



**HAL**  
open science

## Social media processing in crisis response: an attempt to shift from data to information exploitation.

Julien Coche, Guillermo Romera Rodriguez, Aurelie Montarnal, Andrea Tapia, Frederick Benaben

### ► To cite this version:

Julien Coche, Guillermo Romera Rodriguez, Aurelie Montarnal, Andrea Tapia, Frederick Benaben. Social media processing in crisis response: an attempt to shift from data to information exploitation.. HICSS 2021 - 54th Hawaii International Conference on System Sciences, Jan 2021, Online, United States. pp.2285-2294. hal-03095502

**HAL Id: hal-03095502**

**<https://imt-mines-albi.hal.science/hal-03095502v1>**

Submitted on 4 Jan 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

## Social media processing in crisis response : an attempt to shift from data to information exploitation.

Julien Coche  
IMT Mines Albi, France  
[julien.coche@mines-albi.fr](mailto:julien.coche@mines-albi.fr)

Guillermo Romera Rodriguez  
Pennsylvania State University, USA  
[gkr5144@psu.edu](mailto:gkr5144@psu.edu)

Aurelie Montarnal  
IMT Mines Albi, France  
[aurelie.montarnal@mines-albi.fr](mailto:aurelie.montarnal@mines-albi.fr)

Andrea Tapia  
Pennsylvania State University, USA  
[atapia@psu.edu](mailto:atapia@psu.edu)

Frederick Benaben  
IMT Mines Albi, France  
[frederick.benaben@mines-albi.fr](mailto:frederick.benaben@mines-albi.fr)

### Abstract

*Information about the ongoing events is of the utmost importance during emergencies. Previous work in crisis informatics found new ways to pull data from unexploited sources, such as social media. But while the volume of information is crucial, the way the information is reported and provided becomes increasingly important as the volume grows. Yet, little has been achieved on information management. This article proposes a way to automatically organize information from social media data up to decision-makers. This organization is enabled by a metamodel [1] designed to model crucial information in crises. The article is organized as follows. First, the organization of current social media processing systems is presented. Then, the article presents the metamodel used and how it is relevant to organized information in crisis events through the lens of the 6W's [2]. Finally, it walks through the implementation of the proposal based on the two previous parts.*

### 1. Introduction

In our modern societies, crises are common occurrences. They affect every organization (societal, economic, environmental, political...) and impact systems that people rely on.

Despite their changing nature and their various aspects, one common denominator is the confusion they create. People suffering from the crisis are obviously the first victims of the event, but individuals responsible to respond come second. The confusion that crisis generates is mainly due to the need for fast, actionable, and useful information to answer to an unknown scenario. Every bit of information then becomes paramount. While organizations can control

and receive a continuous flow of information during normal operation, they might find themselves in the unknown once a crisis breaks out. The origin may lie in the lack of veracity, speed, or amount of information. Hence why the data used in these situations can be useless.

As a solution to solve the issue of information flow mentioned previously, crisis informatics research groups identified social media as a potent source of data, able to provide relevant insights for crisis management. Social media platforms have an abundance of data made up of billions of users spread around the globe. These platforms act like a network comprised of human sensors capable of providing information on a multitude of events. This network effect has been previously observed [3, 4], specifically how citizens use social media to share what is happening and how events unfold around them and their communities. Moreover, research has shown that individuals tend to share more information during crises [5]. This is why being able to leverage the data created within social media in a digestible and useful way could further help the different actors involved in these crises.

Crisis response time is a factor in its outcome and duration. But we cannot be deceived by thinking crisis are only a problem of information flow or response time. We also need to understand that crisis response is a multi-dimensional environment where different stakeholders are taking part in [6]. Besides, these different actors involved may never have had to collaborate before. Previous research in crisis management studied the flow of information among different actors during exercises [7]. The main issues identified concerned: asymmetric knowledge and experience between the different actors, difficulties in maintaining mutual awareness, uneven workload distribution, and disrupted communications.

The combination of all these factors - information, response time, and multiple stakeholders - means that

there is not a methodology that fits all when it comes to crisis response. Thus, one of the components of crisis response is the ability to re-establish the flow of information needed to respond [8].

However, considering crisis response as an issue of information flow has also its downsides. Part of the issue is not only the amount of data that the system will need to collect and use but also how that data is delivered in an effective and actionable manner for the individuals responsible to manage the crisis. To solve this issue and approach crisis as a problem of information flow we took into account two distinct methodologies. These being the 6W model by Kropczynski et al. [9] and the meta-model by Benaben et al. [1].

As social media data processing systems currently can extract information from the data, few of them can leverage that data around the information that it creates. Previously developed systems did not take into account the additional human cost created. While we do not think about removing the human in the loop, the situation can certainly benefit from better and improved ways to handle the information. It is why our article seeks to address the following: *How can we automatically organize the information extracted from social media to better serve decision making in crisis situations?*

As a potential answer, an information model is used in addition to traditional social media processing systems. Yet, as modeling the information is not enough, the present article relies on a meta-model to organize the information resulting from data processing.

