useful information that helps you understand the situation:
> – "Pray for Boulder, Colorado #boulderflood"
> C. Not related to the floods:
> – "#COstorm you are a funny guy lol"
> D. Not applicable; too short; not readable; or other issues.

M2. Information Type. Instructions and examples:

> Categorize tweets posted during the 2012 Colorado wildfires. Please read them carefully, following links as necessary, and categorize as:
> A. Affected individuals: information about deaths, injuries, missing, trapped, found or displaced people, including personal updates about oneself, family, or others.
> – "Up to 100,000 people face evacuation in Colorado"
> B. Infrastructure and utilities: information about buildings, roads, utilities/services that are damaged, interrupted, restored or operational.
> – "Officials working the #HighParkFire confirmed that several roads are closed"
> C. Donations and volunteering: information about needs, requests, queries or offers of money, blood, shelter, supplies (e.g., food, water, clothing, medical supplies) and/or services by volunteers or professionals.
> – "#Offer Storage Space http://t.co/... #COwildfire"
> D. Caution and advice: information about warnings issued or lifted, guidance and tips.
> – Wildfire warnings issued for six counties Sunday - http://t.co/..."
> E. Sympathy and emotional support: thoughts, prayers, gratitude, sadness, etc.
> – "Pray for Boulder #COwildfire"
> F. Other useful information NOT covered by any of the above categories.
> – "To track fire activity in CO, check this site @inciweb Colorado Incidents http://t.co/..."
> G. Not applicable; not readable;not related to the crisis.

M3. Source. Instructions and examples:

> Categorize tweets posted during the 2013 Queensland floods (Australia). Please read them carefully, following links as necessary, and indicate the most likely source of information for them as:
> A. Eyewitness: if the information originates from eyewitnesses to the event or to response/recovery operations, or from their family, friends, neighbors, etc. :
> – "Just found out my mum is trapped at home, no water, no power, tree's down across roads out of her property near glasshouse mtns"
> – "Outside sounds like it is going to shatter my bedroom windows any sec now #bigwet #qld"
> B. Government: if the information originates from national, regional or local government agencies, police, hospitals, and/or military.
> – "PRT @theqldpremier: UPDATE SCHOOL CLOSURES: An updated school closures list is available now at http://t.co/..."
> C. Non-government: if the information originates from non-governmental and not for profit organizations such as RedCross, UN, UNICEF, etc.
> – "RT @RedCrossAU: Everyone affected by #qldfloods, let people know you're safe: http://t.co/..."
> D. Businesses: if the information originates from for-profit business or corporations such as Starbucks, Walmart, etc.
> – "RT @starbucks: With many partners impacted by OLD floods, consider making (or increasing) donations"
> E. Traditional and/or Internet news or blogs: if the information originates from television channels, radio channels, newspapers, websites or blogs such as CNN, KODA, New York Times, etc.
> – "RT @ABCNews24: #QLDfloods watch: Authorities are preparing for tornadoes in southeast Queensland."
> F. Outsiders: if the information originates from individuals that have NO acquaintances affected by the event, nor are they associate with any organization.
> – "RT @TheBushVerandah: Just heard a farmer had to shoot approx 100 sows at mundubbera ... In preference to them drowning"
> G. Not applicable; not readable; not related to the crisis.

Task Characteristics

For all annotation tasks, we provide examples both in English and the language most commonly used to communicate about the event (if there was a common language used other than English.) Regarding worker selection, the platform we used for crowdsourcing allows us to select workers by country (but not at a sub-country level), so we specified that workers must be from the country where the event took place. In few cases when there were not enough workers to perform the task, we also included workers from neighboring countries having the same official language. We selected workers in this way to ensure that they understand the tweets posted by individuals local to the event, and that they would be more likely able to understand dialects, references to regional and/or local places, and overall be versed in the culture of the area in which the event took place. Additionally, following standard guidelines from this crowdsourcing platform, 20 to 30 tweets per crisis and task were classified by the authors of this paper. We consider all workers whose assessments differ significantly from ours (less than 70% of agreement) as *untrusted*.

Workers were presented with the tweet text, including any links (which they were invited to follow), and then asked to choose a single category that best matched the content of the tweet. To avoid potential ethical concerns on behalf of Twitter users who are likely unaware that their tweets are being collected and analyzed, workers did not have access to the author username, nor the time at which the tweet was sent. In addition, we avoid possible privacy violations by not displaying the username nor the profile picture of persons affected by a given disaster. This practice follows customary procedures used for using crowdsourced annotation of text messages for both information type [4, 10, 24, 34] and information source [14, 34].

Trusted workers took from 10 to 12 seconds to label each tweet (in terms of interquartile mean, which is the figure reported by the crowdsourcing platform). We collect labels from at least 3 different trusted workers per tweet and task, and determine the final label of the tweet by simple majority.

About 15-20 trusted workers participated in each classification step (i.e. a set of 1,000 tweets from a single event and with a single question M1, M2, or M3), with the bulk of the work being done by about 10 of them in each case—with no worker labeling more than 300 items in a classification task, a limit set by us following recommendations from the crowdsourcing provider. The total amount paid to the crowdsourcing platform for the 3 classification tasks was approximately $35 (USD) per event. Payments to specific individual workers depend on how many tasks they performed and on their agreement with the test questions, following an internal procedure of the crowdsourcing provider.

Our first classification task is to identify tweets which are related to a crisis. A tweet may contain a crisis' keywords but be unrelated to it, as some keywords may be quite general, and refer to any number of topics other than the disaster situation. In addition, unscrupulous spammers sometimes exploit the popularity of a crisis hashtag to post promotional content [5]. As a result, the first labeling phase (M1) also has a data cleaning role. For each event we label a set of 1,000 tweets selected uniformly at random. We imposed a minimum threshold of 900 crisis-related tweets per crisis, and in the cases where it was necessary (9 out of 26 crises), we continued labeling tweets until passing the threshold. Next, we kept only the tweets that were related to the crisis (independently of whether they were deemed informative or not), and classified them with respect to *information types* (M2) and *sources* (M3).

Task Evaluation

Tweet classification is a subjective process, especially when performed at a large scale, and with a focus on tweet content. To evaluate to what extent subjectivity affects our results, we performed the following experiment: Two authors *independently* labeled 200 tweets sampled uniformly at random from all the crises. They classified tweets according to information types and sources, by looking at the content of the tweets as displayed in the Twitter platform, including conversations (if any), and looking at links in the tweets, and user profile information from its authors. We also note that authors had background information about each of the events.