## 2. Related Work

The related work section explores the two aspects of the research question. The first one is the automatic extraction of information from social media data during an ongoing emergency event. The second one is the organization of information during crisis events. The first aspect is explored in crisis informatics, through the creation of several social media processing systems to help crisis responders. These social media processing systems will be further explored in the next part, and their functionalities/contributions argued. The second aspect is related to information model engineering to design such social media processing system. It discusses the information of interest that should be extracted and how they relate to each other to improve the situational awareness of crisis stakeholders.

### 2.1. Social media processing systems for crisis response

Many social media processing systems for crisis response have been designed to provide information to decision-makers. The information produced differs from one system to another but can be grouped into 3 main categories:

- Data filtering: tweet related to the crisis, tweet belonging to a category where the tweet was posted by an eye witness of the event [10, 11, 12, 13, 14, 15].
- Semantic enrichment: the spatial (location inference), social (relatives involved) or emotional context of the content [14, 15, 16].
- Clustering of tweets: identification of the most common keywords, aggregation of similar tweets [10, 11, 12, 13, 14, 15, 17].

The first category focuses on filtering the data that comes from social media, specifically Twitter in our case. Previous work [15] has shown that systems that filter automatically information can prove very valuable during emergencies. Additionally, data from social media can provide answers to questions that would be otherwise not possible by traditional means. Zahra et al. [18] studied how to classify these messages from Twitter and how those could be filtered to assess whether the individuals were eyewitnesses or not. Among their findings, they also saw that eyewitnesses tend to use words that relate to their senses.

In a similar vein to our paper presented here and fitting the second category, Zhang et al. [19] proved that it was possible to create a real-time system that allowed for social media data to be collected and filtered at an almost real-time basis. Furthermore, their method allowed for visualizations that allow for further understanding of the data collected. The method created by Zhang et al. allowed the data to be presented within the spatial and temporal context. Other studies such as the one conducted by MacEachren et al. [16], used Twitter data to create a visual analytics tool for crises.

Some of the most modern techniques, that also fit our third category previously mentioned is the AIDR method. Introduced by Irman et al. [12] the system uses machine learning to filter messages from social media to crisis management actors. Moreover, the system is able to successfully classify whether these messages belong to one category or another depending on their content and the context of the crisis. Their approach shows a system that can be altered to fit different requirements and roles depending on the type

of data needed. Other studies [13] have also focused on creating tools that extract and cluster the information according to different parameters and similarities. This approach provides an overview of how different pieces of seemingly unrelated data can be clustered together to create actionable information.

## 2.2. Existing information model for crisis management

To situate this paper in the correct context we need to review what prior research has done in the field of crisis management. Specifically, we need to understand prior research as it pertains to the different ways to manage crises irrespective of whether data from social media was collected or not.

However, grouping and dividing these models into different categories can be a daunting task as models are created for highly specific issues and scenarios in mind, but they all offer a view at different methodologies to manage crises.

We can begin by reviewing knowledge-based models. Knowledge-based models function by grasping the information, or knowledge, to create better crisis management opportunities. Research about this type of crisis management has been abundant [20, 1, 21, 22, 23] and it encompasses many different fields. This type of crisis management response has shown [21] that having a concise knowledge management strategy can help create better solutions and strategies for crisis management. This methodology presented in the previous study pertains to crises that affect public and private companies, which could be argued it might not necessarily translate to other types of crisis. However, Benaben et al. [1] created a meta-model centered around collaborations during crises. Their study shows that knowledge-based meta-models for crisis management would allow for a variety of different crises and not necessarily be relegated to certain situations alone.

Other approaches focus more on creating preventive models that review past data to help decision making for the individuals responsible facing the crisis [24]. This type of model allows for the use of Machine Learning algorithms to help decision-makers and in some situations, it can create realistic scenarios based on previous data to help individuals train [25]. This approach certainly allows for a more flexible and dynamic approach.

## 2.3. Information Summarizing

What this paper aims to accomplish is the creation of an information model that collects the data from social media, refines it, and sends it back to the individuals

responsible for crisis management. To do so, we need to review what previous research has been done in the field of summarizing information for these purposes and how it can help position our research in the field. In a similar vein to crisis management models, the summarizing of information for these purposes has been done in multiple different ways and specific to the research problem tackled.

Having the right information at the right time can be of paramount importance in crisis management [6]. It is why some research has focused specifically on bringing this information in ways that can benefit all the individuals involved. Someren et al. [8] focus its efforts on creating a system that not only provides the right information to the right individuals at the right time but also does it in a way that can be expanded to multiple domains. However, their system still needs the presence of professional input to adjust the model. Other approaches [26] have studied the creation of models in which data is filtered by different processes to create a decision support system that can work in multiple domains while using a base model. When it comes to social media Zahra et al. [18] managed to classify information directly from Twitter, based on whether people were witnesses to a disaster or not. However, a drawback from their study is the fact that this classification and summarizing of the information had to be done manually.