We measure inter-assessor agreement with Cohen's Kappa, resulting in $\kappa = 0.80$ for information type (95% confidence interval CI: $[0.73, 0.87]$) and $\kappa = 0.73$ for source (95% CI: $[0.64, 0.81]$). Customarily, values in this range indicate substantial agreement.

Next, we take all tweets in which both authors agree and compare their joint label with those provided by crowdsource workers. The results are $\kappa = 0.81$ (95% CI: $[0.73, 0.88]$) for information type and $\kappa = 0.72$ for source (95% CI: $[0.62, 0.83]$). Again, these values reflect substantial agreement. The *individual* agreement of authors with workers (which includes cases in which the labels given by authors do not agree) is lower but still substantial ($\kappa = 0.69$ and $\kappa = 0.74$ for information type, $\kappa = 0.57$ and $\kappa = 0.63$ for source).

The conclusion is similar to that of previous work using crowdsourcing labeling (e.g. [14, 41]), crowdsource workers collectively provide reliable labels for social media annotation tasks, at a volume that would be very costly to achieve by other means (in our case, $26 \times 1,000 \times 3 = 78,000$ labels).

This experiment also allows us to evaluate the biases of crowdsourcing labeling. For information type, in 15% of the cases the crowdsourced label does not correspond to the one given by the authors (among the authors this discrepancy is 16%). The most common error of crowdsourcing workers is labeling "Caution and Advice" messages as either "Donations and Volunteering" or "Other Useful Information." For information source, in 17% of the cases the crowdsourced label did not agree with the one of the authors (among authors this discrepancy is 18%). The most common error was labeling "Eyewitness" as "Outsiders" or "Media." This means that in the analysis, we have to consider that "Caution and Advice" and "Eyewitness" may be underrepresented categories, while the other categories we mentioned may be overrepresented. The extent of the total underrepresentation/overrepresentation across all categories, however, is about 15%-17%, and more importantly, is not larger than the discrepancy among the two authors who performed this evaluation.

STEP 5: DATA ANALYSIS

The final step is to perform an analysis of the data annotated by the crowdsourcing workers. We begin by presenting results about the overall distribution of content types across crises, which we connect to the crisis dimensions by mining association rules. Then we consider temporal aspects, as well as the interplay between content dimensions.

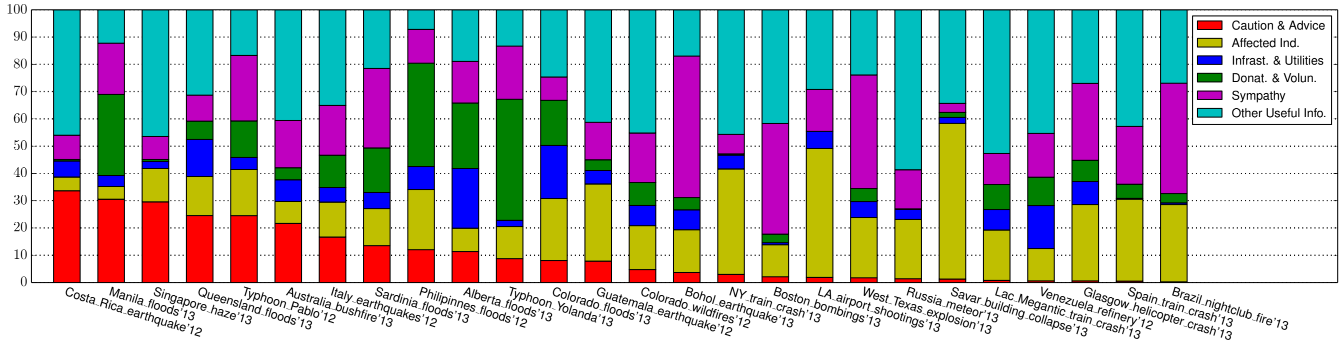
Finally, we show that while substantial variability exists, similar crises tend to have a similar distribution of message types. Though we make no claims that these 26 crises are representative of every event of every type we consider, we do note patterns and consistencies in the proportion of different messages, and present potential explanations about them, to serve as foundations for future explorations.

Content Types vs. Crisis Dimensions

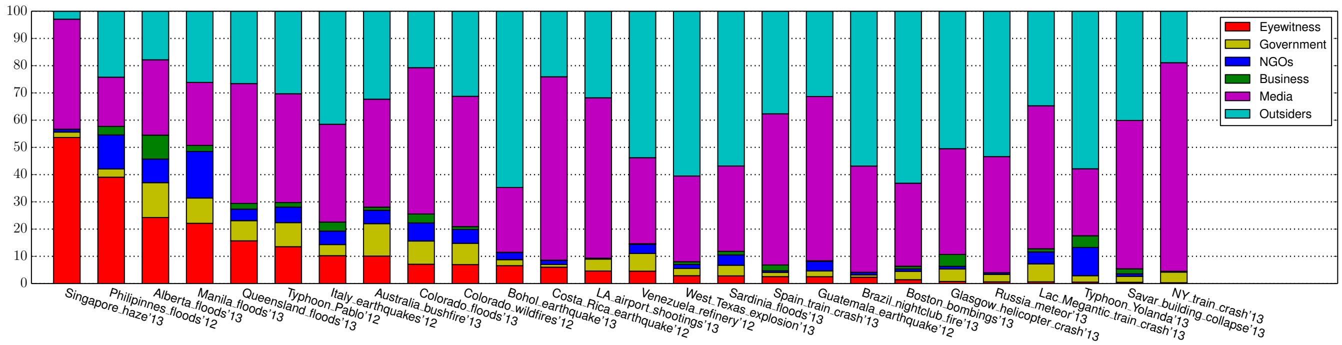
We first present our results regarding relationships between crisis dimensions and the prevalence of categories of information found in tweets.

Informativeness

The proportion of messages in each collection that were found to be about the crisis at hand (i.e. classified in the first two categories of M1) was on average 89% (min. 64%, max. 100%). In this case, one of the most significant factors is whether the keywords/hashtags adopted by people tweeting about the crisis are specific to the event, or were used for other purposes. For instance, `#yolandaph` was specifically used for Typhoon Yolanda, while `#dhaka` (the name of the capital of Bangladesh) was used after the Savar building collapse, but also for other purposes.



(a) Distribution of information types, sorted by descending proportion of *caution and advice* tweets.



(b) Distribution of information sources, sorted by descending proportion of *eyewitness* tweets.

Figure 1. Distributions of information types and sources (best seen in color)

Among these messages, the proportion of informative messages (i.e. those in the first category of M1) was on average 69% (min. 44%, max. 92%). Most of the messages considered "not informative" contained expressions of sympathy and emotional support (e.g. "thoughts and prayers").

Information types

Figure 1(a) shows the distribution of information types found in the tweets related to each crisis. Below, we sort the categories in decreasing order of average prevalence, noting the (wide) range on the proportion of each type.

- *Other useful information*: 32% on average (min. 7%, max. 59%). This "catchall" category is the largest among the information types. An analyst interested exclusively in the remaining categories can skip these messages on the initial pass of analysis. We note that the events in which this category was the least prevalent (i.e., the other categories accounted for more than 80% of the messages) were all diffused. While we do not claim that all, or even most, diffused events will have fewer-than-average amounts of "other useful information" tweets, it is potentially useful to know that this type of tweet is not prevalent in the diffused events we studied.

  The information captured by the "other useful information" category varies significantly across events. For instance, in the Boston bombings and LA Airport shootings in 2013, there are updates about the investigation and suspects; in the West Texas explosion and the Spain train crash, we find details about the accidents and the follow-up inquiry; in earthquakes, we find seismological details.

- *Sympathy and emotional support*: 20% on average (min. 3%, max. 52%). Tweets that express sympathy are present in all the events we examined. The 4 crises in which the messages in this category were more prevalent (above 40%) were all instantaneous disasters. Again, we make no hard-and-fast claims about all instantaneous disasters, but this finding leads us to conjecture that people are more likely to offer sympathy when events are not predicted, take people by surprise, and may cause additional distress due to their unforeseen occurrence.

- *Affected individuals*: 20% on average (min. 5%, max. 57%). The 5 crises with the largest proportion of this type of information (28%–57%) were human-induced, focalized, and instantaneous. These 5 events can also be viewed as particularly emotionally shocking. They resulted in casualties, but a small enough number of casualties to generate many reports regarding specific individuals who lost their lives or suffered injuries.

- *Donations and volunteering*: 10% on average (min. 0%, max. 44%). The number of tweets describing needs or offers of goods and services in each event varies greatly; some events have no mention of them, while for others, this is one of the largest information categories. In our data, tweets about donations and volunteering were more prevalent in Typhoon Yolanda in 2013 (44%) and in the floods in Sardinia, Colorado, Alberta, and Manila in 2013, and in the Philippines in 2012 (16%–38%). In contrast, they were 10% or less for all the human-induced crises we analyzed.

- *Caution and advice*: 10% on average (min. 0%, max. 34%). In instantaneous crises, there is unsurprisingly little information on this type (0%–8%), as these events are often not predicted and only post-impact advice can be present. The only exceptions in our data are the Italy earthquakes in 2012 (17%) — in which the collection covers two earthquakes plus a number of significant aftershocks which happen over an interval of less than 10 days, and Costa Rica earthquake in 2012 (34%) — when tsunami alerts were issued across Central America and parts of South America including even distant countries like Chile. Apart from these two events, the events with the most tweets that include information about caution and advice are caused by diffused natural hazards, and the 5 with the highest fraction from this set are all progressive (22%–31%). Further, baring the meteor that fell in Russia in 2013, we can see a clear separation between human-induced hazards and natural: all human induced events have less caution and advice tweets (0%–3%) than all the events due to natural hazards (4%–31%). The meteor was a rare event that felt like a bomb whose shock wave shattered windows and damaged thousands of buildings, remaining undetected before its atmospheric entry.[14]

- *Infrastructure and utilities*: 7% on average (min. 0%, max. 22%). The crises where this type of information was more than 10% were the Queensland, Alberta, and Colorado floods of 2013, and the Venezuela refinery explosion in 2012. In flood situations, it is common for electricity and water supplies to be cut, and in the case of the refinery explosion, many living in the area were suddenly without electricity due to the massive impact of the discharge.

Sources

In Figure 1(b), we see the distribution of tweet sources, and we observe the following:

- *Traditional and/or Internet media*: 42% on average (min. 18%, max. 77%). Regardless of the event, traditional and Internet media have a large presence on Twitter, in many cases more than 30% of the tweets. The 6 crises with the highest fraction of tweets coming from a media source (54%–76%) are instantaneous, which make "breaking news" in the media.

- *Outsiders*: 38% on average (min. 3%, max. 65%). Depending on the event, the number of "outsiders" can vary. This was in general about 18% or more, with the exception of the Singapore haze in 2013 that had only 3% of tweets from outsiders. The Singapore haze was an event that strongly disrupted the city, but did not result in life-threatening injuries or deaths.

- *Eyewitness accounts*: 9% on average (min. 0%, max. 54%). In general, we find a larger proportion of eyewitness accounts during diffused disasters caused by natural hazards. The 12 events with the highest percentage of eyewitness accounts are all diffused (6%–54%) and the top 6 are also progressive (13%–54%).

- *Government*: 5% on average (min. 1%, max. 13%). A relatively small fraction of tweets include information sourced by government officials and agencies—only for two of the crises we analyze this exceeds 10%. We surmise that this is because governments must verify information before they broadcast it, which takes considerable time [21]. Therefore, government accounts may not have the most up-to-date information in crisis situations. The 7 events with the highest percentage of tweets from governmental agencies are due to natural-hazards, progressive and diffused (7%-13%), which are the cases when the governments intervene to issue or lift warnings or alerts.

- *NGOs*: 4% on average (min. 0%, max. 17%). Like governments, NGOs are also careful to broadcast only verified information. In the human-induced crises we studied there is little NGO activity in Twitter (≈ 4% or less). The highest levels of NGO tweets are seen in natural disasters and all those in which the fraction of such tweets was 6% or more are typhoons and floods.

- *Business*: 2% on average (min. 0%, max. 9%). For the most part, we do not see a large amount of tweet activity from businesses in the disaster situations we studied. The proportion is below 5% for all crises except the Alberta floods in 2013 with 9% of tweets. Furthermore, with only one exception—the Glasgow helicopter crash—the crises with 3% or more tweets from business were diffused.

Association Rules

To systematically search for relationships between the characteristics of crises and the messages in Twitter, we applied an association-rules mining method [11]. To err in the side of caution, we report only the automatically-discovered association rules that are valid for more than 20 out of the 26 crises. To apply this method to numerical data, each category in the information types and sources was divided into two classes: above the median, and below the median.