## 3. Social media data processing

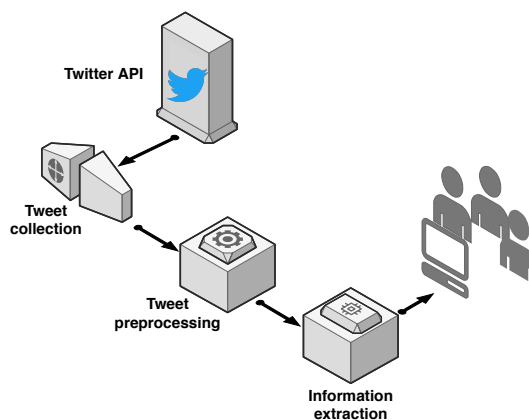


Figure 1. Generic representation of existing social media processing systems.

This section presents the generic structure of the existing social media processing systems, to better contextualize the contributions of this article in the following sections.

Social media are a data source of choice in some

crisis situations. Figure 1 represents the different components that are commonly found in the social media data processing systems mentioned in the Related Works section. Then, the following sub-sections describe the different components represented in the figure.

### 3.1. Twitter API

Most of the research in the field has chosen Twitter as a platform. Twitter is a micro-blogging platform, where each user is allocated a space to post short messages (280 characters) and follow the activity of other users. This platform is mainly used by users to track news<sup>1</sup>. Among other reasons, this choice can be explained as follows: (i) a large number of users (330 million active monthly users worldwide at the time of writing this article<sup>2</sup>), (ii) the ease of access to data through their public API, (iii) the proven sharing of information on this social media at events [5].

### 3.2. Tweet Collection

Twitter offers different methods of collecting the data available on its platform. Most systems are based on the method that allows them to follow in near real-time the messages posted on the platform. It is thus possible to collect data that: contain one of the keywords provided by the user or are posted in a geographical area defined by the user. The keywords used to retrieve the data are crucial to collect relevant data in the case of Twitter. If the keywords are not associated with the current crisis, the collection of relevant data related to the event becomes uncertain. Thus, it is crucial to use relevant keywords when collecting data, and monitor the event to detect emerging trends.

### 3.3. Tweet Pre-processing

Social media data can be different kinds of data. It can be text, photos, or videos. Most of the systems so far were processing text data (the messages posted on Twitter). Text data does not come without its issues, and more often than not, it needs some pre-processing before it can be used. The pre-processing step can be seen as a normalization step of the data. Usually, the pre-processing is composed of:

- Tokenization (how words are split up in the sentence)
- Noise removal (punctuation, hashtags symbol, URLs, "RT" etc.)

<sup>1</sup><https://business.twitter.com/>

<sup>2</sup><https://www.statista.com/>

- Lower casing the text
- Stop-word removal (stop-words are the most common words in a language)

Among all these steps, the most impacted by social media data is the tokenization step. Tokenization consists of splitting up a sequence of characters into smaller pieces called *tokens*. Messages coming from Twitter are unique in their syntax as the platform limits the number of characters to 280. Users, therefore, manage with this constraint by reducing their words, using abbreviations, emoticons, or slang. These particular features of Twitter's messages make most of the existing tokenizers irrelevant. Derczynski et al. [27] highlighted in their article the consequences of using an inadequate tokenizer. Using the generic one, they were only achieving 80% on the F-Measure, while achieving 96% with a tokenizer designed for Twitter's messages.

### 3.4. Information Extraction

Once the collected messages have been pre-processed, each token is associated with a vector representation produced by a language model such as BERT [28] or XLNet [29]. This representation provides then a semantic component to the token allowing the system to actually "understand" the meaning of the token. Thus, the system no longer deals with sequences of cleaned tokens, but with sequences of vectors. These sequences are then passed to a machine learning model which then infer an appropriate label to attribute to the original sentence. Common labels are: crisis-related or not, detected emotions, detected entities, geographical location...

### 3.5. Information Processing

At this stage, the systems provide a result consumed by a user, either an operator or a decision-maker. However, this information from social networks is added to the information obtained from phone calls and reports from the rescue units deployed on-site. This may leave the decision-maker with duplicate or contradictory information, which does not enable decision-making. Also, not all crisis actors have equal access to information. Information can then be fragmented among the different actors, without them necessarily being aware of it. Finally, not all actors involved necessarily use the same vocabulary to describe similar concepts, adding confusion when handling or exchanging information.

To limit the effect of the above-mentioned frictions, the services that can potentially collaborate to conduct joint exercises during the preparation phase. Another

initiative is to work through a common representation of the crisis. Often presented as a map, this representation brings together data under common naming for all actors, thus facilitating the gathering of information [30]. This common map then offers a global vision to all actors. Its interest lies in the graphical representation it provides, facilitating the processing and organization of information by its readers. However, organizing information is still performed by human operators, whereas some events could be automatically linked or new information inferred.