For information types, we found one rule that is valid for 24 out of 26 of the crises: when the geographical spread is diffused, the proportion of caution and advice tweets is above the median, and when it is focalized, the proportion of caution and advice tweets is below the median. For sources, we found one rule that is valid for 21 out of 26 of the crises: human-induced accidental events tend to have a number of eyewitness tweets below the median, in comparison with intentional and natural hazards.

Both rules are possibly related to different levels of access to the area affected by the event and to its surroundings.

Content Redundancy

We next look at content redundancy. Heuristically, we consider two tweets to be near-duplicates if their longest common subsequence was 75% or more of the length of the shortest tweet. Among the sources of information, messages originating from non-governmental organizations and government

[14] http://en.wikipedia.org/wiki/Chelyabinsk_meteor

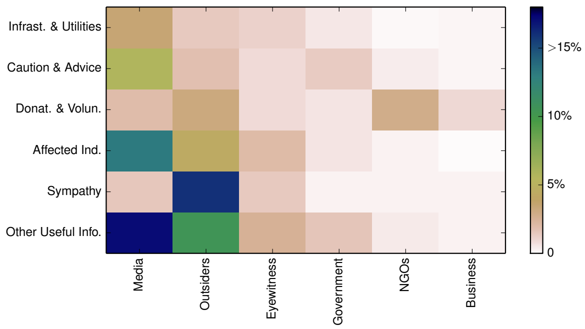


Figure 2. Average distribution of tweets across crises into combinations of information types (rows) and sources (columns). Rows and columns are sorted by total frequency, starting on the bottom-left corner. The cells in this figure add up to 100%.

sources tended to show more redundancy, with the top 3 messages (and their near-duplicates) accounting for ≈20%-22% of the tweets. Among information types, messages of caution and advice, and those containing information about infrastructure and utilities, were the most repeated ones, with the top 3 messages (and their near-duplicates) comprising ≈12-14% of the tweets.

Types and Sources

Some information types are more frequently associated with particular sources, as shown in Figure 2, in which each ⟨type, source⟩ cell depicts the probability that a tweet has that specific combination of information type and source. NGOs and business are more frequently the source of tweets related to donations and volunteering, mostly to ask for resources and request volunteer work (NGOs), or to announce free or discounted goods or services for those affected by a disaster (business).

Tweets from governments are often messages of caution and advice, such as tornado alerts; this agrees with observations in [51] where "preparedness" is the larger category used by government communications. Instead, eyewitness tweets focus on affected individuals. Both government and eyewitness tweets also frequently include a variety of messages that belong to the "other useful information" category. Outsider messages are predominantly about sympathy and support.

Finally, tweets from traditional and Internet media offer a variety of information types including information about affected individuals, and messages of caution and advice. Media are also the most prominent source of information regarding infrastructure and utilities.

Temporal Aspects

We study how the volume of different categories of messages evolves over time, as shown in Tables 5 and 6 (at the end of the paper). We separated crises according to their temporal development (instantaneous vs. progressive), depicting using "spark lines" the total volume of messages over time, and the total volume of messages in each information type and source.[15] This analysis focuses on the differences between the average timestamps of messages in different information categories.[16]

In terms of information types, the messages that arrive first are those of caution and advice, and sympathy and support, roughly in the first 12–24 hours after the peak of the crisis. This is particularly evident in instantaneous crises. Then, messages about affected individuals and infrastructure are most frequent. The last messages to appear are those related to donations and volunteering. Interestingly, this follows the progression in the stages of a crisis from emergency response to early recovery actions [47].

In terms of sources, there are differences depending on the type of temporal development. In *instantaneous* crises, outsiders, media and NGO messages appear early, with other sources following (the temporal position of eyewitness messages varies substantially depending on crisis type). On the other hand, during *progressive* crises, eyewitness and government messages appear early, mostly to warn and advice those in the affected areas, while NGO messages appear relatively late. In addition, there is an interesting temporal complementarity between messages from governments and NGOs that merits to be studied in depth in future work.

Crisis Similarity

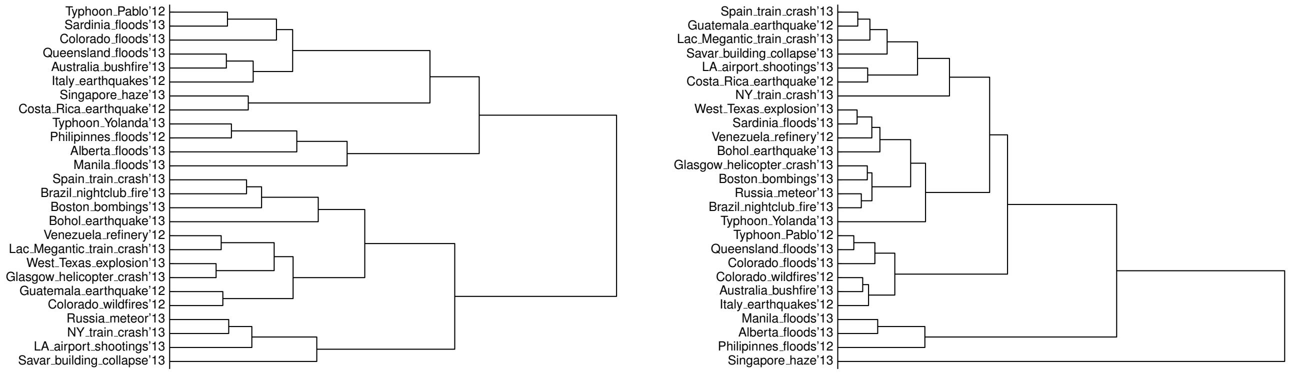
In further seeking links between disaster characteristics and tweet content and source, we apply an unsupervised method; specifically, hierarchical agglomerative clustering. Performing this clustering uncovered groups of crises that have similar content distribution. Given that we compare probability distributions, to measure the similarity between two crisis events we use Bhattacharyya distance (for two discrete distributions p and q this is $-\ln(\sum_{c \in C} \sqrt{p(c)q(c)})$ where C is the set of all classes) which quantifies the overlap between two statistical samples. To combine clusters of crises, we used complete-linkage clustering, which merges those clusters for which the distance between their furthest elements is the smallest.

Figure 3(a) shows the resulting dendrogram when the clustering is done according to the distribution of information types. We see two large clusters: first, the cluster on the bottom is dominated by human-induced crises, while in the one on the top there are only natural hazards. This indicates that, despite the significant variations we have shown, human-induced crises are more similar to each other in terms of the types of information disseminated through Twitter than to natural hazards.

Second, events also cluster depending on how they developed. The cluster at the bottom includes instantaneous events, with one exception: the Colorado wildfires in 2012. This exception may be due to the nature of this particular fire. The combination of heat, drought conditions, and high winds

[15] Each point in the spark line corresponds to a *calendar* day, which explains why in some instantaneous crises the overall curve goes up at the beginning (when the crisis occurs at night).

[16] Peak, average, and median timestamps for each time series in Tables 5 and 6 are available in our data release.



(a) Clusters by information type.

(b) Clusters by source.

Figure 3. Dendrograms obtained by hierarchical agglomerative clustering of crises. The length of the branch points reflect the similarity among crises. We remark that the clusters do not reflect similar messages, but instead similarities in terms of the proportion of different information types and sources in each crisis.

caused the fire to quickly develop, and it claimed 350 houses in just over 12 hours. The cluster on the top includes progressive disasters, with two outliers: Italy earthquakes in 2012—a sequence of earthquakes and aftershocks—and the Costa Rica earthquake in 2012—during which a Caribbean-wide tsunami watch was issued, resulting in a large volume of caution and advice messages that are typically more prominent in progressive crises.

A similar picture emerges in the case of clusters by distribution of sources, shown in Figure 3(b). In this case, there is a large cluster dominated by human-induced crises (on the top), followed by two small clusters encompassing only natural hazards, and the Singapore haze 2013 as an outlier (this haze was caused by a mix of natural and human causes). Further, the large cluster on the top is dominated by instantaneous events (with two exceptions, Typhoon Yolanda and Sardinia Floods in 2013), while in the other clusters the events are progressive, excepting Italy earthquakes in 2012.

Furthermore, while the events development and type arise as the main factors impacting the clusters composition, in both Figures 3(a) and 3(b) we also notice that the clusters are being dominated by either diffused (top cluster by information type and bottom clusters by information source) or focalized events (the remaining clusters). The clusters tendency to encompass events that are similar along all these dimensions is likely explained by the dependency among the crisis dimensions (e.g., typically, the progressive events are also diffused and human-induced crises tend to be focalized).

DISCUSSION

Disasters are common events that occur regularly; the United Nations Office for Coordination of Humanitarian Affairs recorded 394 disasters caused by natural hazards in the 2002–2011 period [1]. While disasters take place often, and may be caused by similar hazards and/or human actions, each event is unique [36] (pag. 5). Regardless of their distinct nature, and of variations in individual reactions and responses, commonalities across crises exist. Sociologists of disaster point out that despite the differences among disaster agents (e.g. flood, earthquake, bomb, fire), there are actions that planning and emergency response teams must take that are independent of these differences [47].

This brings us to an interesting juxtaposition; the types and amounts of information broadcast on Twitter differ across each of the 26 specific crises we studied. This can be viewed as a display of the uniqueness of each event. In some cases the most common tweet in one crisis (e.g. eyewitness accounts in the Singapore haze crisis in 2013) was absent in another (e.g. eyewitness accounts in the Savar building collapse in 2013). Furthermore, even two events of the same type in the same country (e.g. Typhoon Yolanda in 2013 and Typhoon Pablo in 2012, both in the Philippines), may look quite different vis-à-vis the information on which people tend to focus.

Yet, when we look at the Twitter data at a meta-level, our analysis reveals commonalities among the types of information people tend to be concerned with, given the particular dimensions of the situations such as hazard category (e.g. natural, human-induced, geophysical, accidental), hazard type (e.g. earthquake, explosion), whether it is instantaneous or progressive, and whether it is focalized or diffused. For instance, caution and advice tweets from government sources are more common in progressive disasters than in instantaneous ones. The similarities do not end there. When grouping crises automatically based on similarities in the distributions of different classes of tweets, we also realize that despite the variability, human-induced crises tend to be more similar to each other than to natural hazards.

This leads us to believe that we can view Twitter as a medium through which the nuance of disaster events is highlighted or amplified; it is a tool that becomes incorporated into the social construction of the disaster event, and through which we can understand the detailed differences on a large scale when we look closely at Twitter data. At the same time, when we look at those same data at a higher level, we see commonalities and patterns.

Practitioners, including emergency managers, public information officers, and those who develop the tools used by them, should consider that the proportion of tweets that are relevant for a specific purpose will almost invariably be smaller than the proportion of the tweets that are not. For instance, if an analyst or an application focuses on content that is not present in mainstream or other Internet media sources, and wants to exclude content provided by outsiders who are not affected by the crisis, then it will have to skip through 80% of the tweets on average. The same holds for information types. If we group together the four main types we used (affected individuals, donations and volunteering, caution and advice, and infrastructure and utilities), they cover on average 47% of the tweets related to a crisis. This implies that if an application wants to focus on these information types, at least 53% of the messages will have to be discarded. These are lower bounds, as often not all of the tweets of a given type will be relevant for a particular application. *Noise* is a natural consequence of the diversity of information in this medium.

Developers should consider that emergency response includes a set of actions that have to be taken in preparation of any crisis event, plus a broad space for adaptability in response to specific events [47]. Hence, tools to process social media in disaster should consider that there are broad classes of information that are likely to be prevalent, and can be anticipated to occur. At the same time, a substantial volume of messages will depend on specificities of every event, and tools must incorporate methods to adaptively detect and process them.

CONCLUSIONS

Our systematic examination of a diverse set of crisis situations uncovered substantial variability across crises, as well as patterns and consistencies. To the best of our knowledge, this is the largest transversal study on tweets broadcast in response to various international disaster and crisis situations.

Future Work

The high-level patterns we have found lay the foundations for future studies that go into the detail of each specific crisis or each specific information category analyzed.

However, we note that we did not cover all possible crisis situations. For instance, we did not include human-induced progressive or diffused situations, which are less common than the classes we did study. The former (human-induced progressive) mostly refers to politically-driven crises, such as instability leading to demonstrations, riots, and/or civil wars. The latter (human-induced diffused) in recent years have been mostly wars affecting an entire country or region, or less-common, large-scale industrial accidents such as the oil spill in the Gulf of Mexico in 2010. Additionally, the management of a crisis is typically divided into phases: mitigation, preparedness, response and recovery [35, 47]. The work we present here is concerned mostly with the response phase and partially with the recovery phase, as these attract the bulk of social media activities [23]. Language and cultural differences could also be included as explicit crisis dimensions [20, 37], together with temporal factors. Microblogging practices are likely to evolve over the years, and our collections cover a period of just about 20 months. The study of other crisis dimensions, other types of crises and other phases, will certainly deepen our findings.