#### 4. Proposal - Moving from data-driven decisions to information-driven

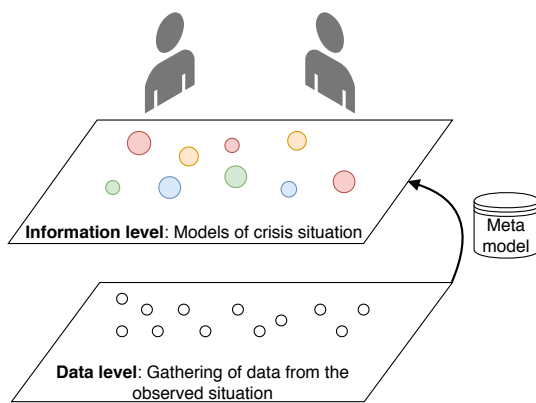


Figure 2. Knowledge management representation and the role of the meta-model.

Moving on now to the contribution of this article: a system that automates the collection and organization of information for crisis management. The contribution is presented through 2 sections: (i) this section presents which information can be extracted from social media data, (ii) the next section presents how the system links the information extracted with the meta-model's classes.

The related works section explored the existing social media processing systems during crises. However, these systems essentially produce information, without considering how they organize or relate to each other. Even today this task has to be carried out by a person.

Yet, information models have been used in several applications to handle a large amount of information carried in large organizations. Among all the models explored, the one proposed by Benaben et al.[1] is interesting as it is in adequacy with the initial problem and is focused on decision making. The following section shows that it is possible to model the information available on social media and that this information has a place in the chosen meta-model.

#### 4.1. Data driven decision making

Currently, decision-makers rely mostly on data made available either by :

- The monitoring of one (or more) operator(s), which is responsible for reading, processing, and linking the information it receives from social media. This approach has the advantage of providing a proper reaction to the incoming information (with the possibility to take initiatives). However, operators are limited regarding the volume of information they can process.
- The aggregation of data by an automated data collection system. This can also include different information such as the most used keywords, the detection of certain subjects (donations, damage...), the location of tweets on a map (if the user turns on his geolocation), or the most representative tweets about certain feelings (anger, fear, joy...).

However, regardless of the option chosen, these systems always occupy human resources, even in the case of an automated data collection system. Human resources time is then consumed in identifying what a particular message refers to, whether it is relevant or not, etc. This is particularly complicated for emergency services as the integration of social media within their organization is recent. Operators, even if they are trained, find themselves navigating through an ocean of data in search of potential information that could be useful later on. The analysis of the flow of information from social media and its organization ultimately relies on an organization that is already struggling in an unknown situation. Yet, when we observe their functioning, these organizations are already accustomed to extracting and organizing information during complex events.

#### 4.2. Moving from data to information

Emergency call centers are already familiar with gathering information from phone calls. They are on the front line for information about the event in progress, its nature, the people involved, the context, etc. In this situation, call centers direct their questions to obtain specific information from the call maker [2]. The authors explain how the operators are looking for answers to what they call the 6W's The 6W's refer to 6 questions: Where, What, Weapons, When, Who, and Why (in this order). From these questions they tried

to extract a coding scheme dedicated to social media content:

- **Where:** Place where help is needed; where an incident occurred; location of victim/suspect/witness; Directional guides/points of reference
- **What:** Type of incident; indications of severity
- **Weapons:** Indication of whether weapons are present and if so, the type of weapons present
- **When:** Time the incident occurred; an indication of whether an incident is still occurring; time lapse
- **Who:** Number of people injured, or number of people armed; Description of victim/perpetrator/vehicle; nature of observer (firsthand vs. secondhand account)
- **Why:** Chain of events that led to the emergency; Rationale for the occurrence that may be helpful to an investigation

This schema then makes it possible to better identify the relevant information. Later, [9] were interested in the distribution of these 6W's in existing data sets collected during multiple crises. The authors thus concluded that, in this particular data set, an important part of the tweets contained information related to the 6W's, except for the Weapon and When questions (Table 1). Assuming

**Table 1. Percent of Tweets Labeled by 6W's**

W Label	Percent
Where	78%
What	88%
Weapons	3%
When	5%
Who	68%
Why	72%

that this distribution generalizes, social media, therefore, do contain a significant amount of information relevant to emergency services, and this information can be clustered into certain categories. Also, these categories are in some cases similar to the concepts represented by some of the classes of meta-models explored above.

### 4.3. Information management through a meta-model

The previous section demonstrated that it is possible to model the data provided by social media. This aggregation is also proving to be useful for emergency

services. However, the concepts mentioned are not organized and cannot interact with each other. In other words, the previous section allows data to be modeled, but there is no structure around it. Also, it is the role of a meta-model to provide this structure that is missing here.

In section 2.2 we presented different meta-models that adopt different viewpoints of crisis management. The meta-model retained here is the meta-model focused on collaboration between the different actors of [1] (Figure 3). This meta-model allows the representation and use of information that enables the different actors to collaborate during the crisis response. The meta-model is composed of the following packages :

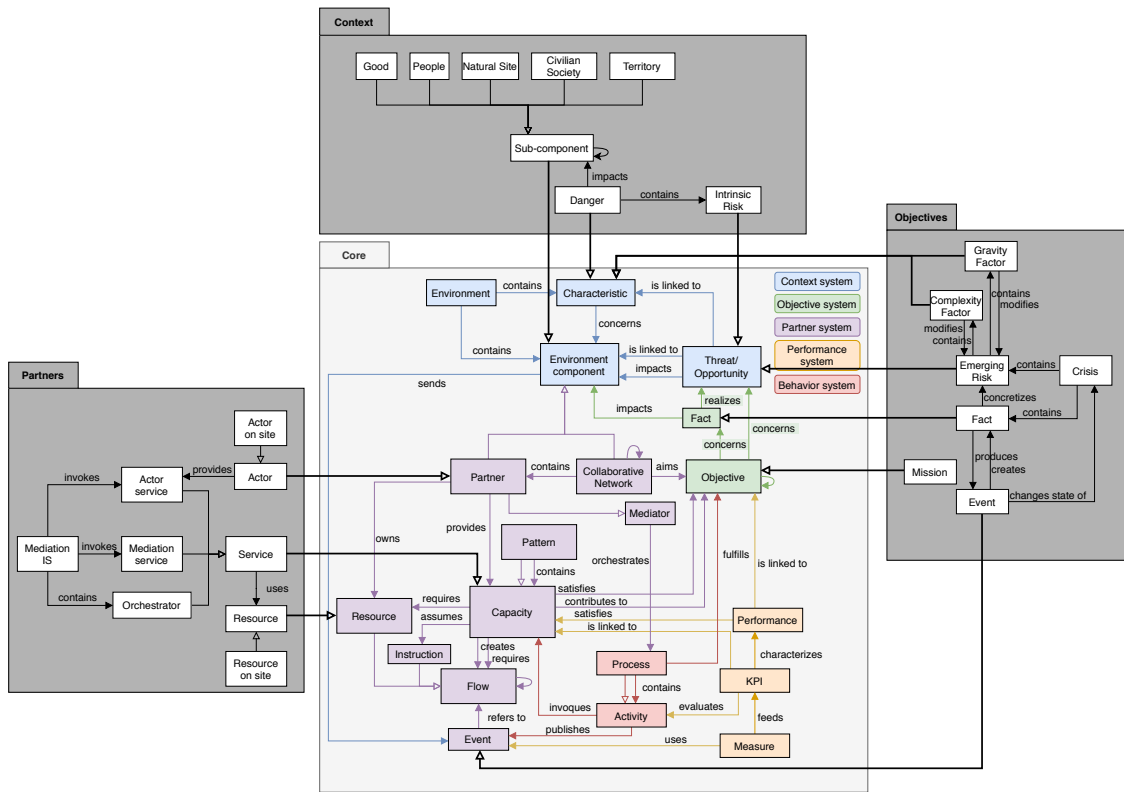
- The Core package allows us to represent collaborative situations by modeling the context, objectives, partners, performance evaluation, and actions.
- The Context package us allows to represent crisis-related context concepts. It includes classes representing the people impacted, administrative areas, natural sites, etc. But also the risks and dangers linked to the environment where the crisis is taking place.
- The Partners package allows to represent the different actors involved, their resources, their capacities, and how these are orchestrated.
- The Objectives package contains the characteristics of the collaboration
- The Behavior package describes the sequence of operations according to the chosen framework (not shown)

Among all the classes that are used on the meta-model, not all the classes can be instantiated using social media data only. However, using the 6W's and the resulting coding scheme proposed by [9], it appears that several classes correspond to some classes. Thus, the classes that the system aims to instantiate are then:

- Sub-component (Context - Where)
- Danger (Context - What)
- Actor (Partners - Who)
- Event (Objectives - Why)

This part has shown that it is possible to model the existing information during an emergency and that meta-models exist to organize the information in a





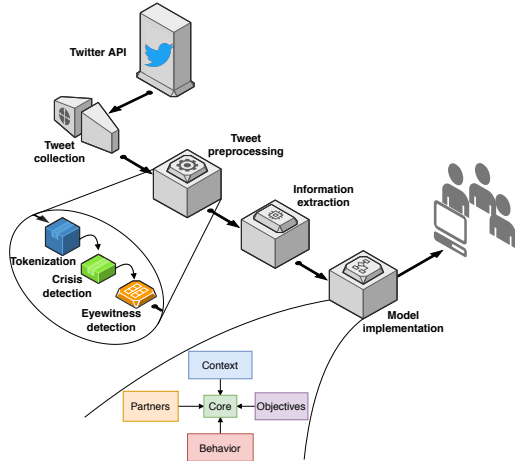
**Figure 3. The crisis meta-model centered around the collaboration of the different actors during an emergency event.**

way that delivers value to the decision-makers. Yet one important question remains: How to implement automatically this model? The next section proposes a potential answer to this question.

### 5. Proposal - Model implementation

The implementation of crisis situation models is based on the data available at the time of the event. For this purpose, it is possible to use different data sources, such as phone calls, the information provided by the units deployed on the ground, or news. However, none of these channels allow for automated implementation and therefore do not implement truly viable crisis models. On the other hand, social media data are already in a digital format and are therefore easier processed by a computer without human interaction to input the data. The implementation of the meta-model classes can then be achieved with the direct use of the data (the message of a tweet for example) or the metadata of the message. These metadata include:

- The time the message was sent
- The number of reactions to the message



**Figure 4. Social media processing systems proposed to implement the meta-model from [1]**

- Information on the user who sent the message (giving access to his social network)
- The geographical position of the user, according to the parameters he defined.