Methodologically, we asked crowdsource workers to match each tweet to one specific class. This simplifies the labeling process and makes the presentation of the results clearer. When workers associate a tweet to multiple classes, it may be possible that the distributions change. Employing professional emergency managers as annotators instead of crowdsource workers may lead to further results. Finally, assessing the quality, credibility, or veracity of the information in each tweet is relevant for most of the potential consumers of this data. However, we note that in these cases the cost of the annotation would certainly increase—or the amount of labeled data would decrease.

Data Release

The tweets used in this research, and the labels collected through the crowdsourced annotation, are available for research purposes at http://crisislex.org/

Acknowledgments

We thank Patrick Meier for his valuable advice while defining this project. We are indebted to Karl Aberer for his support. We thank the reviewers for their detailed feedback. Alexandra Olteanu was partially supported by the grant *Reconcile: Robust Online Credibility Evaluation of Web Content* from Switzerland through the Swiss Contribution to the enlarged European Union.

REFERENCES

1. World humanitarian data and trends. Tech. rep., UN OCHA, Dec. 2013.
2. Acar, A., and Muraki, Y. Twitter for crisis communication: lessons learned from Japan's tsunami disaster. *Int. J. of Web Based Communities* (2011).
3. Adams, D. S. Policies, programs, and problems of the local Red Cross disaster relief in the 1960s. Tech. rep., University of Delaware, Disaster Research Center, 1970.
4. Aramaki, E., Maskawa, S., and Morita, M. Twitter catches the flu: Detecting influenza epidemics using Twitter. In *Proc. of EMNLP* (2011).
5. Benevenuto, F., Magno, G., Rodrigues, T., and Almeida, V. Detecting spammers on Twitter. In *In Proc. of CEAS* (2010).
6. Bruns, A., Burgess, J. E., Crawford, K., and Shaw, F. #qldfloods and @qpsmedia: Crisis communication on Twitter in the 2011 South East Queensland floods. Tech. rep., ARC Centre, Queensland University of Technology, 2012.
7. Caragea, C., McNeese, N., Jaiswal, A., Traylor, G., Kim, H., Mitra, P., Wu, D., Tapia, A., Giles, L., Jansen, B. J., et al. Classifying text messages for the Haiti earthquake. In *Proc. of ISCRAM* (2011).
8. Carr, L. J. Disaster and the sequence-pattern concept of social change. *American Journal of Sociology* (1932).

9. Carvin, A. *Distant Witness*. CUNY Journalism Press, 2013.

10. Castillo, C., Mendoza, M., and Poblete, B. Information credibility on Twitter. In *Proc. of WWW* (2011).

11. Cohen, W. W. Fast effective rule induction. In *Proc. of ICML* (1995), 115–123.

12. De Choudhury, M., Diakopoulos, N., and Naaman, M. Unfolding the event landscape on Twitter: Classification and exploration of user categories. In *Proc. of CSCW* (2012).

13. Denef, S., Bayerl, P. S., and Kaptein, N. A. Social media and the police: tweeting practices of British police forces during the August 2011 riots. In *Proc. of CHI* (2013).

14. Diakopoulos, N., De Choudhury, M., and Naaman, M. Finding and assessing social media information sources in the context of journalism. In *Proc. of CHI* (2012).

15. Esser, F., and Hanitzsch, T., Eds. *The Handbook of Comparative Communication Research*. ICA Handbooks. Routledge, Apr. 2012.

16. Fischer, H. W. *Response to disaster: fact versus fiction & its perpetuation—the sociology of disaster*. University Press of America, 1998.

17. Fraustino, J. D., Liu, B., and Jin, Y. Social media use during disasters: A review of the knowledge base and gaps. Tech. rep., Science and Technology Directorate, U.S. Department of Homeland Security, 2012.

18. Gerlitz, C., and Rieder, B. Mining one percent of Twitter: collections, baselines, sampling. *M/C Journal* (2013).

19. González-Bailón, S., Wang, N., Rivero, A., Borge-Holthoefer, J., and Moreno, Y. Assessing the bias in communication networks sampled from Twitter. *Social Networks 38* (July 2014), 16–27.

20. Hong, L., Convertino, G., and Chi, E. H. Language matters in Twitter: A large scale study. In *Proc. of ICWSM* (2011).

21. Hughes, A. L. Participatory design for the social media needs of emergency public information officers. In *Proc. of ISCRAM* (2014).

22. Hughes, A. L., St. Denis, L. A., Palen, L., and Anderson, K. Online public communications by police and fire services during the 2012 hurricane Sandy. In *Proc. of CHI* (2014).

23. Imran, M., Castillo, C., Lucas, J., Patrick, M., and Rogstadius, J. Coordinating human and machine intelligence to classify microblog communications in crises. In *Proc. of ISCRAM* (2014).

24. Imran, M., Elbassuoni, S. M., Castillo, C., Diaz, F., and Meier, P. Extracting information nuggets from disaster-related messages in social media. In *Proc. of ISCRAM* (2013).

25. Joseph, K., Landwehr, P. M., and Carley, K. M. Two 1%s don't make a whole: Comparing simultaneous samples from Twitter's streaming API. In *Social Computing, Behavioral-Cultural Modeling and Prediction*. 2014.

26. Kanhabua, N., and Nejdl, W. Understanding the diversity of tweets in the time of outbreaks. In *Proc. of WOW workshop* (2013).

27. Kumar, S., Morstatter, F., Zafarani, R., and Liu, H. Whom should I follow?: Identifying relevant users during crises. In *Proc. of Hypertext* (2013).

28. Leavitt, A., and Clark, J. A. Upvoting hurricane Sandy: event-based news production processes on a social news site. In *Proc. of CHI* (2014).

29. Metaxas, P., and Mustafaraj, E. The rise and the fall of a citizen reporter. In *Proc. of WebSci* (2013).

30. Monroy-Hernández, A., boyd, d., Kiciman, E., De Choudhury, M., and Counts, S. The new war correspondents: The rise of civic media curation in urban warfare. In *Proc. of CSCW* (2013).

31. Morstatter, F., Pfeffer, J., and Liu, H. When is it biased?: assessing the representativeness of Twitter's streaming API. In *Proc. of Web Science Track at WWW* (2014).

32. Morstatter, F., Pfeffer, J., Liu, H., and Carley, K. M. Is the sample good enough? comparing data from Twitter's streaming API with Twitter's Firehose. In *Proc. of ICWSM* (2013).