While metadata are easy to process and interpret, they



are also insufficient to instantiate the model. For example, they do not allow us to understand the nature of the event or the people involved in the action described. For this reason, the information contained in the message itself is used to extract more information.

To implement the crisis situation model based on social media data, the system shown in Figure 4 is used. It is largely inspired by existing systems (Figure 1) presented in section 3. The most notable differences are as follows:

- The *pre-processing* step benefits from previous work on the classification of tweets. It is composed of: (i) a module that determines whether the message evokes an event of interest, and (ii) a module that determines whether the message was posted by an eyewitness to the event. The addition of these two modules makes it possible to filter out some of the noise and ensure that the final information is at a minimum truthful.
- The *information extraction* step identifies the classes of the meta-model that are potentially present in the tweet.
- The *meta-model implementation* step creates and organizes the instances of the meta-model classes.

If the steps of pre-processing (filtering of relevant messages) by tweets classification systems [12, 18] and creation/organization of instances of the meta-model [1] are already explored, the link between these two systems is at the heart of our contribution.

The previous section showed that it is possible to find data that reflects the meta-model classes in the social media stream. These data can therefore be used to instantiate it. However, there is a need for a way to detect and match the data we are interested in with their respective classes. This bridge between the data and the model is at the heart of our contribution.

The approach proposed here is to look for the presence of keywords in the messages. These keywords refer to the classes that are being instantiated. If the keyword is present in the message, then the class corresponding to the keyword is instantiated. Unlike traditional approaches to social media data processing, this approach does not focus on the message as a whole, to provide a label, but on the different words that compose the message. However, there are some challenges, the most notable of which are the following:

- The keyword list has to be exhaustive
- This method ignores the context in which the keyword is used

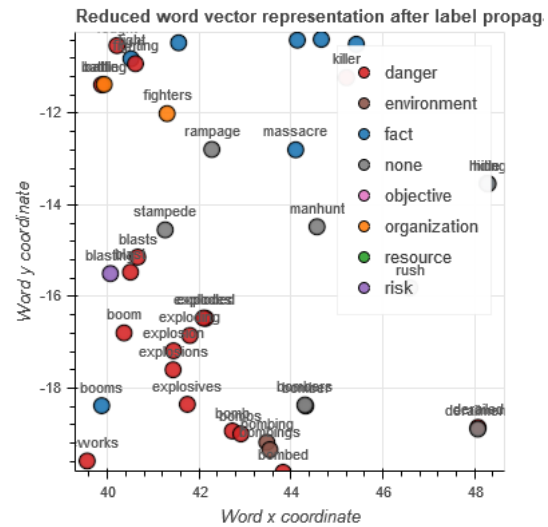


Figure 5. Results obtained from [31] on the propagation of associations in the vector space representing a vocabulary.

Coche et al. [31] presents a method to answer the initial problem while addressing the first challenge. To do this, the authors use two vocabularies: a short set of keywords associated with a concept and a much larger one without any association. All words are associated with vector representations. These representations form a vector space in which semantically close words are close in terms of distance. Then, each word that is semantically close to a word associated with a concept receives the association by propagation. Figure 5 shows the result of the propagation of labels around the "source word of origin". Thus, the list of keywords associated with MM concepts is no longer only composed of the keywords having an association, but also of all the synonyms of these keywords or semantically close words. In their experiments, the authors went from 154 words with an association to a concept, to 882, without losing the quality of associations.

This solution allows linking relevant entities contained in text messages to the meta-model classes. Once the instances are generated, an extra step consists of identifying the relationships between the different instances. This will allow us to identify more easily the different ongoing sub-events of the crisis that are asking for a response from the emergency services. From a broader perspective, one can also imagine completing the resulting model with data coming from other sources such as sensors [32] to provide a broader representation of the event. With a sufficiently exhaustive representation of the crisis, the inference of events or the suggestion of optimized decisions then

become possible.

## 6. Conclusion

### 6.1. Summary

The particular context of handling social media data in crises is unique. The dislocation of the organization and the characteristic lack of information defaults many existing practices. Among these practices are decision support systems, which, although widely adopted today and in daily use, appear to be out of place in unforeseen situations. However, recent research in crisis computing has raised hopes that these systems will be used.

This article explored the problem of automatically organizing the information extracted from social media to better serve decision making in crisis situations. First, the introduction defined the terms crisis and social media, and how both interact during an event. From this point, the related works section presented in the first part the existing social media processing systems already proposed and which kind of information they are providing. In the second part, we looked at previous attempts to organize information during emergency events. The following section then introduced the reader to the collection of data and the challenges that appear in doing so. We then justified the existence of relevant data on social media that can be modeled as useful information. Moreover, this information matches the classes used in the chosen meta-model. Finally, the last section presents the contribution, namely, a mean to match entities contained in messages posted on social media with the classes of the meta-model.

### 6.2. Future work

This article explored one possible answer to this problem. Yet, the proposed solution has some limitations that need to be addressed in future work. First, since the system is based on a method of matching keywords in a sentence, it is not designed to capture the context in which the keyword is used. A keyword could then be used in a context that could not be relevant.

A possible solution would be to train a neural network model to detect the relevant entities contained in sentences. Such models can be trained to detect and understand the semantic of a whole sentence. However, this solution would require a large amount of labeled data to be truly effective.

As a final word, it is interesting to point out that despite tremendous efforts from the scientific community to develop decision support systems for crisis response, only a few ideas made their way to crisis cells. It might then be interesting to explore:

- Did the developed systems really answered the needs of crisis management organizations?
- What are the feedbacks of organizations using similar systems?
- What prevents the adoption of the features explored by academics?

## 7. Acknowledgments

This research is partially supported by the French National Research Agency - ANR, through the MACIV project (Projet-ANR-17-CE39-0015), NSF BIGDATA grant 1741370. We also wish to thank our anonymous reviewers for their constructive comments, which helped improve the presentation of this paper.

## References

- [1] F. Bénaben, M. Lauras, S. Truptil, and N. Salatgé, "A Metamodel for Knowledge Management in Crisis Management," in *2016 49th Hawaii International Conference on System Sciences (HICSS)*, pp. 126–135, Jan. 2016. ISSN: 1530-1605.
- [2] J. Kropczynski, J. Coche, E. Obeysekare, F. Bénaben, R. Grace, S. Halse, A. Montarnal, and A. Tapia, "Identifying Actionable Information on Social Media for Emergency Dispatch," *Proceedings of the ISCRAM Asia Pacific*, p. 11, 2018.
- [3] T. Sakaki, M. Okazaki, and Y. Matsuo, "Earthquake shakes Twitter users: real-time event detection by social sensors," *WWW '10: Proceedings of the 19th international conference on World wide web*, p. 851, 2010.
- [4] T. Sakaki, M. Okazaki, and Y. Matsuo, "Tweet Analysis for Real-Time Event Detection and Earthquake Reporting System Development," *IEEE Transactions on Knowledge and Data Engineering*, vol. 25, pp. 919–931, Apr. 2013.
- [5] J. I. Blanford, J. Bernhardt, A. Savelyev, G. Wong-Parodi, A. M. Carleton, D. W. Tittley, and A. M. MacEachren, "Tweeting and tornadoes.," in *ISCRAM*, 2014.
- [6] P. Lagadec, "La gestion des crises," *Outils de réflexion al'usage des décideurs*, Ediscience International, Paris, 1991.
- [7] L. G. Militello, E. S. Patterson, L. Bowman, and R. Wears, "Information flow during crisis management: challenges to coordination in the emergency operations center," *Cognition, Technology & Work*, vol. 9, no. 1, pp. 25–31, 2007.
- [8] M. van Someren, N. Netten, V. Evers, H. Cramer, R. de Hoog, G. Bruinsma, *et al.*, "A trainable information distribution system to support crisis management," in *2nd International Conference on Information Systems for Crisis Response and Management (ISCRAM)*, 2005.
- [9] J. Kropczynski, S. Halse, R. Grace, N. Elrod, D. Caragea, C. Caragea, and A. Tapia, "Refining a Coding Scheme to Identify Actionable Information on Social Media," *Proceedings of the 16th International Conference on Information Systems for Crisis Response and Management*, p. 7, 2019.