33. Munro, R., and Manning, C. D. Short message communications: users, topics, and in-language processing. In *Proc. of DEV* (2012).

34. Olteanu, A., Castillo, C., Diaz, F., and Vieweg, S. CrisisLex: A lexicon for collecting and filtering microblogged communications in crises. In *Proc. of ICWSM* (2014).

35. Palen, L., and Liu, S. B. Citizen communications in crisis: anticipating a future of ICT-supported public participation. In *Proc. of CHI* (2007).

36. Perry, R. W., and Quarantelli, E. L. *What is a disaster?: New answers to old questions*. Xlibris Corporation, 2005.

37. Poblete, B., Garcia, R., Mendoza, M., and Jaimes, A. Do all birds tweet the same?: characterizing Twitter around the world. In *Proc. of CIKM* (2011).

38. Prelog, A. J. *Social Change and Disaster Annotated Bibliography*. PhD thesis, Department of Economics, Colorado State University, 2010.

39. Qu, Y., Huang, C., Zhang, P., and Zhang, J. Microblogging after a major disaster in China: a case study of the 2010 Yushu earthquake. In *Proc. of CSCW* (2011).

40. Shaw, F., Burgess, J., Crawford, K., and Bruns, A. Sharing news, making sense, saying thanks: Patterns of talk on Twitter during the Queensland floods. *Australian Journal of Communication* (2013).

41. Snow, R., O'Connor, B., Jurafsky, D., and Ng, A. Y. Cheap and fast—but is it good?: Evaluating non-expert annotations for natural language tasks. In *Proc. of EMNLP* (2008).

42. Sreenivasan, N. D., Lee, C. S., and Goh, D. H.-L. Tweet me home: exploring information use on Twitter in crisis situations. In *Online Communities and Social Computing*. 2011.

43. Starbird, K., Muzny, G., and Palen, L. Learning from the crowd: Collaborative filtering techniques for identifying on-the-ground twitterers during mass disruptions. In *Proc. of ISCRAM* (2012).

44. Starbird, K., and Palen, L. "Voluntweeters": Self-organizing by digital volunteers in times of crisis. In *Proc. of CHI* (2011).

45. Starbird, K., Palen, L., Hughes, A. L., and Vieweg, S. Chatter on the red: what hazards threat reveals about the social life of microblogged information. In *Proc. of CSCW* (2010).

46. Thomson, R., Ito, N., Suda, H., Lin, F., Liu, Y., Hayasaka, R., Isochi, R., and Wang, Z. Trusting tweets: The Fukushima disaster and information source credibility on Twitter. In *Proc. of ISCRAM* (2012).

47. Tierney, K. J. Disaster preparedness and response: Research findings and guidance from the social science literature. Tech. rep., University of Delaware, Disaster Research Center, 1993.

48. Truelove, M., Vasardani, M., and Winter, S. Towards credibility of micro-blogs: characterising witness accounts. *GeoJournal* (2014), 1–21.

49. Vieweg, S. *Situational Awareness in Mass Emergency: A Behavioral and Linguistic Analysis of Microblogged Communications*. PhD thesis, University of Colorado at Boulder, 2012.

50. Vieweg, S., Hughes, A. L., Starbird, K., and Palen, L. Microblogging during two natural hazards events: What Twitter may contribute to situational awareness. In *Proc. of CHI* (2010).

51. Wukich, C., and Mergel, I. A. Closing the Citizen-Government communication gap: Content, audience, and network analysis of government tweets. *Social Science Research Network Working Paper Series* (Aug. 2014).

APPENDIX

- Table 4 containing the list of keywords used to collect data from each crisis.
- Tables 5 and 6 depicting temporal distributions of tweets on each crisis for each type, and for each source.

Table 4. List of keywords used to collect data for each of the crises in this study.