- [10] C. Caragea, N. McNeese, A. Jaiswal, G. Traylor, H.-W. Kim, P. Mitra, D. Wu, A. H. Tapia, L. Giles, B. J. Jansen, and J. Yen, "Classifying Text Messages for the Haiti Earthquake," *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management*, p. 10, 2011.
- [11] M. A. Cameron, R. Power, B. Robinson, and J. Yin, "Emergency situation awareness from twitter for crisis management," *21st International Conference Companion on World Wide Web - WWW '12 Companion*, no. July 2015, p. 695, 2012.
- [12] M. Imran, C. Castillo, J. Lucas, P. Meier, and S. Vieweg, "AIDR: artificial intelligence for disaster response," in *Proceedings of the 23rd International Conference on World Wide Web - WWW '14 Companion*, (Seoul, Korea), pp. 159–162, ACM Press, 2014.
- [13] Z. Ashktorab, C. Brown, M. Nandi, and A. Culotta, "Tweedr: Mining Twitter to Inform," *Proceedings of the 11th International Conference on Information Systems for Crisis Response and Management*, p. 5, 2014.
- [14] H. Purohit and A. Sheth, "Twitris v3: From Citizen Sensing to Analysis, Coordination and Action," *Seventh International AAAI Conference on Weblogs and Social Media*, p. 2, 2013.
- [15] F. Abel, C. Hauff, G.-J. Houben, R. Stronkman, and K. Tao, "Semantics + Filtering + Search = Twitcident. Exploring Information in Social Web Streams," in *Proceedings of the 23rd ACM Conference on Hypertext and Social Media*, HT '12, (New York, NY, USA), pp. 285–294, ACM, 2012. event-place: Milwaukee, Wisconsin, USA.
- [16] A. M. MacEachren, A. Jaiswal, A. C. Robinson, S. Pezanowski, A. Savelyev, P. Mitra, X. Zhang, and J. Blanford, "SensePlace2: GeoTwitter analytics support for situational awareness," in *2011 IEEE Conference on Visual Analytics Science and Technology (VAST)*, (Providence, RI, USA), pp. 181–190, IEEE, Oct. 2011.
- [17] J. Rogstadius, M. Vukovic, C. A. Teixeira, V. Kostakos, E. Karapanos, and J. A. Laredo, "CrisisTracker: Crowdsourced social media curation for disaster awareness," *IBM Journal of Research and Development*, vol. 57, pp. 4:1–4:13, Sept. 2013.
- [18] K. Zahra, M. Imran, and F. O. Ostermann, "Automatic identification of eyewitness messages on twitter during disasters," *Information processing & management*, vol. 57, no. 1, p. 102107, 2020.
- [19] X. Zhang, S. Kelly, and K. Ahmad, "The slandail monitor: Real-time processing and visualisation of social media data for emergency management," in *2016 11th International Conference on Availability, Reliability and Security (ARES)*, pp. 786–791, IEEE, 2016.
- [20] Z. Jia, Y. Shi, Y. Jia, and D. Li, "A framework of knowledge management systems for tourism crisis management," *Procedia Engineering*, vol. 29, pp. 138–143, 2012.
- [21] W.-T. Wang and S. Belardo, "Strategic integration: A knowledge management approach to crisis management," in *Proceedings of the 38th Annual Hawaii International Conference on System Sciences*, pp. 252a–252a, IEEE, 2005.
- [22] D. Blackman, M. Kennedy, and B. Ritchie, "Knowledge management: the missing link in dmo crisis management?," *Current Issues in Tourism*, vol. 14, no. 4, pp. 337–354, 2011.
- [23] P. Racherla and C. Hu, "A framework for knowledge-based crisis management in the hospitality and tourism industry," *Cornell Hospitality Quarterly*, vol. 50, no. 4, pp. 561–577, 2009.
- [24] S. Bennani, A. Maalel, H. B. Ghézala, and M. Abed, "Towards a decision support model for the resolution of episodic problems based on ontology and case bases reasoning: application to terrorism attacks," in *2017 IEEE/ACS 14th International Conference on Computer Systems and Applications (AICCSA)*, pp. 1502–1509, IEEE, 2017.
- [25] E. Grois, W. H. Hsu, M. Voloshin, and D. C. Wilkins, "Bayesian network models for generation of crisis management training scenarios," in *AAAI/IAAI*, pp. 1113–1120, 1998.
- [26] Y. Han and W. Xu, "An ontology-oriented decision support system for emergency management based on information fusion," in *Proceedings of the 1st ACM SIGSPATIAL International Workshop on the Use of GIS in Emergency Management*, pp. 1–8, 2015.
- [27] L. Derczynski, A. Ritter, S. Clark, and K. Bontcheva, "Twitter part-of-speech tagging for all: Overcoming sparse and noisy data," in *Proceedings of the International Conference Recent Advances in Natural Language Processing RANLP 2013*, pp. 198–206, 2013.
- [28] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding," *arXiv:1810.04805 [cs]*, Oct. 2018. arXiv: 1810.04805.
- [29] Z. Yang, Z. Dai, Y. Yang, J. Carbonell, R. R. Salakhutdinov, and Q. V. Le, "XLNet: Generalized Autoregressive Pretraining for Language Understanding," in *Advances in Neural Information Processing Systems 32* (H. Wallach, H. Larochelle, A. Beygelzimer, F. d. Alché-Buc, E. Fox, and R. Garnett, eds.), pp. 5754–5764, Curran Associates, Inc., 2019.
- [30] J. Coche, A. Montarnal, A. Tapia, and F. Benaben, "Actionable Collaborative Common Operational Picture in Crisis Situation: A Comprehensive Architecture Powered with Social Media Data," in *Collaborative Networks and Digital Transformation* (L. M. Camarinha-Matos, H. Afsarmanesh, and D. Antonelli, eds.), IFIP Advances in Information and Communication Technology, (Cham), pp. 151–162, Springer International Publishing, 2019.
- [31] J. Coche, A. Tapia, A. Montarnal, and F. Benaben, "Automatic Information Retrieval from Tweets: A Semantic Clustering Approach," *Proceedings of the 17th International Conference on Information Systems for Crisis Response and Management*, pp. 134–141, 2020.
- [32] A. Fertier, A. Montarnal, A.-M. Barthe-Delanoë, S. Truptil, and F. Bénaben, "Real-time data exploitation supported by model- and event-driven architecture to enhance situation awareness, application to crisis management," *Enterprise Information Systems*, vol. 14, pp. 769–796, July 2020.