Year	Crisis Name	Keywords	#Kw
2012	Italy earthquakes	earthquake italy; quake italy; #modena; #sanfelice; san felice; modena terremoto; modena earthquake; modena quake; #norditalia; terremoto italia; #terremoto;	11
2012	Colorado wildfires	#cofire; #boulderfire; #colorado; #wildfire; #waldocanyonfire; #waldofire; #waldocanyon; colorado springs; #highparkfire; #flagstafffire; #littlesandfire; #treasurefire; #statelinefire; #springerfire; #lastchancefire; #fourmilefire; #4milefire; #fourmilecanyonfire; #boulderfire; #bisonfire; colorado wildfire; colorado wildfires; colorado fire; colorado fires; boulder fire; boulder fires; boulder wildfires; boulder wildfires;	28
2012	Philipinnes Floods	#rescueph; #reliefph; #floodsph; #prayforthephilippine; manila flood; manila floods; philippine floods; philippine flood; #floodph; #phalert; #habagat;	11
2012	Venezuela refinery explosion	paraguana refinery; venezuela refinery; paraguaná refinery; #paraguana; #paraguaná; amuay refinery; venezuelan refinery; #amuay; paraguaná refinería; paraguana refineria; amuay refineria; amuay refinería; #falcon; #falcón; refinería venezuela; refineria venezuela; refinería paraguana;	17
2012	Costa Rica earthquake	#temblorcr; #terremotocr; #costarica; #terremoto; costa rica quake; costa rica earthquake; costa rica temblor; costa rica terremoto; #creq; costa rican quake; costa rican earthquake; #quake; #earthquake;	13
2012	Guatemala earthquake	#sismo; #guatemala; tiemblaenguate; temblorgt; terremotogt; temblor guatemala; terremoto guatemala; sismo guatemala; earthquake guatemala; quake guatemala; #sanmarcos; #terremotoguatemala; #tremorguatemala;	13
2012	Typhoon Pablo	#pabloph; #reliefph; #bopha; #typhoonpablo; #typhoonbopha; typhoon bopha; typhoon pablo; #bopha; #pablo; #typhoon; #walangpasok; #mindanao; #visayas; #hinatuan; #rescueph; #pablosafetytips; #cdo; #strongerph;	18
2013	Brazil nightclub fire	#forçasantamaria; boate kiss; #boatekiss; #santamaria; #tragédiaemsm; #tragediaemsm; #todosdesejamforçasasantamaria; #brazilfire; #brazil fire; brazil nightclub; #brasildesejaforçasasvitimasdesantamaria; #prayforsantamaria; #prayforbrazil;	13
2013	Queensland Floods	#qldflood; #bigwet; queensland flood; australia flood; #qldfloods; queensland floods; australia floods; queensland flooding; qld flood; qld floods; qld flooding; australia flooding;	12
2013	Russian Meteor	#метеорит; #meteor; #meteorite; russia meteor; russian meteor; #russianmeteor; #chelyabinsk; #челябинск;	8
2013	Boston Bombings	boston explosion; boston explosions; boston blast; boston blasts; boston tragedies; boston tragedy; prayforboston; boston attack; boston attacks; boston terrorist; boston terrorists; boston tragic; bostonmarathon; boston marathon; boston explosive; boston bomb; boston bombing; dzhokhar; tsarnaev; marathon attack; marathon explosion; marathon explosions; marathon tragedies; marathon tragedy; marathon blasts; marathon blast; marathon attacks; marathon bomb; marathon bombing; marathon explosive;	30
2013	Savar building collapse	#savar; #bangladesh; bangladesh collapse; #ranaplaza; savar bangladesh; savar collapse; rana plaza;	7
2013	West Texas Explosion	#westexplosion; west explosion; waco explosion; texas explosion; texas fertilizer; prayfortexas; prayforwest; waco tx; west tx; west texas; waco texas; #west; #waco; westexplosion; west explosion; waco explosion; tx explosion; fertilizer explosion; prayfortexas; prayforwest; westtx; wacotx; west texas; waco texas; west tx; waco tx; texas fertilizer; west fertilizer; waco fertilizer;	29
2013	Alberta Floods	alberta flood; #abflood; canada flood; alberta flooding; alberta floods; canada flooding; canada floods; #yycflood; #yycfloods; #yycflooding; calgary flood; calgary flooding; calgary floods;	13
2013	Singapore Haze	#sghaze; singapore haze; #hazyday; blamethehaze; mustbehaze; #sg #haze; singapore #hazy;	7
2013	Lac-Megantic train crash	#lacmegantic; #lacmégantic; #lacmég; #lacmeg; #tragedielacmegantic; #tragédielacmégantic; #mégantic; lac mégantic; lac megantic; quebec train explosion; quebec train derailment; quebec train crash; quebec oil train; canada train oil; canada train oil; canadian train oil;	16
2013	Spain train crash	compostela train; spain train; tren compostela; españa tren; #santiagocompostela; #accidentesantiago;	6
2013	Manila Floods	baha manila; #maringph; #rescueph; #reliefph; #floodsph; #prayforthephilippine; manila flood; manila floods; philippine floods; philippine flood; #floodph; #phalert; #safenow; #trafficph; #habagat; #maring; #maringupdates;	17
2013	Colorado Floods	#cofloodrelief; colorado floods; colorado flooding; #coloradoflood; #coflood; #opcoflood; #boulderflood; #longmont;	8
2013	Australia wildfires	#nswfires; #nswbushfire; #nswbushfires; #nswrfs; #sydneybushfire; #sydneyfire; #sydneyfires; #sydneybushfires; nsw #bushfire; #redoctober; australia #bushfire; #faulconbridge; #nswrfs; #bushfire sydney; nsw fire; #prayforaustralia; #prayfornsw; australia fire; sydney fire; nsw fires; australia fires; sydney fires; prayfornsw;	23
2013	Bohol earthquake	#phquake; #pheq; #phtrenchquake; philippines earthquake; philippines quake; ph earthquake; ph quake; #phtrenchquake; #prayforthephilippines; #rescueph; #reliefph; #tabangbohol; #tabangcebu; #bohol; #cebu; prayforvisayas; prayforbohol; #lindol;	18
2013	Glasgow helicopter crash	#prayerforglasgow; #helicopter; glasgow helicopter; #clutha; helicopter crash;	5
2013	LA Airport Shootings	lax shooting; lax shootings; lax shooter; lax suspect; #laxshooting; lax airport; #lax; airport shooting; airport shootings; #losangeles airport; lax victims;	11
2013	NYC train crash	#newyork derailment; ny derailment; nyc derailment; #metronorth derailment; #spuyten duyvil; #nyctrain; new york derailment; metro north derailment; #metronorth derailment; ny train crash; nyc train crash; newyork train crash; york train crash; #metronorth train crash; metro north crash; ny train derailed; york train derailed; nyc train derailed;	18
2013	Sardinia Floods	sardinia floods; sardinia flooding; cyclone cleopatra; #cyclonecleopatra; #sardinia; sardegna alluvione; #cleopatra alluvione; #sardegna;	8
2013	Typhoon Yolanda	#typhoonyolanda; #yolandaph; #yolanda; #haiyan; #tracingph; #floodph; #safenow; #rescueph; #reliefph; typhoon yolanda; typhoon haiyan; typhoon philippines; #typhoonhaiyan; #typhoonaid; #philippines; #typhoon; #supertyphoon; #redcrossphilippines; #yolandaactionweekend; rescue ph; typhoon ph; super typhoon;	22

Table 5. Temporal distribution of tweets across information sources (top) and types (bottom) for the progressive events we analyzed. The 3 most frequent sources, respectively, information types, per crisis are highlighted in green. The red vertical bar indicates the peak volume for all tweets related to each event. (Best seen in color.)

Name	Alberta floods'13	Australia wildfires'13	Colorado floods'13	Colorado wildfires'12	Manila floods'13	Philipinnes floods'12	Queensland floods'13	Sardinia floods'13	Singapore haze'13	Typhoon Pablo'12	Typhoon Yolanda'13
Overall											
Eyewitness											
Government											
NGOs											
Business											
Media											
Outsiders											
Affected Ind.											
Infrast. & Utilities											
Donat. & Volun.											
Caution & Advice											
Sympathy											
Other Useful Info.											

Table 6. Temporal distribution of tweets across information sources (top) and types (bottom) for the instantaneous events we analyzed. The 3 most frequent sources, respectively, information types, per crisis are highlighted in green. The red vertical bar indicates the peak volume for all tweets related to each event. (Best seen in color.)

Name	Bohol earthq.'13	Boston bombings'13	Brazil nightclub fire'13	Costa Rica earthq.'12	Glasgow helicopter crash'13	Guatemala earthq.'12	Italy earthq.'12	LA airport shootings'13	Lac-Megantic train crash'13	NYC train crash'13	Russian meteor'13	Savar building collapse'13	Spain train crash'13	Venezuela refinery explosion'12	West Texas explosion'13
Overall															
Eyewitness															
Government															
NGOs															
Business															
Media															
Outsiders															
Affected Ind.															
Infrast. & Utilities															
Donat. & Volun.															
Caution & Advice															
Sympathy															
Other Useful Info.															